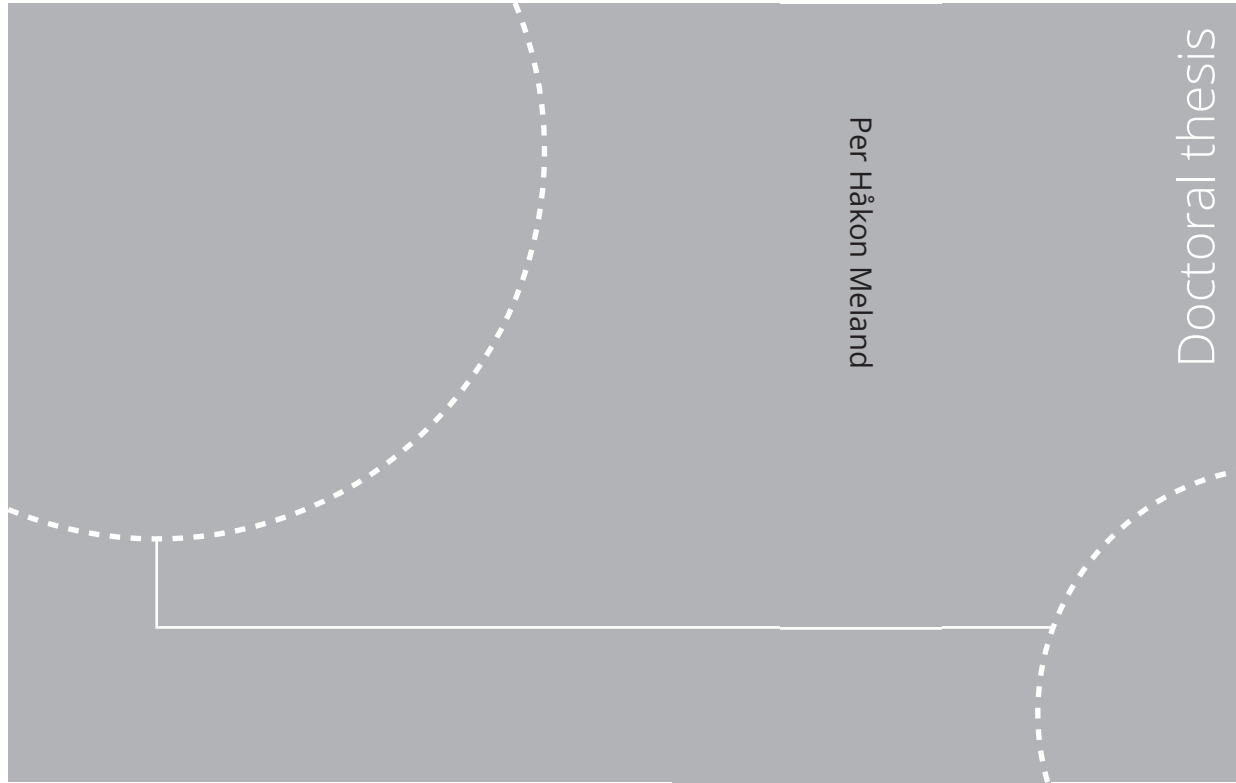


Doctoral theses at NTNU, 2021:329

Per Håkon Meland

# Storyless cyber security

Modelling threats with economic incentives



Doctoral theses at NTNU, 2021:329

**NTNU**  
Norwegian University of  
Science and Technology  
Thesis for the degree of  
Philosophiae Doctor  
Faculty of Information Technology  
and Electrical Engineering  
Department of Computer Science

 NTNU

 **NTNU**  
Norwegian University of  
Science and Technology

 **NTNU**  
Norwegian University of  
Science and Technology

ISBN 978-82-326-5412-3 (printed ver.)  
ISBN 978-82-326-6362-0 (electronic ver.)  
ISSN 1503-8181 (printed ver.)  
ISSN 2703-8084 (electronic ver.)

Per Håkon Meland

# Storyless cyber security

Modelling threats with economic incentives

Thesis for the degree of Philosophiae Doctor

Trondheim, October 2021

Norwegian University of Science and Technology  
Faculty of Information Technology  
and Electrical Engineering  
Department of Computer Science



Norwegian University of  
Science and Technology

**NTNU**

Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Information Technology  
and Electrical Engineering  
Department of Computer Science

© Per Håkon Meland

ISBN 978-82-326-5412-3 (printed ver.)  
ISBN 978-82-326-6362-0 (electronic ver.)  
ISSN 1503-8181 (printed ver.)  
ISSN 2703-8084 (electronic ver.)

Doctoral theses at NTNU, 2021:329



Printed by Skipnes Kommunikasjon AS

## DEDICATION

*For my family,  
four, twenty-four, twenty-two,  
for my friends,  
for your health.*



## ABSTRACT

Cyber risk management is about identifying, assessing and reducing risk to an acceptable level. With systems that have been in operation for some time, we might be able to make qualified risk estimations and treat them in a cost-efficient manner based on the previous events and experiences. However, with storyless systems, such estimations become more of a guesswork and it is hard to determine how much and what kind of security is good enough. Additionally, both old and new systems are exposed to an evolving threat environment where relying on the Maginot lines of the past could lead to brutal consequences in the future.

The purpose of this PhD study has been to investigate new methods for managing cyber security risks without too much reliance on historical events. These methods belong to an area found in the intersection between threat modelling and security economics. The former is about anticipating attacks and imagining what can go wrong, often taking the mindset of an adversary. The latter is concerned about how economic mechanisms shape security.

The overall research approach of the study leans towards practice-based research, where interventions and designs contribute to local practices as well as generalized knowledge. Following the principles of pragmatism, a mix of quantitative and qualitative research methods have been applied for empirical inquiry, covering problem investigation, artefact creation and evaluation. The study has complemented ongoing projects that are addressing threats and technology development within the aviation and maritime fields, and included cyber insurance as an application area for risk transfer to third parties. A general limitation is the assumed rational behaviour of both attackers and defenders, which do not cover all types of cyber threats. Furthermore, there are ethical concerns restricting the research methods and openness of results related to cyber crime investigations.

The results have been published as a collection of papers and show that subjective estimations can be supported by economic incentives when identifying threats, the likelihood of their occurrence and ways of treating them. For instance, by focusing on the capabilities that are needed for the different attack stages, we can spend less time and obtain a higher degree of reusability compared to modelling specific attack paths. Just as there is no one-solution-fits-all for threat modelling, we cannot use data types and sources for economic incentives uncritically.

We have documented some of these strengths and weaknesses related to a given set of threats, and encourage to expand this work to support the cyber risk management discipline.

## PREFACE

This thesis is submitted to the *Norwegian University of Science and Technology* (NTNU) for the partial fulfilment of the requirements for the degree of *Philosophiae Doctor* (PhD).

The grant from the Research Council of Norway is of type *Institute PhD* (STIP-INST), which is a new type of scholarship in Norway. The purpose for this arrangement is to strengthen the role of the technical-industrial research institutes in the Norwegian PhD education, and to ensure industry relevance by connecting the work to established applied research communities. The grant belongs to project number 259869.

This doctoral work has been conducted at the independent research organisation SINTEF Digital, Department for *Software Engineering, Safety and Security* (SESS) and NTNU, Faculty of *Information Technology and Electrical Engineering* (IE), Department of Computer Science (IDI). The work has been performed under the supervision of Professor Guttorm Sindre. Professor Letizia Jaccheri and Associate Professor Karin Bernsmed were assigned as co-supervisors.





## ACKNOWLEDGEMENT

My initial thanks go to my closest family. To my parents, Karen and Torbjørn, who put food and computers on my table ever since I was a little boy. To my brother, Erlend, who set the standard in education. To the missis, Mari, for tolerating me. To my pack of children, Edvard, Emil and Nina, for keeping the noise down while I was writing this thesis. To the rest of my family and friends, thank you for your support and encouragement in the past few years.

Many thanks to Professor Guttorm Sindre, my supervisor at NTNU, who's always excited and made time available for discussions and feedback throughout the PhD period. I've also appreciated the support of my co-supervisors Professor Letizia Jaccheri and Associate Professor Karin Bernsmed.

I would like to thank all co-authors of the publications contributing to this thesis. I feel I've been lucky to have had the opportunity to work with and learn from so many esteemed people from all over the world. I hope we will be able to continue collaborating in the future.

Thank you to the trinity consisting of the Research Council of Norway, my employer SINTEF and NTNU for providing funds, courses, administration and taking care of all the various practicalities needed for this kind of PhD work. My colleagues and fellow PhD students have ensured a superb working environment and enlightening discussions both related to this research discipline and other aspects of life.

I'm grateful for the opportunity I've had to combine my PhD work with several research and development projects, especially CySiMS, CySiMS-SE, CyberSec4Europe, Iris Precursor, Iris Service Evolution Study and SESAR PJ05. I would like to thank all professionals involved in these projects for allowing this research to be integrated into a practical context. Thank you to the STERNA project, led by Dr. Ivonne Herrera and Associate Professor Tarcisio Saurim, for teaching me about resilience engineering, inviting me to visit the *Universidade Federal do Rio Grande do Sul* (UFRGS), *Pontifícia Universidade Católica do Rio Grande do Sul* (PUCRS) and *Universidade Federal da Bahia* (UFBA), and allowing me to teach about cyber resiliency.

Special thanks go to Nanna Unhammer at Willis Towers Watson and Jeff Cohen at Advisen (a Zywave company) for providing cyber insurance data for me to work with and access to customers. Also, I really appreciate the many informal discussions on cyber insurance challenges

I've had with Thomas Schnitzer and colleagues at Swiss Re.

Many thanks to Dr. Rune Storesund and Professor Emeritus Karlene Roberts at the *Center for Catastrophic Risk Management (CCRM)* for sponsoring my stay as a visiting scholar at UC Berkeley. Also, thank you to Ann Cleaveland and Matthew Nagamine at the *Center for Long-Term Cybersecurity (CLTC)* for letting me join their office space there. I only wish it would have been possible to stay for longer.

There have been many students at NTNU and INSA de Rennes (supervised by maître de conférences Barbara Fila) that have supported me with tool development and complementary research. You have all done a fantabulous job and we have learned a lot from each other. I wish you great success in your careers and personal lives.

Lastly, I have to send a NACK to the population of bats in the Hubei province in the People's Republic of China. In the future, please take your vitamins and try to avoid getting eaten by humans.

# CONTENTS

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	The title explained . . . . .	3
1.2	Research approach overview . . . . .	3
1.3	Overview of papers . . . . .	4
1.4	Structure of the thesis . . . . .	8
<b>2</b>	<b>Area of concern</b>	<b>9</b>
2.1	Threat modelling . . . . .	13
2.2	Security economics . . . . .	16
2.3	Limitations . . . . .	20
<b>3</b>	<b>Problem setting</b>	<b>21</b>
3.1	Aviation . . . . .	24
3.2	Maritime . . . . .	28
3.3	Cyber insurance . . . . .	33
<b>4</b>	<b>Research questions</b>	<b>37</b>
4.1	Main research question . . . . .	37
4.2	Sub-questions . . . . .	37
4.3	Paper-specific questions . . . . .	39
<b>5</b>	<b>Conceptual framework</b>	<b>41</b>
5.1	Research paradigms . . . . .	41
5.2	From theory to practice . . . . .	43
<b>6</b>	<b>Methods of empirical inquiry</b>	<b>45</b>
6.1	Problem investigation methods . . . . .	48
6.2	Artefact creation . . . . .	49
6.3	Evaluation methods . . . . .	50
6.4	Mapping of research methods . . . . .	54
<b>7</b>	<b>Contributions</b>	<b>55</b>
7.1	The use of the Contributor Roles Taxonomy . . . . .	55
7.2	Primary papers contribution . . . . .	56
7.3	Secondary papers contribution . . . . .	73
7.4	Posters contribution . . . . .	77
<b>8</b>	<b>Discussion</b>	<b>79</b>

8.1	Addressing the research questions . . . . .	79
8.2	Ethical issues . . . . .	98
8.3	Future opportunities and recommendations . . . . .	101
<b>9</b>	<b>Conclusion</b>	<b>103</b>
	<b>Bibliography</b>	<b>107</b>
	<b>Index</b>	<b>139</b>
<b>A</b>	<b>Primary papers</b>	<b>145</b>
A:	‘Attribute decoration of attack–defense trees’ . . . . .	147
B:	‘Mitigating risk with cyberinsurance’ . . . . .	184
C:	‘Visualizing cyber security risks with bow-tie diagrams’ . . . . .	195
D:	‘Facing uncertainty in cyber insurance policies’ . . . . .	217
E:	‘When to treat security risks with cyber insurance’ . . . . .	231
F:	‘An experimental evaluation of bow-tie analysis for security’ . . . . .	255
G:	‘Demand side expectations of cyber insurance’ . . . . .	283
H:	‘An Experimental Analysis of Cryptojacking Attacks’ . . . . .	293
I:	‘Cyber Attacks for Sale’ . . . . .	311
J:	‘The Ransomware-as-a-Service economy within the darknet’ . . . . .	319
K:	‘Breaking the cyber kill chain by modelling resource costs’ . . . . .	330
L:	‘A Systematic Mapping Study on Cyber Security Indicator Data’ . . . . .	347
M:	‘Assessing cyber threats for storyless systems’ . . . . .	375
<b>B</b>	<b>Posters</b>	<b>401</b>
<b>C</b>	<b>Awards</b>	<b>405</b>

## LIST OF FIGURES

1.1	The modern Panzer II (left) and Panzer I mobile units passing through Ardenne Forest in 1940. Photo: Bundesarchiv [5] (CC-BY-SA 3.0) . . . . .	2
1.2	Generic structure of engaged scholarship. Adapted from Mathiassen [15] . . . . .	4
2.1	The cybermen. Photo used with permission from [51] . . . . .	10
2.2	A domain model of central terms . . . . .	11
2.3	The set of activities involved in risk management. Adapted from ISO/IEC 27005 [57]	12
2.4	The intersection of threat modelling and security economics . . . . .	13
2.5	An attack prediction and forecasting methods taxonomy. Adapted from Husák et al. [101] . . . . .	15
2.6	The security of the fish depends on the fishermen and their potential profit . . . . .	18
2.7	A sample of advertisements found on a darknet marketplace . . . . .	19
3.1	The memorial built after the attacks on the World Trade Center towers. Photo by Meland . . . . .	24
3.2	The problem areas for aviation are focused on air-ground communication and remote tower operations . . . . .	26
3.3	Trial simulating remote air traffic control. Photo by Meland with the courtesy of SAAB and LfV . . . . .	27
3.4	The concentration of intentional maritime cyber threats from the last decade. Data from Meland et al. [43] . . . . .	29
3.5	The problem areas for maritime are focused on ship-ship and ship-shore communication	30
3.6	One of the ships tested with the CySiMS PKI solution and VDES radio. Photo courtesy of Kongsberg Seatex . . . . .	31
3.7	Different roles involved in cyber insurance . . . . .	34
4.1	Hierarchy of research questions . . . . .	39
5.1	A structural anatomy model for practice research. Adapted from Goldkuhl [248] . . . . .	44
6.1	Research context and potential contribution. Adapted from Gregor and Hevner [254]	50
6.2	Screenshot from the bow-tie modelling software artefact. . . . .	51
6.3	Benchmarking results of the time (years) it takes to mine a single Monero coin on different systems . . . . .	53
7.1	A high-level overview of the relationships between the primary papers contribution	58
7.2	A high-level overview of the secondary papers contribution . . . . .	74

8.1	A spaghetti-diagram showing how the primary papers relate to the research questions	81
8.2	Second-hand price for ECDIS software obtained from eBay.com . . . . .	89
8.3	A value chain for the RaaS economy as presented in paper J . . . . .	92

## LIST OF TABLES

4.1	Sub-questions and their type . . . . .	38
5.1	Mapping towards methodological pragmatism (MP) . . . . .	43
6.1	Why qualitative research was needed . . . . .	46
6.2	Why quantitative research was needed . . . . .	47
6.3	Mapping between research methods and primary papers . . . . .	54
7.1	Summary of secondary papers . . . . .	75
7.2	Summary of posters . . . . .	78
8.1	Economic incentive data used in primary papers . . . . .	86
8.2	Economic incentives and threat modelling . . . . .	93





## INTRODUCTION

*Everybody tries to pass the buck*

---

Ross Anderson [1]

Prior to the Second World War, France invested heavily in building a wall of strong fortifications towards the German border known as the *Maginot Line* [2]. It featured state-of-the-art defence capabilities, but these proved costly to maintain and led to underfunding elsewhere. Furthermore, the fortifications did not cover the Ardennes Forest, which was considered impenetrable due to the difficult terrain and the fact that no invading army had used that route in the past. Meanwhile, the Germans developed new attack capabilities based on aerial warfare and highly mobile armoured units. During May 1940, these mobile units swiftly passed through the Ardenne Forest with support from the skies above (Figure 1.1), and France was compromised. Since then, the Maginot Line has become a well-known metaphor for expensive efforts that offer a false sense of security [3, 4].

Today, large organisations' spending on cyber security are increasing twice the rate compared to all other types of information technology [6], and at the same time, the number of security incidents just continue to increase. According to Paté-Cornell et al. [7], there is little information about the effectiveness of adopted security interventions and the priorities among them. This is supported by Woods and Böhme [8], who show that research has inconsistently demonstrated how such interventions reduce risk. This is not sustainable in the long run, and we need to avoid as many poor security investments as possible.



**Figure 1.1:** The modern Panzer II (left) and Panzer I mobile units passing through Ardenne Forest in 1940. Photo: Bundesarchiv [5] (CC-BY-SA 3.0)

An inherent challenge with cyber security is the lesser relevance of retrospective analysis compared to other fields that deal with risks, such as safety, finance and insurance. This is arguably due to facts such as:

- Cyberattacks represent a relatively new phenomenon, there is a lack of historical data, and organisations are not eager to share information about incidents.
- Technology and threats increase and develop so fast that historical data become irrelevant even after a relatively short time. In 2006, the security company McAfee Labs counted an average of 25 threats a day. In 2016, at the start of my PhD study, the number was more than 300 threats per minute [9]. By the end 2020, the number had risen to 419 threats per minute [10].
- The likelihood of cyberattacks is hard to predict accurately since these are not randomly triggered unfortunate events, but rather depending on issues such as attacker motivation and capabilities, which are outside the control of an organisation.

To overcome these challenges, there is a need for better methods for quantifying cyber security risks, so that more informed decisions can be made for security investments. To quote Denning in [11]: “Security is a bottomless pit; you can only do so much. But it’s important to do the right things - the things that will make a difference.”

The goal of this PhD study has been to address the following main research question: *How can modelling threats and economic incentives improve cyber risk management?*

There is a need to accept the general unavailability of reliable historical data, and instead build on data about the present to project the future. Identifying reliable data sources and applying these in models for attacker and defender costs will be of benefit when estimating likelihood and consequence of unwanted cyber security events. The PhD study has complemented ongoing research projects that are addressing threats and technology development within the aviation and maritime fields, and included cyber insurance as an application area for risk transfer to third parties.

## 1.1 The title explained

The title of this thesis introduces a new term, *storyless cyber security*. The word “storyless” is by Merriam-Webster dictionary defined as “being without a story” [12], meaning there is no (his-)story or records related to the following noun, which is “cyber security”. The title represents situations where one cannot rely on the past to know the present or future. This is in particular relevant with new technologies, threats or application domains. So, instead of driving forward by looking in the rear-view mirror, there is a need to look for obstacles ahead, anticipate dangers around the corner and be prepared to take evasive manoeuvres.

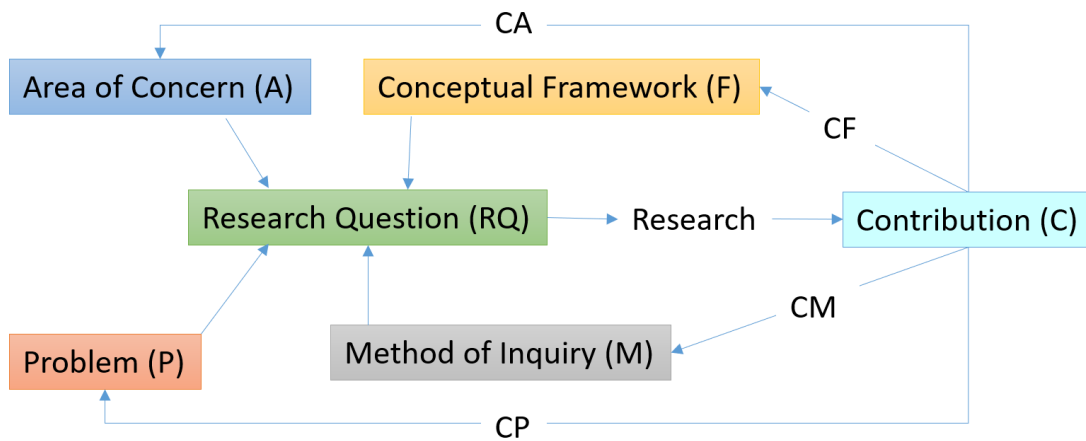
The subtitle, *modelling threats with economic incentives* refers to the idea of analysing cyber threats in a new way, taking positive and negative incentives into account. This does not replace existing threat modelling methodologies, but should be seen as an additional tool to reduce uncertainty in such processes.

## 1.2 Research approach overview

This institute PhD work follows Costley and Fulton’s [13] principles of *practice-based research*, which are aimed at professional doctorate candidates. Practice researchers “are often, but by no means exclusively, mid-career professionals, coming to the research with a wealth of experience and a variety of projects already completed.” Furthermore, the researcher is more than an insider but inside the research and knowledge is generated through addressing problems which occur in practice. Being a somewhat seasoned researcher myself, where threat modelling takes one form or the other in almost every project, it was natural choice to select an approach that sought to make life easier for me and my peers.

The overall research design is based on the *engaged scholarship* model by Van de Ven [14], which combines knowledge from several fields to produce practical contributions that address

complex real-world problems. Mathiassen [15] has defined a generic structure of engaged scholarship studies, drawing on Checkland’s model of scientific inquiry [16], as shown in Figure 1.2. The central component *Research question* (RQ) is developed based on known real-world *Problems* (P) and the related *Area of concern* (A) from the literature. The research question is addressed using a *Conceptual Framework* (F) with a suitable *Method of inquiry* (M). The conceptual framework serves as the key intellectual vehicle for answering the RQ, drawing on relevant theory and analytic frameworks. The method of inquiry consists of the specific research methods used in the analysis. The *Research* leads to the *Contribution* (C), that should be of benefit to the general area of concern (contribution CA) and to the specific problems owned by key stakeholders (contribution CP). It is also possible that the contribution can improve the conceptual framework (contribution CF) as well as the method of inquiry (contribution CM), but this has not been an intentional concern for this PhD study.



**Figure 1.2:** Generic structure of engaged scholarship. Adapted from Mathiassen [15]

### 1.3 Overview of papers

The contributions take the form of a number of research papers, that have either been submitted, accepted or published. Costley and Fulton [13] refer to this as the *continental model*, which ensures that “the research has been conducted soundly, securely, ethically and with a robust methodology”.

The contributing papers are divided into the categories *primary* and *secondary* papers, as well as papers presented as *posters*.

### 1.3.1 Primary papers

The primary papers have the closest relationship to the main research question, and are included both as summaries in Chapter 7 and in their complete form in Appendix A.

- A** A. Bagnato, B. Kordy, P. H. Meland and P. Schweitzer, ‘Attribute decoration of attack–defense trees,’ *International Journal of Secure Software Engineering (IJSSE)*, vol. 3, no. 2, pp. 1–35, 2012. DOI: <https://doi.org/10.4018/jsse.2012040101>
- B** P. H. Meland, I. A. Tondel and B. Solhaug, ‘Mitigating risk with cyberinsurance,’ *IEEE Security & Privacy*, vol. 13, no. 6, pp. 38–43, 2015. DOI: <https://doi.org/10.1109/MSP.2015.137>
- C** K. Bernsmed, C. Frøystad, P. H. Meland, D. A. Nesheim and Ø. J. Rødseth, ‘Visualizing cyber security risks with bow-tie diagrams,’ in *International Workshop on Graphical Models for Security (GramSec)*, Springer, 2017, pp. 38–56. DOI: [https://doi.org/10.1007/978-3-319-74860-3\\_3](https://doi.org/10.1007/978-3-319-74860-3_3)
- D** P. H. Meland, I. A. Tøndel, M. Moe and F. Seehusen, ‘Facing uncertainty in cyber insurance policies,’ in *International Workshop on Security and Trust Management*, Springer, 2017, pp. 89–100. DOI: [https://doi.org/10.1007/978-3-319-68063-7\\_6](https://doi.org/10.1007/978-3-319-68063-7_6)
- E** P. H. Meland and F. Seehusen, ‘When to treat security risks with cyber insurance,’ *International Journal on Cyber Situational Awareness*, vol. 3, no. 1, pp. 39–60, 2018. DOI: <https://doi.org/10.22619/ijcsa.2018.100119>
- F** P. H. Meland, K. Bernsmed, C. Frøystad, J. Li and G. Sindre, ‘An experimental evaluation of bow-tie analysis for security,’ *Information & Computer Security*, vol. 27, no. 4, pp. 536–561, 2019. DOI: <https://doi.org/10.1108/ICS-11-2018-0132>
- G** U. Franke and P. H. Meland, ‘Demand side expectations of cyber insurance,’ in *2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, IEEE, Jun. 2019, pp. 1–8. DOI: <https://doi.org/10.1109/CyberSA.2019.8899685>
- H** P. H. Meland, B. H. Johansen and G. Sindre, ‘An experimental analysis of cryptojacking attacks,’ in *Nordic Conference on Secure IT Systems (NordSec)*, Springer, 2019, pp. 155–170. DOI: [https://doi.org/10.1007/978-3-030-35055-0\\_10](https://doi.org/10.1007/978-3-030-35055-0_10)
- I** P. H. Meland and G. Sindre, ‘Cyber attacks for sale,’ in *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, IEEE, 2019, pp. 54–59. DOI: <https://doi.org/10.1109/CSCI49370.2019.00016>
- J** P. H. Meland, Y. F. F. Bayoumy and G. Sindre, ‘The Ransomware-as-a-Service economy within the darknet,’ *Computers & Security*, vol. 92, no. May 2020, 2020. DOI: <https://doi.org/10.1016/j.cose.2020.101762>

- K** K. Haga, P. H. Meland and G. Sindre, 'Breaking the cyber kill chain by modelling resource costs,' in *International Workshop on Graphical Models for Security (GramSec)*, Springer, 2020, pp. 111–126, ISBN: 978-3-030-62230-5. DOI: [https://doi.org/10.1007/978-3-030-62230-5\\_6](https://doi.org/10.1007/978-3-030-62230-5_6)
- L** P. H. Meland, S. Tokas, G. Erdogan, K. Bernsmed and A. Omerovic, 'A systematic mapping study on cyber security indicator data,' *Electronics*, vol. 10, no. 9, p. 1092, 2021. DOI: <https://doi.org/10.3390/electronics10091092>
- M** P. H. Meland, D. A. Nesheim, K. Bernsmed and G. Sindre, 'Assessing cyber threats for storyless systems,' *Submitted to Information Security and Applications*, 2021, ISSN: 2214-2126

### 1.3.2 Secondary papers

The secondary papers can be thought upon as tools for situational inquiry and gaining domain knowledge. Since these have a more supportive and exploratory nature to my research, as well as presenting early results, they are not included in the thesis itself, but can be retrieved by following their reference.

- N** C. Frøystad, K. Bernsmed and P. H. Meland, 'Protecting future maritime communication,' in *Proceedings of the 12th International Conference on Availability, Reliability and Security*, Association for Computing Machinery, 2017, pp. 1–10, ISBN: 9781450352574. DOI: [10.1145/3098954.3103169](https://doi.org/10.1145/3098954.3103169)
- O** K. Bernsmed, C. Frøystad, P. H. Meland, T. A. Myrvoll *et al.*, 'Security requirements for SATCOM datalink systems for future air traffic management,' in *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, IEEE, 2017, pp. 1–10. DOI: <https://doi.org/10.1109/DASC.2017.8102083>
- P** P. H. Meland and F. Seehusen, 'When to treat security risks with cyber insurance,' in *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, IEEE, 2018, pp. 1–8, ISBN: 978-1-5386-4565-9. DOI: <https://doi.org/10.1109/CyberSA.2018.8551456>
- Q** Y. F. F. Bayoumy, P. H. Meland and G. Sindre, 'A netnographic study on the dark net ecosystem for ransomware,' in *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, IEEE, 2018, pp. 1–8. DOI: <https://doi.org/10.1109/CyberSA.2018.8551424>
- R** K. Bernsmed, M. G. Jaatun and P. H. Meland, 'Safety critical software and security - how low can you go?' In *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*, IEEE, 2018, pp. 1–6, ISBN: 978-1-5386-4113-2. DOI: <https://doi.org/10.1109/DASC.2018.8569579>

- S** P. H. Meland, K. Bernsmed, C. Frøystad, J. Li and G. Sindre, ‘An experimental evaluation of bow-tie analysis for cybersecurity requirements,’ in *ESORICS 2018 International Workshops, CyberICPS 2018 and SECPRE 2018*, Springer, 2018, pp. 173–191, ISBN: 978-3-030-12786-2. DOI: [https://doi.org/10.1007/978-3-030-12786-2\\_11](https://doi.org/10.1007/978-3-030-12786-2_11)
- T** Ø. J. Rødseth, P. H. Meland, C. Frøystad and O. V. Drugan, ‘PKI vs. Blockchain when securing maritime operations,’ *European Journal of Navigation*, vol. 18, no. 3, pp. 4–11, 2018, ISSN: 1571-473-X. [Online]. Available: <http://hdl.handle.net/11250/2612306>
- U** M. Branlat, P. H. Meland, T. E. Evjemo and A. Smoker, ‘Connectivity and resilience of remote operations: Insights from air traffic management,’ in *REA Symposium on Resilience Engineering Embracing Resilience*, 2019, ISBN: 978-91-88898-95-1. DOI: <https://doi.org/10.15626/rea8.15>
- V** T. Myklebust, P. H. Meland, T. Stålhane and G. K. Hanssen, ‘The Agile RAMSS lifecycle for the future,’ in *Proceedings of the 29th European Safety and Reliability Conference (ESREL)*, 2019. DOI: [https://doi.org/10.3850/978-981-11-2724-3\\_0170-cd](https://doi.org/10.3850/978-981-11-2724-3_0170-cd)
- W** Ø. J. Rødseth, C. Frøystad, P. H. Meland and K. Bernsmed, ‘The need for a public key infrastructure in international shipping,’ in *International Maritime and Port Technology and Development Conference (MTEC)*, 2019
- X** Ø. J. Rødseth, C. Frøystad, P. H. Meland, K. Bernsmed and D. A. Nesheim, ‘The Need for a Public Key Infrastructure for Automated and Autonomous ships,’ in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, vol. 929, 2020. DOI: <http://dx.doi.org/10.1088/1757-899X/929/1/012017>
- Y** G. Bour, K. Bernsmed, R. Borgaonkar and P. H. Meland, ‘On the certificate revocation problem in the maritime sector,’ in *Nordic Conference on Secure IT Systems (NordSec)*, Cham: Springer International Publishing, 2021, pp. 142–157, ISBN: 978-3-030-70852-8. DOI: [https://doi.org/10.1007/978-3-030-70852-8\\_9](https://doi.org/10.1007/978-3-030-70852-8_9)
- Z** D. A. Nesheim, Ø. J. Rødseth, B. M. v. Zernichow, P. H. Meland and K. Bernsmed, ‘Secure, Trustworthy and Efficient Information Exchange – Enabling Added Value through The Maritime Data Space and Public Key Infrastructure,’ in *the 20th Conference on Computer Applications and Information Technology in the Maritime Industries (COMPIT’21)*, 2021
- Æ** P. H. Meland, K. Bernsmed, E. Wille, Ø. J. Rødseth and D. A. Nesheim, ‘A retrospective analysis of maritime cyber security incidents,’ *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation*, vol. 15, no. 3, pp. 519–530, 2021, ISSN: 2083-6473. DOI: <https://doi.org/10.12716/1001.15.03.04>



### 1.3.3 Posters

Posters have been used to present a visual and high-level view of my research at conferences. They are included in Appendix B, with an accompanying extended abstract in their conference proceedings.

- Ø P. H. Meland, 'Combining threat models with security economics,' in *The 11th Norwegian Information Security Conference (NISK)*, IEEE, 2018. [Online]. Available: <https://ojs.bibsys.no/index.php/NISK/article/view/570/486>
- Å P. H. Meland, 'Resilient cyber security through cybercrime market analysis,' in *REA Symposium on Resilience Engineering Embracing Resilience*, 2019, ISBN: 978-91-88898-41-8. [Online]. Available: <https://open.lnu.se/index.php/rea/article/view/1975/1695>

## 1.4 Structure of the thesis

As already mentioned, the thesis is a compilation of papers. In order to give a holistic view of the work, there is an initial *capstone*<sup>a</sup> that precedes the papers. The capstone chapters have been organised according to the components of engaged scholarship structure as defined by Mathiassen [15]. The area of concern, its state of the art and central definitions are explained in Chapter 2, followed by the problem setting in Chapter 3. The overall research question is broken down into more specific ones in Chapter 4. The conceptual framework, or framing, is discussed in a reflective manner in Chapter 5, and the methods of empirical inquiry are explained in Chapter 6. The contributions are summarized in Chapter 7 and an overall discussion of these are given in Chapter 8. Chapter 9 contains the thesis conclusion. The index contains clickable references to central topics throughout the capstone and definitions of abbreviations. Appendix A contains the primary papers, Appendix B the posters and Appendix C the associated awards.

---

<sup>a</sup>A capstone is the final (often decorative) brick put on a building [13].

## AREA OF CONCERN

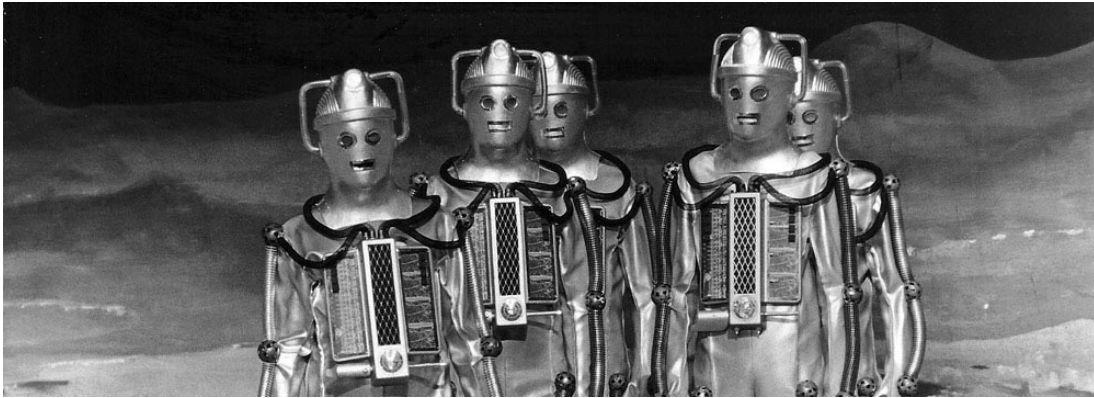
*Cybermen can survive more efficiently than animal organisms. That is why we will rule the galaxy.*

---

Cyberman, Dr Who, 1975

This PhD research is positioned within the overall area of *cyber security*. There are several more or less similar definitions of this term in the literature, and it is often used interchangeably with the term *information security* as shown by von Solms [46]. All security is about the protection of *assets* from possible harm resulting from various *threats* and *vulnerabilities*. With cyber security, we focus on systems that include *cyber resources* as assets. As defined by the *National Institute of Standards and Technology* (NIST) [47], a cyber resource “creates, stores, processes, manages, transmits, or disposes of information in electronic form and which can be accessed via a network or using networking methods.” The interdependent network of information technology infrastructures is often referred to as *cyberspace*, coined in 1982 by the science fiction writer William Gibson [48]. The word *cyber* by itself has had different meanings throughout history. Before becoming a prefix for words to do with the Internet (e.g., *cyberwar*, *cybercrime*, *cyberprank*, *cybersex*, *cyber monday*), it was used related to robot technology (e.g. the notorious *cybermen* that first appeared in the Dr Who series in 1966, see Figure 2.1) [49]. Cyber comes from *cybernetics* (control and communication theory) [50], which again is derived from the Greek κυβερνᾶν (meaning “steersman”).

Today, cyber security is also a concern for any asset that can be reached via cyberspace, including equipment, vehicles, people and the natural environment. The terms *cyber physical*



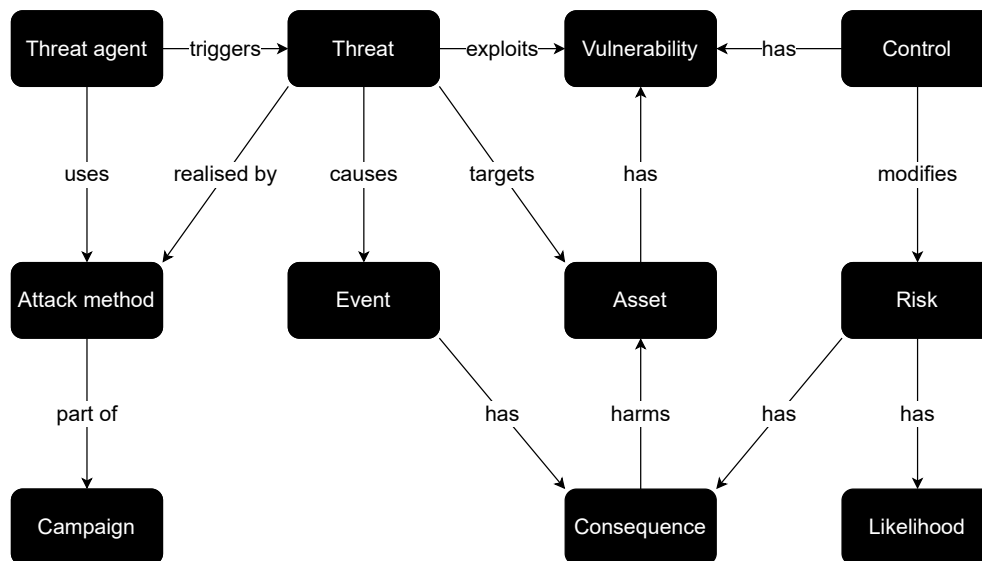
**Figure 2.1:** The cybermen. Photo used with permission from [51]

[52] and *socio-technical systems* [53] are often used when such tangible assets are involved. The term *cyber resilience* has become popular to describe organisational resilience against cyber threats, which typically include improved risk governance, incident response procedures, monitoring and threat information sharing. NIST [47] emphasize that in a systems security engineering context, *cyber resiliency* is about building systems so that they have “the ability to anticipate, withstand, recover from, and adapt to adverse conditions, stresses, attacks, or compromises on systems that include cyber resources”. They compare *cyber resilient systems* with the human body, having an immune and self-repair system that allows mission-essential functions to withstand and recover from infections and injuries. Hence, understanding the limitations of both humans and machines are fundamental when managing threats.

In the context of this PhD thesis, the following definitions paraphrased from the *International Organization for Standardization (ISO) / International Electrotechnical Commission (IEC) 27000* vocabulary [54] are especially important:

- A *threat* is the potential cause of an unwanted incident, which can result in harm to a system or organisation.
- An *attack (method)* is the attempt to destroy, expose, alter, disable, steal or gain unauthorized access to or make unauthorized use of an asset. An attack can be part of a *campaign* consisting of several activities.
- A *vulnerability* is a weakness of an asset or control that can be exploited by one or more threats.
- A *risk* is a positive or negative deviation from the expected, often expressed as the combination of *consequence* of an event and the associated *likelihood* (if likelihood is a numerical value between 0 and 1, the term *probability* is used instead).
- A *control* is a measure that is modifying risk.

Figure 2.2 shows how these terms relate to each other. It is partly based on the *Information System Security Risk Management (ISSRM)* domain model by Dubois et al. [55].



**Figure 2.2:** A domain model of central terms

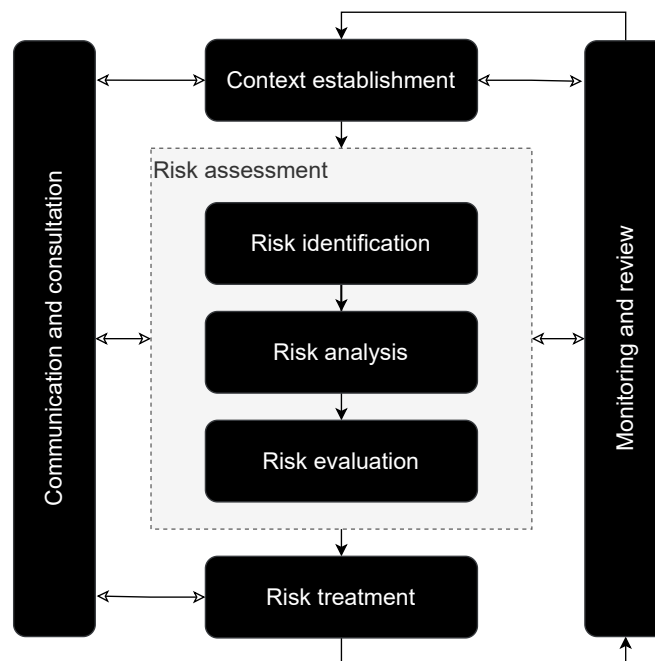
Additional central definitions from Dubois et al. are:

- An *asset* is anything that has value to an organisation, and thus needs to be protected.
- A *threat agent* can potentially cause harm to an asset. A threat agent triggers a threat and is thus the source of a risk.
- An *event* is the combination of a threat and one or more consequences.

NIST uses the term *risk management* to describe the process of identifying risk, assessing risk, and taking steps to reduce risk to an acceptable level [56]. This is practically the same definition as in the ISO/IEC 31000- and 27000-series ([57, 58]), which describe the risk management process as the set of activities shown in Figure 2.3. Here, solid arrowheads show the process flow between activities, while hollow arrowheads show information flow between concurrently ongoing activities. *Context establishment* specifies the basic criteria (such as evaluation, impact and acceptance), scope and boundaries, as well as the organisation responsible of the risk management. The goal of risk assessment is to quantify or qualitatively describe the risks, and consists of:

- Identifying risks, more precisely identifying assets, threats, existing controls, vulnerabilities and consequences.
- Analysing the risk level through some qualitative and/or quantitative method.
- Evaluating the risks against the criteria from the context establishment.

Obtaining a list of risks does nothing to secure a system by itself, so the next step in the process is to treat the risks using a combination of:

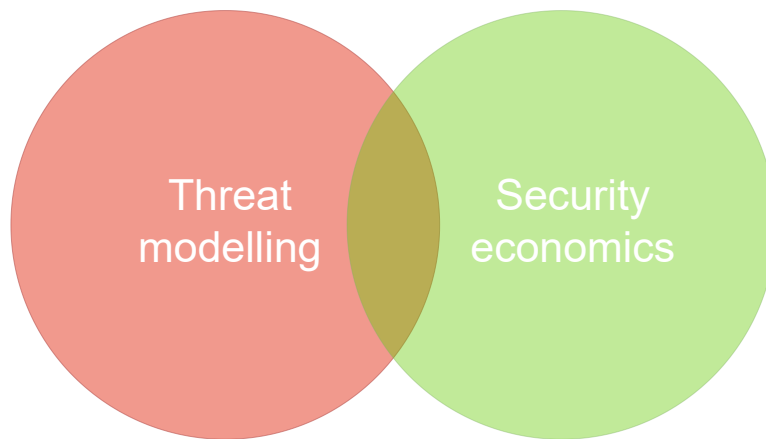


**Figure 2.3:** The set of activities involved in risk management. Adapted from ISO/IEC 27005 [57]

- *Risk modification*, which is about introducing, removing or altering security controls. Controls can provide the following types of protection; *correction, elimination, prevention, impact minimization, deterrence, detection, recovery, monitoring and awareness*.
- *Risk retention*, which is the decision of taking no further action for the assessed risk.
- *Risk avoidance*, which is about withdrawing from activities where the risk can occur or changing the operating conditions of the activities.
- *Risk sharing*, which involves sharing the risk with an external party, for instance through sub-contracting or insurance.

*Communication and consultation*, as well as *monitoring and review*, are activities that take place in parallel throughout the risk management process to share information and identify internal or external changes to the risk picture.

Within the area of cyber security, this PhD study focuses on two specific areas, or more precisely, the intersection between them. These two are *threat modelling* and *security economics* as depicted in Figure 2.4. They are usually considered in isolation, and it is the combination that creates focus [59] and is needed to address the complex problems within these fields. The established body of knowledge for this intersection is limited, therefore the sections below give an overview for both of them separately.



**Figure 2.4:** The intersection of threat modelling and security economics

## 2.1 Threat modelling

*Threat modelling*<sup>a</sup> belongs to the more general field of *security modelling*, which is concerned about “identifying system behaviour, including any security defences; the system adversary’s power; and the properties that constitute system security” [60]. Security modelling comes in many different forms and flavours, but they all share the common aim of understanding security issues so they can be dealt with effectively.

In 2000, Schneier [61] described threat modelling as a way of imagining the vast vulnerability landscape of a system and ways to attack it. He also made a point that this is something hard to do and only comes with experience. In the years to follow, threat modelling became commonly known as a central part of the Microsoft *Security Development Lifecycle* (SDL) [62–64]. In 2010, Steven [65] added additional steps to the SDL and re-defined threat modelling to be “the process of enumerating and risk-rating malicious agents, their attacks, and those attacks’ possible impacts on a system’s assets”, and repeated the message that this is something perceived as difficult and costly to perform. Besides the Microsoft SDL, threat modelling is a vital component in other cyber security frameworks as well. Comprehensive surveys by Bodeau et al. [66] and Magar [67] include the *NIST Framework for Improving Critical Infrastructure Cybersecurity* [68], *NIST SP 800-30* [69], *CBEST Intelligence-based Testing* [70], *Control Objectives for Information and Related Technologies* (COBIT), [71] *Cyber Prep methodology* [72], and *Operationally Critical Threat, Asset, and Vulnerability Evaluation* (OCTAVE) [73] to name a few.

In November 2020, a *Threat Modeling Manifesto* [74] was published by a diverse set of security experts from this field. It defines threat modelling as “analyzing representations of a system to highlight concerns about security and privacy characteristics”, and sums up some

---

<sup>a</sup>Threat *modelling* and threat *modeling* are used interchangeably in the literature depending on US or UK English writing style.

of the key issues that have appeared in the literature over the years. Among these are the four questions (originally published by Shostack [64]) that threat modelling should focus on:

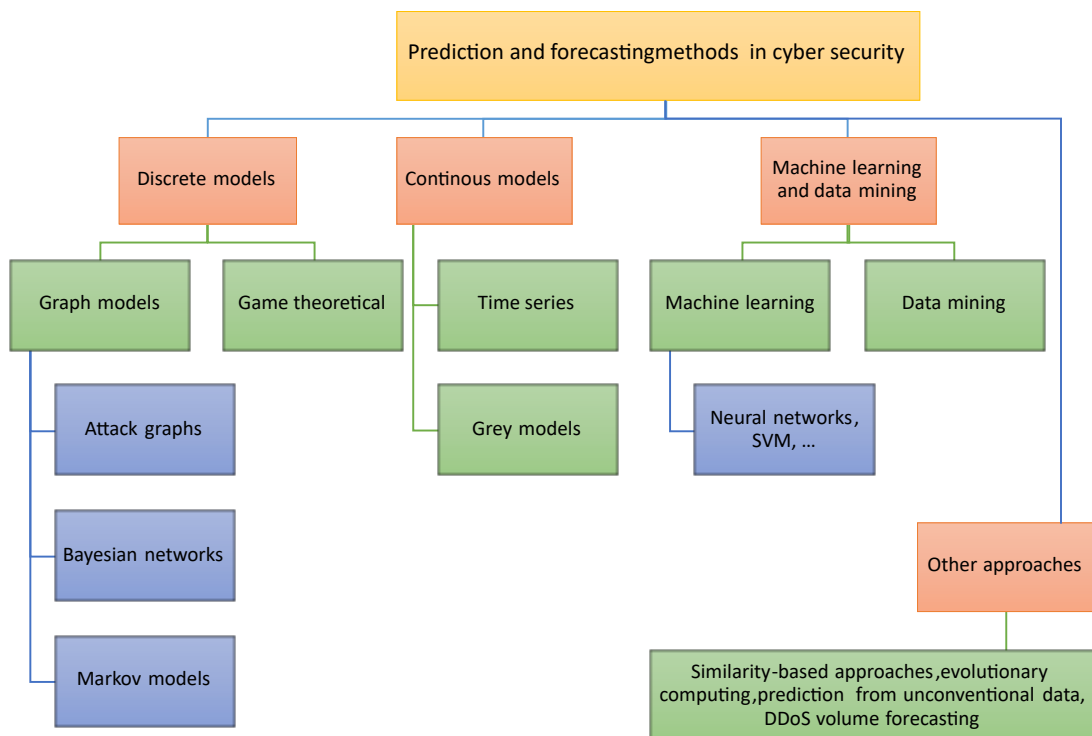
- 1) What are you building?
- 2) What can go wrong?
- 3) What should you do about those things that can go wrong?
- 4) Did you do a decent job of analysis?

While question 1 requires knowledge that you can (hopefully) find within the organisation that develops the system, question 2 requires a different mindset, namely trying to think like an attacker and exploit possible vulnerabilities within the system. For question 3, you need to switch back to defence again, eliminate vulnerabilities and build barriers that can withstand or cope with attacks. *Security requirements* [75, 76] are meant to guide this development process. Question 4 is difficult to assess, since threat modelling has a lot to do with prediction. Both when it comes to predicting what attackers might do and how well security barriers/controls/measures perform. Also, our systems and society exist in an evolving threat environment, so a system that might be regarded as secure at one point in time can be very insecure at a later stage. This requires that threat models are updated regularly to give us good analysis of the current and foreseen future situation.

Threat models are typically expressed through diagrams with different dialects of notations. Which one to choose usually depends on what you want to focus on, the level of abstraction/details and personal preference (e.g. familiarity). To quote Shostack [64]: “different diagrams will help in different circumstances”. A similar definition can be found in the Threat Modeling Manifesto [74], which states “it is better to create multiple threat modeling representations because there is no single ideal view, and additional representations may illuminate different problems”.

Many representations are tree-based and graph-based notations [77]. *Fault tree analysis* (FTA) [78], *event tree analysis* (ETA) [79], *attack trees* [80, 81], *defence trees* [82], *attack-defence trees* [83], *attack-fault trees* [84] and *capabilities-based attack trees* [85] are examples of the former, and they systematically refine attacker’s (or defender’s) goals into easily understandable actions. The typical process of quantitative risk assessment then consists of decorating basic actions with attributes and synthesizing cost values using a bottom-up approach [17, 83, 86]. *Cause-consequence analysis* (CCA) [87], *attack graphs* [88, 89] and *CORAS* [90] are examples of graph-based notations, and they typically enumerate all known paths that an attacker can take to intrude into a system. In a survey from 2017, Hong et al. [91] analyse the usability and practical applicability of different types of graphical security models. They argue that the diversity of models may confuse the users and limit adoption. A later survey from 2019 by Widell et al. [92] shows recent advances in graphical security modelling with focus on the application of formal methods.

Extensions to risk assessment techniques by means of stochastic tools such as *Markov chains* [93, 94], *Hidden Markov models* [95], *Bayesian networks* [96–99] and *Petri nets* [100] are well studied. Husák et al. [101] place such methods within their taxonomy of prediction and forecasting methods, as illustrated in figure 2.5. This taxonomy is based on a survey that also includes non-graphical methods. It can also be further specialised, for instance, attack graphs can be divided into *logical attack graphs*, *state-based attack graphs*, *hierarchical attack graphs*, *conservative attack graphs*, *multiple-prerequisite graphs* and *exploit dependency graphs* as shown by Barrère et al. [102].



**Figure 2.5:** An attack prediction and forecasting methods taxonomy. Adapted from Husák et al. [101]

Other security modelling approaches can be seen as extensions or additions to models where the original purpose is to document e.g., regular use, architecture or processes. For instance, Sindre and Opdahl [103] show how *misuse case* diagrams extend regular *Unified Modeling Language (UML) use case diagrams* [104], adding inverse use cases, which can be considered as threats, and mis-actors, who are malicious threat agents instantiating the misuse cases. Similarly to misuse cases, McDermott and Fox [105] have suggested *abuse cases* for expressing threats using the standard UML use case notation. In their approach, abuse cases are kept in separate models. Another example of extension is shown by Meland and Gjære [106], using error and escalation events to manage threats in *Business Process Model and Notation (BPMN) process*



diagrams, as well as in *collaboration* and *choreography* diagrams. *Data flow* diagrams [107] (aka *threat model* diagrams) is the preference within the Microsoft SDL. Here the focus is on data flows rather than control flows, and specifying which data stores and processes exchange data makes it suitable for determining the attack surface of the system.

Several threat models for analysis of attack and defence scenarios have been created in the past, with notable application in the domains of *supervisory control and data acquisition* (SCADA) systems [108], voting systems [109], Internet related attacks [110–112], secure software engineering [113], and socio-technical attacks [114–116].

## 2.2 Security economics

*Security economics* is a relatively new field, by many considered born from Anderson’s paper “Why information security is hard – an economic perspective” from 2001 [117]. The main argument from this paper is that security is not just shaped by technical measures, but also by economic mechanisms. Examples of such are [117–119]:

- *Perverse or misaligned incentives*, where someone can easily spend money on protecting his own computer, while refusing to spend a dime on preventing the same computer from attacking other because there is little incentive to do so. In other words, it pays off to be selfish (often referred to as *moral hazard*).
- *Information asymmetry or hidden-action problems*, where one of the parties involved in a transaction knows more about the quality of a product than the other, or can impact the outcome with unobservable actions. In other words, it pays off to play unfair.
- *Externalities*, which are side-effect of economic transactions that may have positive or negative effects on third parties. In other words, you may become collateral damage no matter what you do. There more specific *network externality* comes into play when the value of the network grows more than linearly in the number of users, for instance when people choose a technology due to its dominant market position, not because of actual quality or security (“winner takes it all”).

The development of the field, up until 2012, has been summarized by Anderson in the paper “Security economics – a personal perspective” [120], which succeeded the previous survey papers he wrote with Moore [119, 121, 122]. Another survey by Cordes [123] was published in 2011, and the EU project IPACSO wrote a state-of-the-art report [124] in 2016. Since the start, the research community has grown and the scope of the field has broadened, combining security with different subfields of economics, psychology and neighbouring humanities subjects. For instance, *security econometrics* is about considering how much time, money or effort should be devoted to security given that time, money and good people are always in short supply [125].

Making the right security investments depends on gathering relevant data and making trade-offs between cost and level of protection. The NIST *Risk Management Guide for Information Technology Systems* [126] defines that the purpose of a *cost-benefit analysis* is to demonstrate that the costs of implementing risk-treatments can be justified by the reduction in risk-level. This type of analysis has been extensively researched and documented in the past, see e.g. [127–131], however it is a problem not easily solved. A report by the *European Union Agency for Cybersecurity* (ENISA) [132] shows that there are different models of calculating classical concepts such as *annual loss expectancy* (ALE) and *return on security investment* (ROSI), but the challenge is to populate the variables of these models with accurate and meaningful values [124, 133]. Shorten, Smith and Paté-Cornell [134] point out that despite the range of available cyber security tools and techniques, there is significant uncertainty surrounding their risk reduction value.

It is not just a question of how much to invest, but also where to invest in security. The term *utility* is central in order to understand the motivations of both defenders and attackers. Economists refer to this as a form of satisfaction from consuming goods and services [135]. A *utility function* can be used to compare benefits and costs of an investment [131]. As seen from the defender's point of view, cyber security is usually not something that generates income by itself, but one can think of it as a benefit that reduces costs/loss as a consequence of cyber crime [136]. Investments can be of different nature, and aligning them will in many cases be based on security metrics (see [137–141]), cyber insurance/risk transfer (see [142–146]), information sharing (see [147, 148]), and liability assignment (see [149]). Laube and Böhme [150] have surveyed the economic literature on the strategic aspects of defenders' information sharing, and Schatz and Bashroush [133], Alexeev et al. [151] and Kissoon [152] provide up-to-date reviews of the literature on optimal investment in cyber security. Recently, Gordon et al. [153] have divided the literature on cyber security investments into three main streams, namely:

- The trade-offs among combinations of security related expenditures given a fixed budget.
- Expenditures related to cyber insurance to transfer risks associated with security breaches.
- Deriving the optimal amount (budget) that should be invested in cyber security activities.

Security economics also covers the *econometrics of wickedness* [120], meaning the economic incentives that exists in the underground economy and for people performing cyber attacks. Schechter and Smith [154] use the term *economic threat modeling*. They compare this with fishing as illustrated in Figure 2.6, where the central questions are:

- 1) How difficult it is to catch fish?
- 2) How much are consumers willing to pay for fish?

Just as the security of the fish depends on the number of fishermen and their resources (rods, lines, nets) at their disposal, the security of the system depends on the number of threat agents



**Figure 2.6:** The security of the fish depends on the fishermen and their potential profit

who stand to profit from attacking. The desired security level can be found by “quantitatively determining the point at which the costs to a potential attacker outweigh the benefits of attack” [154]. A series of papers and reports have been written on this topic, such as [155–164]. The argument that Anderson makes in the 2020 edition of his “Security Engineering” book [1] is that “if you’re going to protect systems from attack, it’s a good idea to know who the attackers are, how many they are, where they come from, how they learn their jobs and how they’re motivated. This brings us to the *economics of cybercrime*.” Similarly, Casey et al. [165] argue that we need to focus more on the threat agents and their the economic motivations. Knowing the profile and capabilities of perpetrators is necessary to select the right kinds of mitigations that will impact them economically. Schechter and Smith [154] also point out that traditional threat models fall short because they do not provide a quantitative measure of how much security is enough to deter a given adversary.

The mentioned literature above is in line with *rational choice theory*, which is based on the idea that criminals will consider and evaluate their decisions before they commit a crime, and is useful for understanding the motivations of cyber criminals and countering them with deterrence policies [166]. Buldas et al. [167] refers to the *rational attacker’s paradigm*, which assumes that rational attackers:

- 1) Do not attack if the attack-game is unprofitable, and
- 2) choose the most profitable ways of attacking.

Geer, Jardine and Leverett [168] exemplify this paradigm with the fact that not many attackers wasted time designing malware for MacOS until it accounted for at least one-sixth of the OS market. The biggest rewards reside where the highest concentration of victims are to be found.

In many ways, executing a cyber attack is like creating an online start-up business. There are initial investments, like hiring staff, procuring or renting hardware equipment and network access, developing/purchasing software, training, advertisement (see Figure 2.7) and getting hold of large quantities of energy drinks. Additionally, one might have to set aside funds for bribes, lawyer expenses and fines. If someone is paying for the attack, there might be some money up front, but if the income depends on the attack success, there is a significant uncertainty related to the actual return of investment. This implies that even attackers have to make rational choices on how much to invest and where to invest in attacks.

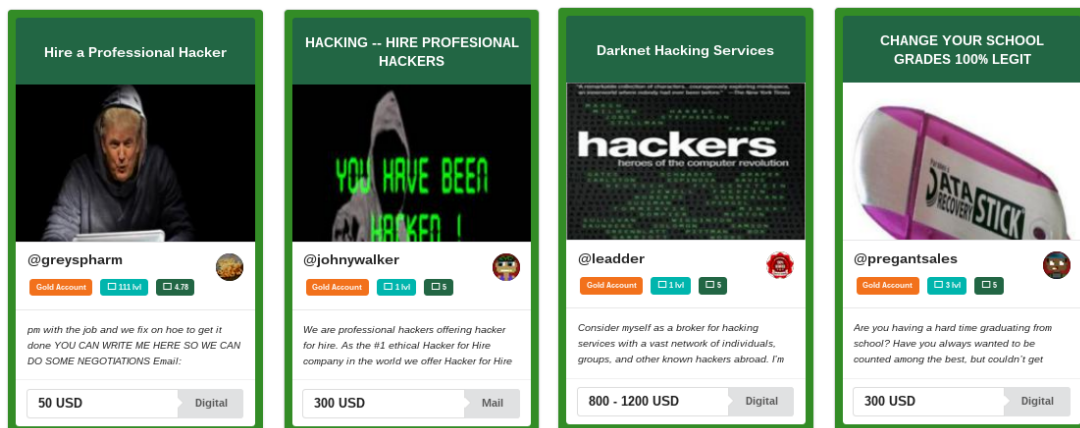


Figure 2.7: A sample of advertisements found on a darknet marketplace

When the utility is greater for the attacker than the defender, the situation may be said to favour the attacker, and vice versa [169]. In most cases, the attackers have the advantage, and as pointed out by Grobman and Cerra [9], there is an asymmetrical and unfair battle between these two sides. The attacker only needs to succeed once and can choose the time, place and method of his liking. The defender is one step behind and needs to prepare against all possible threats and be right 100% at all time. Hence, it is much more expensive to defend against a cyber attack than running it. Also, the damage costs are disproportionally high compared to the attack reward [136]. Kshetri [162] draws on a *dynamic choice model* of criminal behaviour [170] and shows how rational attackers make use of utility functions. In this case, there are also metrics of another nature than on the defender's side. An example is *opportunity costs* [171], which is not a direct cost but a type of loss, where the attacker has invested in something with a lower utility than an alternative. For instance, when an attacker would have made more money from attacking victim B instead of victim A, or when a legit job is more lucrative than crime. Kshetri also include the more intangible *psychic costs and benefits* of committing crime.

The use of *game theory* has become a popular way of modelling the strategies and interactions between cyber security opponents. Grossklags et al. [172] have described various types of

security games that support security decision-making, and extends this by modelling attackers as active and strategic economic actors [163]. More recently, Wang et al. [173] use a non-cooperative zero-sum game to model the attack and defence of the network. Both sides adopt an action strategy to generate reward and cost. When no player can improve their situation by changing strategy while the others stay constant, we have a *Nash equilibrium*. This is a concept that the economist John Forbes Nash Jr. received both a Nobel (1994) and Abel (2015) price for, and is fundamental in most game theoretic approaches [174, 175]. According to Shiva et al. [176], one of the central limitations for many such approaches is the assumption that all sides have near perfect information about the current state of the game. As we know, attackers have the advantage of playing unfair, hence information about their knowledge and actions are inaccurate. This reduces the prediction capabilities and practical applicability of the models. There has been a number of survey papers in the merging field of game theory and cyber security, such as [177–181].

## 2.3 Limitations

When combining threat modelling and security economics, there is an assumption that both defenders and attackers have a rational behaviour ([154]), driven by motives that will be of direct or indirect economic benefit. In many cases, attacks are not financially motivated, but due to e.g., political or religious reasons, personal revenge, or just plain fun. As pointed out by Sen et al. [182]; “methodological strategy of getting to actual behaviour via the concept of rationality has deep problems, though its advantages are also clear enough”. For instance, Clayton et al. [183] show that cyber crime concentration may be caused by non-economic factors related to copy cat criminals and uncaring attack host sites. Events arising from forces of nature, acts of God(s), accidents, mistakes and random events are also not within the main focus of this PhD study. However, such events may expose systems to economically motivated attacks, for instance, an accident could put a system into a safe state operating mode, which has less security controls running. Therefore, the combination and causality of malicious and non-malicious events is something that should be considered within the overall risk assessment process.

## PROBLEM SETTING

*There is no problem so bad  
that you can't make it worse*

---

Chris Hadfield

While the area of concern is part of the academic literature, the problem setting is based on real-world phenomena [15]. For this PhD research, the following phenomena are central to the problem setting:

- **Poor security investments:** The costs of cybercrime represent the greatest transfer of economic wealth in history, with an expected global growth of 15% per year going from 3 trillion USD in 2015 reaching 10.5 trillion USD by 2025 [184]. This is happening despite the fact that the global ICT security budget is growing every year [6, 185]. This disturbing trend makes it fundamental for our economy to manage cyber security threats in a more cost-effective way.
- **Little help from hindsight:** There is a general lack of historical data for cyber attacks, which again hinders the development of realistic models in cyber security [99]. As pointed out by Anderson et al. [149]: “Crime statistics are problematic enough in the traditional world, but things are harder still online.” At the same time, the technology and threats increase and develop so fast that historical data become irrelevant even after a relatively short time [186–189]. Risk quantification techniques that rely on historical data are therefore difficult to apply.

- **Malicious threat agents:** The likelihood of cyberattacks is hard to predict accurately since these are not randomly triggered unfortunate events, but rather depending on issues such as motivation and attacker capabilities, which are outside the control of an organisation. Most threat actors nowadays are motivated by economic motives and a vast economic ecosystem has developed around the business model of cybercrime [190].
- **Safety versus security:** Though many organisations have a long tradition dealing with safety-critical systems, they tend to be underprepared when it comes to cyber security threats. According to Abdo et al. [191], “existing approaches for industrial risk analysis ignore cyber-security”. Threats arise when these systems evolve from isolated entities to globally exposed cyber-physical systems. Safety and security have different traditions, standards, vocabularies and people addressing them. Instead of treating safety and security in separated processes, we need to learn more about the interaction [1].

With increasing systems complexity and number of attack methods, as well as criminal profit, these problems are constantly becoming harder to tackle. To overcome this, there is a need for better methods for quantifying cyber security risks, so that more informed decisions can be made for security investments. This must be in combination with, and not in conflict with, existing concerns related to risks. Though we can make pretty good estimates on consequences following a cyber event, the likelihood factor is a hard challenge. To quote Böhme et al. [188]: “Models of cyber risk *arrival* need to be more predictive.” Ahrend and Jirotko [192] are aligned with this, stating that “cyber security defenders need to make more informed decisions regarding what threats to mitigate and how to mitigate them” and “to do so requires defenders to *anticipate* threat actors’ behaviour”. Almukaynizi et al. [193] show that predicting cyber security events has received an increasing attention, and argue that predictions should be transparent and interpretable to allow human-in-the-loop-driven decisions.

Threat modelling is a means to support prediction, but as pointed out by Doynikova and Kotenko [194], practical use of techniques such as attack graphs tends to fail due to uncertainty of input data. Choosing one of the threat modelling techniques mentioned in Chapter 2 over another or inventing new ones is therefore not likely to improve the situation significantly by itself. Without proper input, we cannot expect good results. Unfortunately, the accuracy of crystal balls tends to be limited in this matter. Pure qualitative predictions made by experts are heavily dependent on experience and can be influenced by personal idiosyncrasies [191]. Hence, we are looking for data-driven ways of making these informed decisions. As shown by Brown et al. [195], there are many possible data sources about cyber threats, including sharing communities, open-source and commercial sources. The term used in this context is *threat intelligence*, which is any evidence-based knowledge about threats that can inform decisions [196]. The term can be further defined into the following sub-domains [197, 198]:

- *Strategic threat intelligence* is high-level information consumed by decision-makers, such as financial impact of cyber activity or attack trends, historical data or predictions regarding the threat activity.
- *Operational threat intelligence* is information about specific impending attacks against the organisation.
- *Tactical threat intelligence* is about how threat actors are conducting attacks, for instance attacker tooling and methodology.
- *Technical threat intelligence* (TTI) is more detailed information about attacker tools and methods, such as low-level indicators that are normally consumed through technical resources (e.g., *intrusion detection systems* (IDS) and malware detection software).

Wagner et al. [199] analyse and compare 30 threat intelligence platforms, such as the *Malware Information Sharing Platform* (MISP), Microsoft interflow, IBM X-force and McAfee threat intelligence exchange. Through interviewing 30 cyber security experts, Tundis et al. [200] have assessed existing *open source intelligence* (OSINT) sources. A lot of the research work today is dedicated to the sharing of low-level indicators of compromise/TTI, such as malware hash values or malicious IP addresses. This is useful for the *detecting* an ongoing attack rather than prediction, and therefore not so relevant in the scope of this PhD thesis. It is worth stressing that prediction and detection are complementing features, as prediction is a best-effort hit-and-miss game, while detection might be too late. For instance, a study by Griffioen et al. [201] showed that for 17 such open source feeds, it takes at least 20 days before active indicators become available. They also found that they have biases towards certain countries and that blocking listed IP addresses can yield large amounts of collateral damage.

This PhD work has been conducted together with a set of SINTEF-driven research projects with relevant industry stakeholders. The projects belong to the maritime and aviation sectors, and have provided real application environment consisting of stakeholders, organisational systems, technical systems, users, threats, assets and opportunities for improvements. In both sectors there are emerging technologies and digitalisation trends that can make ships and aircraft more exposed to cyberattacks. Consequently, this may threaten crew, passengers, goods, equipment and the environment. Inherent activities within these projects have been threat modelling and risk management in order to make cost-effective security prioritisation. A more sector-independent set of problem owners related to the phenomena above have been found within cyber insurance. This was a new product in Europe and in particular Norway at the start of the PhD work, and both insurers and insureds have been lacking experience with how to use this mechanism as a cost-effective risk transfer option.

The sections below describe the cyber security problems of the maritime, aviation and cyber insurance areas in more detail.



### 3.1 Aviation

*Aviation* is the activities surrounding mechanical flight and the aircraft industry [202]. Though aviation has had a long history of focussing on safety concerns, cyber security has not been integrated in the same way yet. This is arguable due to physical risks (e.g., bombs, guns, rocket-propelled grenades) being more prominent following events such as when three airliners (two Boeing 767 and one Boeing 757) were hijacked and used as weapons during the 9/11 terror attack in 2001 [203] (see Figure 3.1), the Malaysia Airlines Flight 17 (MH17) was shot down while flying over Ukraine in 2014 [204] and a laptop rigged as a bomb exploded inside the Daallo Airlines Flight 159 over Somalia in 2016 [205].



**Figure 3.1:** The memorial built after the attacks on the World Trade Center towers.  
Photo by Meland

Technology onboard aeroplanes has not been so much reliant on cyber elements, hence there have been very few incidents caused by cyber attacks. Sampigethaya and Poovendran [206] point to cyber limitations such as:

- The pre-departure operations are paper-based.
- The communication between the pilot and *air traffic control* (ATC) is mainly voice over

VHF radio.

- Flight trajectories are pre-planned based on clearances, not readily adaptive to continuous dynamics from bad weather, emergencies and traffic congestion.

However, through digitalisation in the aviation industry, this is now changing, and the remedies include:

- Satellite-based aircraft navigation and sharing of position.
- 4D trajectory management [207] (a real-time 3D route with time constraints).
- Digital data links between air traffic control and pilot (systems).
- A global information network of real-time air traffic control and meteorological data.

The driving force between the research and development of this technology in Europe is the *Single European Sky ATM Research* (SESAR) project [208], while the U.S. counterpart is called *Next Generation Air Transportation System* (NextGen) [209].

The increased integration between cyber-elements and aircraft systems warrants “surgical consideration” of potential risks from cyber and physical threats [206], something that is reflected in the recent aviation regulations. However, with new technology there is little relevant history to base these considerations on. Though there have not been many significant incidents related to cyber events, the threat is real and the damage potential just as significant as with a bomb onboard. Already in 2013, both Boeing and NATO ranked cyber-terrorism as one of the foremost threats to international aviation [210].

Both before and during my PhD study I have been involved in research and development related to some of the upcoming solutions for European airspace. The *Iris Precursor* project [211] was developing a new satellite communication link between the cockpit and ATC operators, effectively allowing 4G trajectory management and less reliance on voice communication. Working with the security layer on top of this link challenged us during risk assessments and when describing safety cases<sup>a</sup> as we had to focus on future and evolving threats for the next 10-20 years. In the subsequent *Iris Service Evolution* project [213], we repeated this exercise in an even more realistic setting that also included airborne trials.

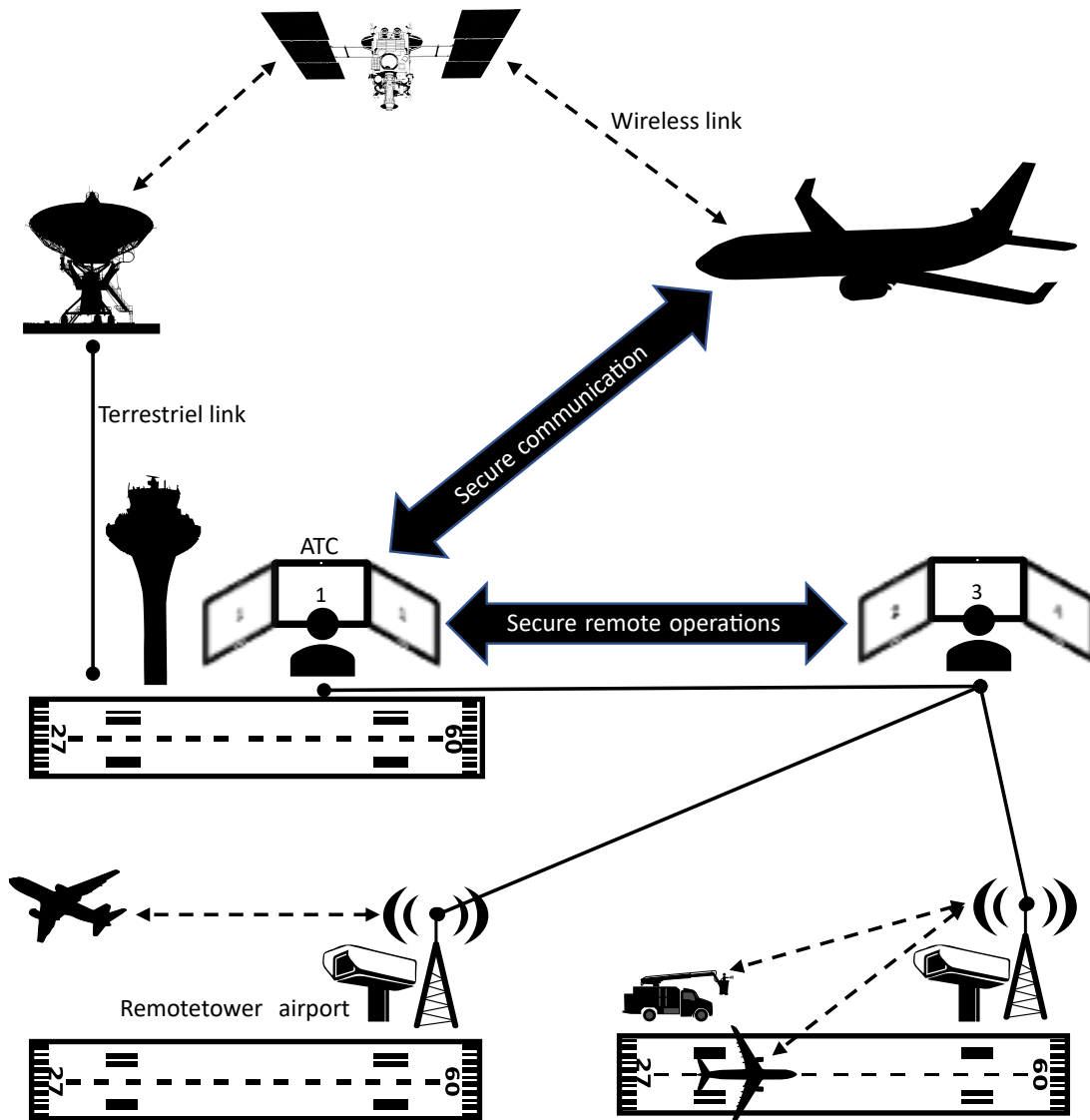
My second link to aviation has been the SESAR project *PJ05 Remote Tower for Multiple Airports* [214], which is developing the concept of remotely controlling multiple airports from a single physical location using new technology and working procedures. Instead of having an ATC operator present at every small and medium-sized airport, locally mounted cameras and a range of other sensors are instead providing data to control centres that manage the air and ground activities. A single ATC operator can then manage up to three airports on completely different locations depending on the number of movements and system state. Just as with

---

<sup>a</sup>A *safety case* is a way of creating confidence in a critical system, often without mathematical or statistical proof [212].

air-ground communication, this technology requires an increased reliance of cyber elements, and opens up to a new breed of threats that can be triggered from anywhere in the world.

Figure 3.2 illustrates the relevant focus areas related to the Iris programme and Remote Tower project, namely secure communication over a satellite datalink and secure remote operations between airports and ATC operators. Figure 3.3 was taken during a validation trial in Växjö (Sweden) during the spring of 2019. Here, the objective was to identify problems that could occur during handover of airport control between ATC operators. Similar trials were conducted in Asker (Norway) and Braunschweig (Germany).



**Figure 3.2:** The problem areas for aviation are focused on air-ground communication and remote tower operations



**Figure 3.3:** Trial simulating remote air traffic control. Photo by Meland with the courtesy of SAAB and LFV

## 3.2 Maritime

The *maritime* domain is defined as “all areas and things of, on, under, relating to, adjacent to, or bordering on a sea, ocean, or other navigable waterway, including all maritime-related activities, infrastructure, people, cargo, and vessels and other conveyances” [215]. According to Kontovas and Psaraftis [216], the *International Maritime Organisation* (IMO) has recognised that the whole philosophy of using historical data for *Formal Safety Assessment* (FSA) cannot be used for new system designs. Furthermore, it is undesirable to wait for new incidents to happen in order to measure the effects of newly implemented risk controls. Besides from a long tradition of safety-focus, there are many common traits and challenges compared to aviation, but the cyber security problems seem to be in a more severe state. An analysis by ENISA [217] lists the following key issues:

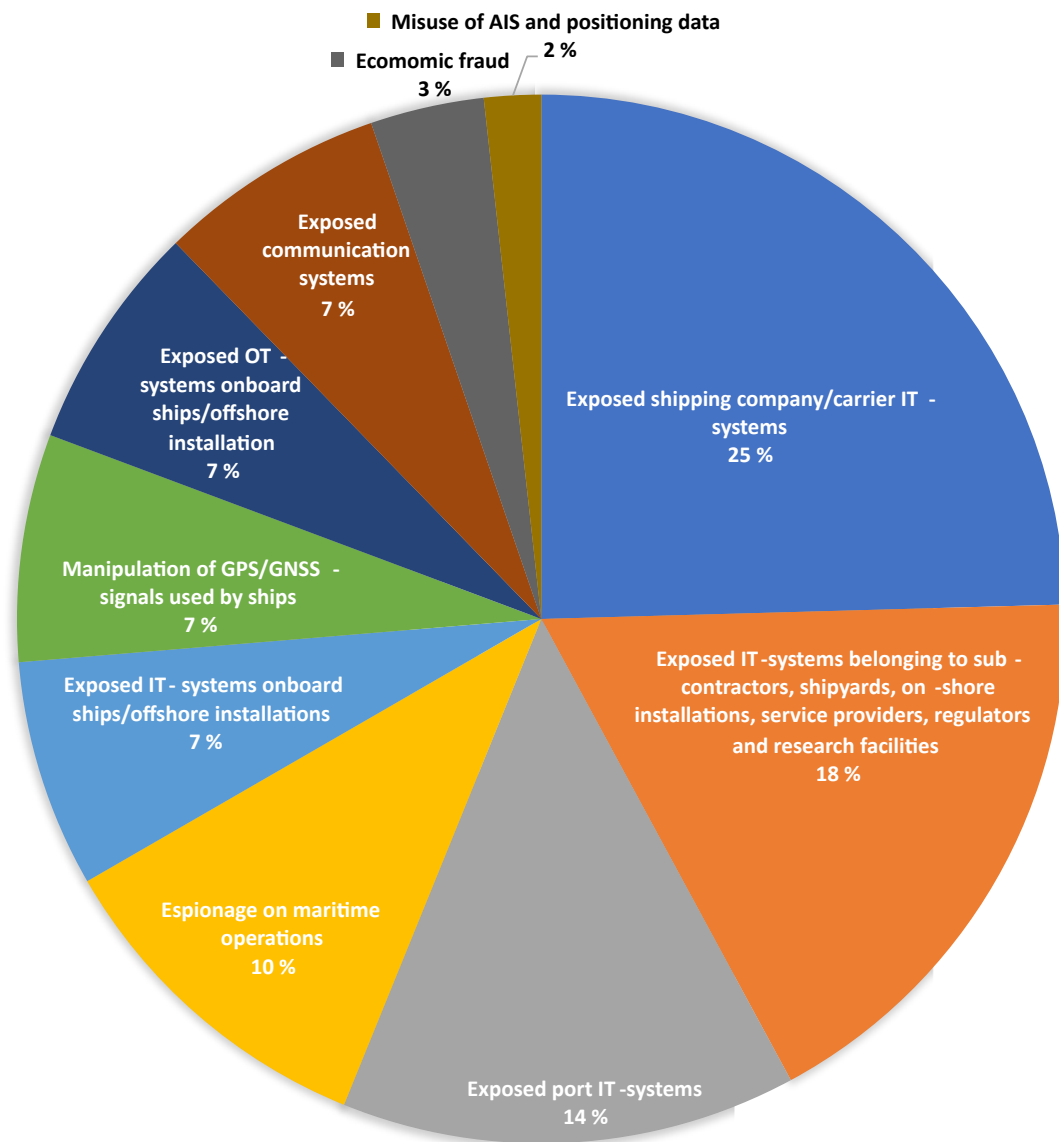
- The awareness and focus of cyber security in the maritime community has unfortunately been very low.
- A generally high complexity in the ICT infrastructure and fast technology development impair focus on security features.
- A fragmented maritime governance context that is not able to coordinate risks associated to cyber threats.
- The current regulations only include provisions related to safety and physical security concepts, while cyber security elements are insufficiently considered.
- There is no holistic approach to maritime cyber risks, expectations are set in an ad-hoc manner and only parts of the actual risks are considered.
- Being a highly cost-driven industry, there is an overall lack of direct economic incentives to implement good cyber security.

Ship operations represent high values and incidents can have severe consequences. An accident with large container ships or offshore structures may cost as much as \$1 billion [218]. If a cyber attack causes a ship to block, e.g., the approach to Rotterdam, Antwerp or Hamburg, the direct and societal costs could become much higher. A recent event where the Suez canal was blocked by the stuck “Ever Given” cargo vessel costed in trade about \$400 million an hour, or \$9.6 billion a day [219].

Due to the criticality of shipping, the future faces cyber threats from hostile states and economically motivated organised crime. IMO has stated that shipping is expected to become the “next playground for hackers” [220].

In early 2021, a retrospective analysis of maritime cyber threats and incidents from the last decade was published in a report by Meland et al. [221] and a subsequent publication [43]. Figure 3.4 shows the top 10 threats based on the frequency of reported actual incidents.

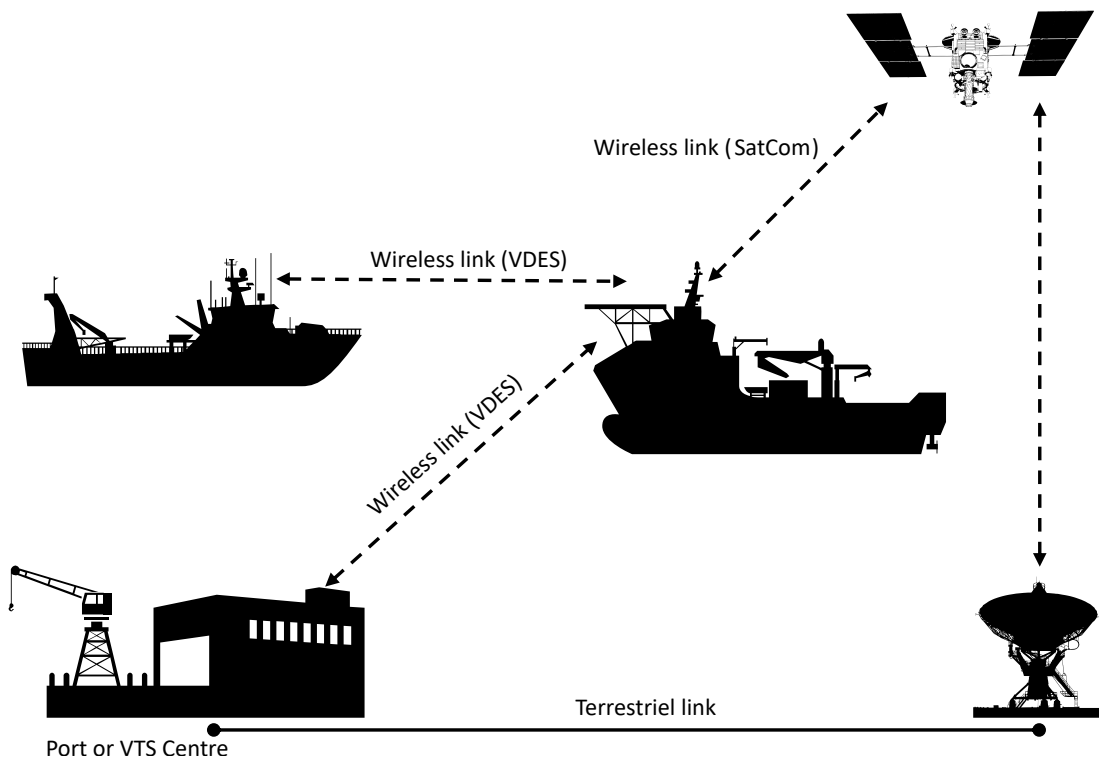
It is worth noting that the majority of them have been against the land-based operations and IT-infrastructure, while operational technology (OT) onboard ships has not been struck so hard yet. Most notably was the ransomware attack on the shipping company Maersk in 2017 [222], which affected more than one fifth of the world's shipping operations, including 76 ports. More than 4000 servers, 45000 PCs and 2500 applications had to be re-installed, and Maersk estimates their losses to be close to \$ 300 Million. This is regarded as the most devastating cyber attack in history.



**Figure 3.4:** The concentration of intentional maritime cyber threats from the last decade. Data from Meland et al. [43]

The maritime scope is a bit narrower in the context of my PhD study, which has been

about secure ship-ship and ship-shore communication for merchant ships (which excludes recreational pleasure crafts and military naval vessels). More specifically, the projects *Cyber Security in Merchant Shipping (CySiMS)* [223] and the follow-up *Cyber Security in Merchant Shipping Service Evolution (CySiMS-SE)*, both funded by the Research Council of Norway, have worked with cyber threats and a new technology for digital radio transmissions called *VHF Data Exchange System (VDES)* [224]. In parallel, the European project *CyberSec4Europe* [225] contained a maritime use case in which we were involved. The common goal of these projects has been to demonstrate and operationalise security for the radio communication solution and integrate it with the onboard computer architecture. The solution includes a *Public Key Infrastructure (PKI)* and the necessary hardware/software for secure information exchange across systems on the bridge, off-bridge and on-shore. Figure 3.5 depicts some of the possible communication links that each have different characteristics and need to be secured. The *vessel traffic service (VTS) centre* in the figure is similar to air traffic control for aviation, and has the responsibility of preventing incidents and accidents by monitoring and regulating ship traffic in defined areas along the coast [226]. Figure 3.6 shows one of the vessels that have been used for technology validation in the Trondheimsfjord area.



**Figure 3.5:** The problem areas for maritime are focused on ship-ship and ship-shore communication

What makes such communication challenging is related to the many natural limitations that



**Figure 3.6:** One of the ships tested with the CySiMS PKI solution and VDES radio.  
Photo courtesy of Kongsberg Seatex

one encounters for the maritime domain, such as:

- Ships are seldom continuously connected as wireless coverage may be very limited far from shore. This condition makes it difficult to use security solutions that rely on constant network access.
- Even when ships have a network connection during the voyage, they usually have a relatively low bandwidth and high communication latency. The cost of communication can also be a limiting factor, e.g. for satellite links.
- Ships regularly call on ports in other countries than where they are flagged. Some flag states might refuse to communicate with each other, and have incompatible technology or security solutions.
- Every ship is unique, meaning that they have a lot of specifically made software and different hardware onboard to support operations at sea. This makes it challenging to establish common security regulations and guidelines, and also very expensive to change/upgrade/patch the systems.
- Today, the *human-in-the-loop* is the main safeguard against information that has been tampered with. Increased automation and autonomy onboard ships will require improved mechanisms for authenticity and integrity checks of cyber elements.



All these factors lead to various security trade-offs, where cyber risks and economic implications are central inputs.

### 3.3 Cyber insurance

*Cyber insurance* can be defined as “the transfer of financial risk associated with network and computer incidents to a third party” [142], and is a mechanism to deal with contingent consequences of a cyber event. An *insuree* pays an annual fee to be covered by an insurance policy in case of such unforeseen events. An *insurer* is able to sell insurance policies by doing *actuarial analysis*, which is a mathematical analysis that normally takes past losses of similar insurees and projects them into the future to determine the reserves an insurer needs to keep and the rates to charge [227].

The *InSecurance* [228] project (2015-2016) was funded by SINTEF to investigate the challenges of cyber insurance, which was virtually a non-existing product in Norway at that time. A technical report from 2015 [229] gave an overview of academic and grey literature, which was limited to 36 scientific publications and a few reports, white papers and newspaper articles. This report identified a number of knowledge gaps and recommendations for further research. We learned that there was particular need for improved ways of doing risk quantification when there is little reliable actuarial data, and it became a natural extension to address this challenge as part of the PhD work.

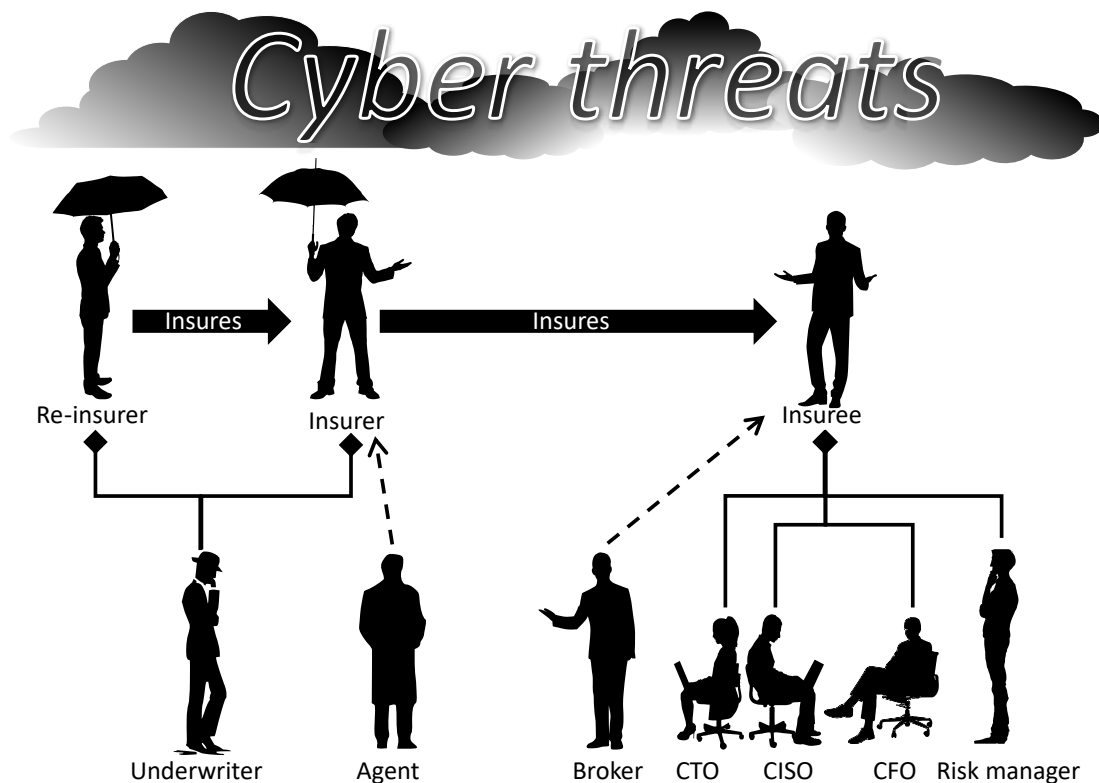
Though different forms of insurance can be traced 4-5 millennia back in time [230], cyber insurance is a relatively new product that has struggled to achieve a strong foothold despite of a substantial market potential [190]. Insurees are having problems determining whether the risks, coverage, price and terms are right for them, while insurers need good measurements, indicators and cost data for cyber threats on a macroeconomic level. A recent survey from the U.S. by Romanosky et al. [231] shows that policy pricing is to a large extent based on guesswork and looking at competitors. It is the lack of robust actuarial data that has been the most acknowledged reason for the somewhat limited success of the cyber insurance market [231–234]. Carfora et al. [235] have reviewed the recent literature on cyber risk management in the actuarial field, pointing at the immense difficulties to insure cyber risk and the lack knowledge about the quantification of cyber attacks in financial terms.

There are several different roles or stakeholders involved in (cyber) insurance, and in addition to insurers and insurees, these can be described as:

- A *re-insurer* insures all or a portion of an insurers risk under a contract. Re-insurers worry about risk concentration and events that will have national or global consequences.
- An *underwriter* classifies the potential insurees into risk pools and establish policy fees. Underwriters are in need of actuarial data and security profiles of insurees.
- An *agent* is a non-employee of an insurer that is authorised to sell policies. Agents often sell all kinds of policies and may have limited expertise on cyber threats and how to assess

the security of an insuree.

- A *broker* solicits insurance policies on behalf of the insuree, and may have limited expertise on cyber security.
- A *chief financial officer (CFO)* is an executive responsible of the financial planning of an organisation. CFOs mainly worry about keeping expenses and potential loss as low as possible.
- A *chief technology officer (CTO)* is an executive who manages the technology in an organisation and has a close relationship with the IT department, but may not know much about insurances.
- A *chief information security officer (CISO)* is an executive responsible of the (analogue and digital) information security of an organisation. CISOs are often not involved in cyber insurance considerations.
- A *risk manager* handles the insurance portfolio for an organisation, but will often not have a technical background and cyber risk is a new concern.



**Figure 3.7:** Different roles involved in cyber insurance

Figure 3.7 depicts the roles mentioned above, where lines with diamonds symbolise composition relationships and dashed arrows show dependencies. An insuree with a large budget typically has dedicated people fulfilling the specialised roles, while smaller organisations combine roles

(e.g., CTO, CISO and risk manager) or do not have them at all. Representatives of these roles have to a varying degree been involved in contributions of this thesis, supplying real-world problems, personal opinions, experiences and datasets.



## RESEARCH QUESTIONS

*The kinds of question we ask are as many as  
the kinds of things which we know*

---

Aristotle, 350 B.C.E

The research questions are based on the problem setting and open up to research contributions into the area of concern [15]. The first section below defines the main research question, before we look into the types and definitions of the particular and interrelated sub-questions. Furthermore, we explain the relationship to the paper specific questions.

### 4.1 Main research question

The objective of a research question is to provide a focus to the research. My work began with a main research question developed during the first months, and it has essentially been stable throughout the whole PhD period:

**Main RQ:** How can modelling threats and economic incentives improve cyber risk management?

### 4.2 Sub-questions

The *sub-(research-)questions* (SQ) are of a more specific nature than the main research question, and have been somewhat more dynamic throughout the course of the PhD period, both in terms of number and content. Table 4.1 enlists that became stable in the end of the study.

**Table 4.1:** Sub-questions and their type

ID	Description	Type
SQ1	What can we do to remedy situations with limited historical data of cyber security events?	Explanatory, Design
SQ2	What data can be used to model security economic incentives?	Descriptive, Relational, Predictive
SQ3	How can economic incentives enrich existing threat modelling techniques?	Design
SQ4	How can the cost of risk treatment options be balanced with cyber insurance?	Predictive, Design
SQ5	What can we learn about attack trends from studying phenomena within the cybercrime economy?	Descriptive, Explanatory, Predictive

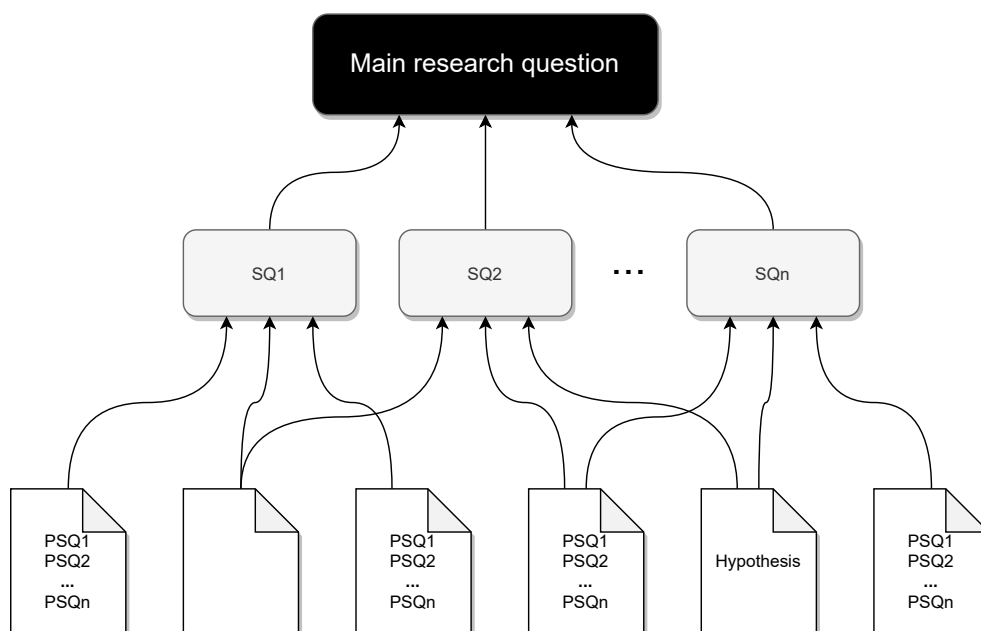
A survey on research question classifications by Dillon [236] shows that there are many different ways of doing that, and none that are perfect. To classify the ones enlisted in Table 4.1, a more modern scheme from the *research methods tool box* developed by the University of Twente [237] has been applied. The reason for this selection is that their separation between empirical and applied questions seems very much compatible with the engaged scholarship thinking. The following types of questions are defined:

- **Normative** questions define how desirable or good something is. Answers are usually more philosophical than empirical.
- **Empirical** questions have the goal to infer and generalise. These are in particular relevant for contribution to the area of concern. We can specify further:
  - **Descriptive** questions are about finding directly observable or inferential facts.
  - **Relational** questions involve examining causal or non-causal relationships between variables.
  - **Explanatory** questions are about explaining the causes for something.
- **Applied** questions seek to find solutions to specific problems, i.e. contribution to practice. These can be distinguished as:
  - **Predictive** questions are about finding out what will happen in the still unknown future.
  - **Remedy** questions find solutions to specific problems based on previous research.

- **Design** questions are about finding solutions to problems but not necessarily based on previous research.

### 4.3 Paper-specific questions

In order to guide the work of the individual papers, we have defined a third tier of questions or hypotheses. These are named *paper-specific questions* (PSQ), and they typically contribute to one or more of the sub-questions as depicted in Figure 4.1. Not all of the papers have well-defined research questions as such. They can instead have for instance a hypothesis or a goal. However, they are still able to contribute to the second tier of research questions.



**Figure 4.1:** Hierarchy of research questions

The paper specific questions can be found in the summary of the individual papers in Chapter 7 where applicable.





## CONCEPTUAL FRAMEWORK

*So the universe is not quite as you thought it was.  
You'd better rearrange your beliefs, then. Because  
you certainly can't rearrange the universe*

---

Isaac Asimov, Nightfall

As already stated in Section 1, my PhD research approach follows the principles of practice-based research and has an engaged scholarship design [14]. This implies that the work is characterised by a “creative process in which you discover and evaluate different ways to frame and publish your research by iteratively collecting and interpreting knowledge and evidence, exploring and testing ideas, and discovering and evaluating alternatives” [15]. Engaged scholarship can take on different forms, but in my case, it is *design and evaluation research*, also referred to as *design science* (DSR), that has been instrumental. Such studies create artificial knowledge of artefacts, policies, or programs for solving practical problems. Design science “is concerned with how things ought to be, with devising structures to attain goals” [238], in contrast to explanatory sciences that attempt to describe, explain, predict social systems. The next sections outline the theoretical basis this approach is based on.

### 5.1 Research paradigms

A *paradigm* is “a set of assumptions, concepts, values, and practices that constitutes a way of viewing reality” [239]. According to Costley and Fulton [13], practice research needs to reflect

the complexity of the world and answer to both academia and industry, and thus it is sensible to borrow from different paradigms. Because of this, the term paradigm is often replaced by *approach*.

Design science is derived from the engineering discipline, and is by many not fully accepted as a research paradigm by itself. According to Weber [240] and Hevner et al. [241], this is probably due to the failed theoretical contributions, and they show examples on how DSR has been combined with formal theory in order to make a contribution to the knowledge base besides focusing entirely on the problem solution.

In [242], Goldkuhl, discusses the meta-scientific debate concerning the rivalling trenches of *interpretivism* and *positivism*. He argues that interpretivism is in the lead when it comes to qualitative research, and that there is a lack of competitors. For instance, *critical research* is seen as a competitor by some, but others argue that it is a variant within interpretivism. In the same paper, he presents a third, new option, *pragmatism*, and draws on the literature to define it, going from the philosophy of John Dewey to much more recent publications on how information systems can use this paradigmatic base for action research and design research rather than positivism. The pragmatism paradigm is associated with action, intervention and constructive knowledge, which should be useful in practice. In contrast, interpretivism is more about understanding a phenomenon, and positivism about observing natural phenomena.

Pragmatism is a broad field by itself, and has been divided into 13 kinds in a classical article by Lovejoy in 1908 [243]. For my research approach, it is *methodological pragmatism* (MP), as defined by Goldkuhl [242], that has been most relevant. Key points on how MP is related to my PhD are shown in Table 5.1.

### 5.1.1 Was Grounded Theory relevant?

The basic principle of *Grounded Theory Method* (GTM) is the construction of theory from data, meaning that you do not start with a hypothesis or clearly defined research question, but more a notion about an issue that may (or may not) be present in a collection of qualitative data. To phrase this another way, GTM is the discovery of emerging patterns in data [244]. For me, the question was whether this is relevant for my research or not. Working with cyber attacks, discovering patterns is definitely relevant, but traditional threat intelligence usually works with quantitative data (logs in particular). For economic incentives, such as attack(er) costs, there are no good quantitative sources of such, and I had the idea to try to collect this from e.g. online communities in a qualitative way. Initially, I thought this will not fall under GTM, since I was not looking for a new theory, but simply sampling cost data that can be applied in threat models. However, during a workshop I attended [245], Professor Natalie Levina made the prediction that “with the growth of archival data, GTMs will increasingly use quantitative analysis techniques”.

**Table 5.1:** Mapping towards methodological pragmatism (MP)

MP key points	PhD relevance
Concerned about how (new) knowledge is created.	New knowledge is needed on the use of economic incentives as a part of cyber threat analysis. Traditional actuarial/historical data are difficult to obtain and rely on.
The researcher is active in creating data and theories.	I have been collecting data and created models in which data have been applied.
The researcher is participating in practice in order to explore and observe the effects and success of different tactics.	My PhD work was related to ongoing development projects that my organisation was doing for the industry. This means active participation in artefact development and evaluation.
Several methods and method combinations are used to the research purpose and empirical situation.	I have applied a mixed methods approach (see Chapter 6), as there are limited data for quantitative studies and stakeholder opinions needs to be collected in qualitative studies.

With that respect, GTMs was not really relevant for collecting cost data, but for discovering behavioural patterns that contradicted some of the common beliefs for certain cyber threats.

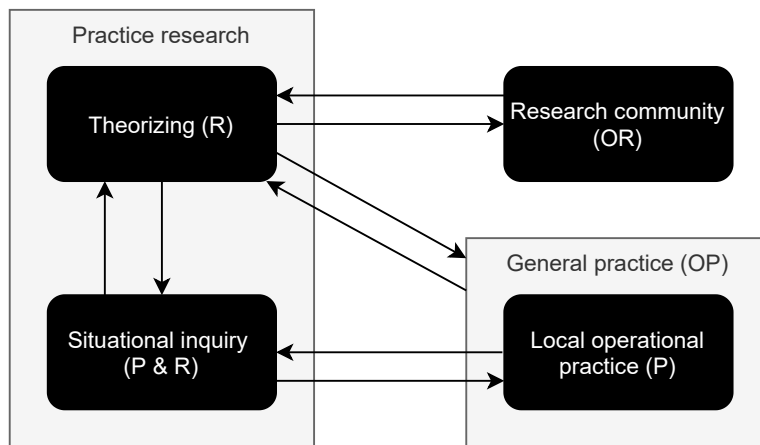
## 5.2 From theory to practice

There is a lot of overlap between *action research* and design science, but also quite a few differences [246]. The former is concerned with changes in an organisation and comes from the social science field. According to Coleman in [13], the involved researchers are typically researching their own practice or a situation in which they are personally/professionally involved.

With the latter, as already mentioned above, the focus is on the construction of new and useful artefacts, and it originates from the engineering discipline. In many cases, the introduction of an artefact will lead to an organisational change, therefore one can argue that both camps are involved. Others will argue that action research is more about the introduction of “normal design practice” [247], hence the novelty of the artefact is not significant.

Related to the problem setting of in my PhD work (see Chapter 3), new technologies are already changing the organisations, for instance going from analogue, voice-based communication to digital solutions and new services. My contributions were not the drive force for this new communication technologies themselves, but related to models for making informed decisions on how to develop them and balancing risks with costs. As a result, there was undoubtedly some overlap between action research and design science.

Since I have been looking at several domains with similar challenges and needs, I needed a way to combine results from each. Goldkuhl [246] argues that practice research is suitable for this, as it includes both intervention and design, and contributes to local practices as well as generalised knowledge. Figure 5.1 shows his proposed anatomy of practice research, where the arrows indicate interaction. *Researchers* (R) are producers of practice research; they are active in theoretical and empirical work. Together with (local) *practitioners* (P) they pursue *situational inquiries*. The target group is both *local operational practitioners* (P) as well as *general practitioners* outside (OP) the case studies. Generalised knowledge from *theorizing* (R), aimed for the wider *research community* (OR). Goldkuhl points out that both researchers and practitioners may have the driving role, and that there is a continual back and forth movement between situational inquiry and theorizing [248]. Theorizing provides abstract knowledge, while situational provides situational/empirical data.



**Figure 5.1:** A structural anatomy model for practice research. Adapted from Goldkuhl [248]

Practice research is meant to be an encompassing approach embracing different research approaches and compatible with engaged scholarship. It is still a relatively new approach, but can be seen as more suitable than a “pure” DSR study in my research.

## METHODS OF EMPIRICAL INQUIRY

*My methodology is not knowing what I'm  
doing and making that work for me*

---

Stone Gossard

This PhD study applies a *mixed method approach*, where quantitative and qualitative research methods are used in combination. Plowright [249] uses the following definition of mixed methods: “Mixed methods is the collection of different types of data using more than one method, approach or strategy derived from more traditional research paradigms or perspectives that draw on different epistemologies and explanations that inform and underpin knowledge claims.” This is in accordance with the principles of pragmatism (see Section 5.1), that addresses the question “what is the best approach to getting the most appropriate information we need to deal with a real-world practical problem?” [13].

The choice of using both qualitative and quantitative methods was also inspired by a lecture on research methods held by Professor Kalle Lyytinen [245]. He discussed different reasons for involving qualitative and quantitative methods, and I have summarised these along with a mapping towards my PhD study in Tables 6.1 and 6.2. The latter table also includes some supplementary points from Creswell [250].

**Table 6.1:** Why qualitative research was needed

General characteristics for qualitative research	PhD relevance
You need to ask what, why, how.	It is important to identify which incentives are useful, what kind of data can be used as input, what are the data sources. Furthermore, there should be a clear benefit of including incentives in threat modelling and a set of identified best practices.
There are motivations and underlying reasons that cannot be directly observed.	Security economics involves aspects from psychology and behavioural science to make predictions on attack likelihood.
There are new observations on behaviour.	Attacks evolve and attackers will behave differently over time, just as people being attacked can have unpredictable actions.
Need to get close to the phenomenon / context.	We have been working closely with stakeholders and their context, and obtained detailed knowledge about domain specific threats and modelling techniques.
Deal with processes / mechanisms.	Threat modelling is part of risk analysis and directly influences decision-making processes.
Need to understand meaning / experience of the actors.	Design science and evaluation involve actors, their opinions and experience when introducing new artefacts.
Do not exactly know what the issue is / Deal with multi-level phenomena / Need to understand the beaming, buzzing confusing world.	All risk quantification is hard by nature and a complex problem. Exploratory research is therefore a good option when it is difficult to pinpoint all problematic issues from the start, there are various phenomena on both the attacker and defender side in play, and data can be covert, out-of-date or misaligned.

**Table 6.2:** Why quantitative research was needed

General characteristics for quantitative research	PhD relevance
There are hard core facts and numbers involved.	This study accessed concrete monetary cost data related to cyber security events. These are by insurance companies regarded as the best data available, but far from perfect. In the future, even more incident data are expected to become available due to new legislation and ongoing initiatives for information sharing.
Behaviour facts can be measured, e.g., usage, frequency, amount bought or used.	Some behaviour data related to attackers can be mapped to threat models, e.g., number of attacks, number of sold attack tools and number of motivated attackers.
Pricing decisions are made based on the data.	Data can be used in risk models to estimate cost/benefit for security investments. For instance, underwriters use this kind of data to estimate premiums for cyber insurance.
Data are used to support major decisions.	The data have been used to aid decision making for different technology development projects. Also, stakeholders looking at the macroeconomic consequences of extensive attacks have an interest in the data.
There are factors or variables that influence an outcome (causality).	There are relationships between the different economic incentives, for instance high attack cost with low reward will cause low utility. Some of these factors have been studied and applied to threat estimations.



My strategy related to mixed methods has been to apply a *concurrent transformative approach* [250]. This has a perspective related to participatory research, and means a concurrent collection of both quantitative and qualitative data collected together, and that the data are mixed through merging, connecting or embedding. The methodological choices have been based on the suitability for the study at hand, and driven by the research questions.

The following sections give an overview of the research methods that have been applied for the primary papers. I have made a rough separation between *problem investigation methods*, *artefact creation* and *evaluation methods*. The first category is concerned about gaining knowledge about problems in the area of concern, the second about design and implementation of artefacts and the third about how evaluations have been conducted.

## 6.1 Problem investigation methods

A combination of methods has been used to analyse the problem setting. The goal has been to further describe and diagnose (explain) the problems seen from the stakeholders' side and observed phenomena from the literature. Generalised knowledge can be seen as a contribution about the causes and priorities for problems to be solved.

### 6.1.1 Literature study

Literature studies are essential components in every study to get an overview of related work and build on existing results. Initially in my PhD period, and as a part of the course "DT8114 - PhD Seminar in Computer and Information Science", I developed a body of knowledge from central articles related to the area of concern. This was used as a basis when working with the papers later on, which all contain tailored background information for the study at hand.

Paper L is a dedicated literature study work. Here, we specifically employed a *systematic mapping study* method, which is used to describe an existing research field without assessing specific details from each publication [251]. This allowed us to see focus areas, gaps and trends related to data-driven security indicator data.

### 6.1.2 Interview

*Interviews* with stakeholders, and in particular *semi-structured interviews* have been important for the situational inquiry in many of the primary papers. This is a method that usually follows a literature study, and the interviewees are in practice selected based on appropriateness and availability [252]. An interview guide is developed to help the interviewer and structure the questions. Open-ended questions allow the interviewee to follow interesting knowledge paths

(hence the “semi” in semi-structured). After the interview itself, considerable time is spent on transcription and data analysis. Anonymisation/pseudonymisation of data is usually a part of this process as well.

### 6.1.3 Netnography study

According to Kozinets [253], netnography consists of a particular set of actions for doing research within and about social media. It is centred around the study of *online traces*, which is any kind of media that people leave on the Internet.

In the research context of the papers, we have primarily been interested in traces related to cybercrime found in darknet markets and forums. Through unobtrusive online observation we have collected empirical data, used induction to look for patterns, and finally identify phenomena and trends. Though netnography is mostly rooted in qualitative research, we have included quantitative data related to trade activities as well.

## 6.2 Artefact creation

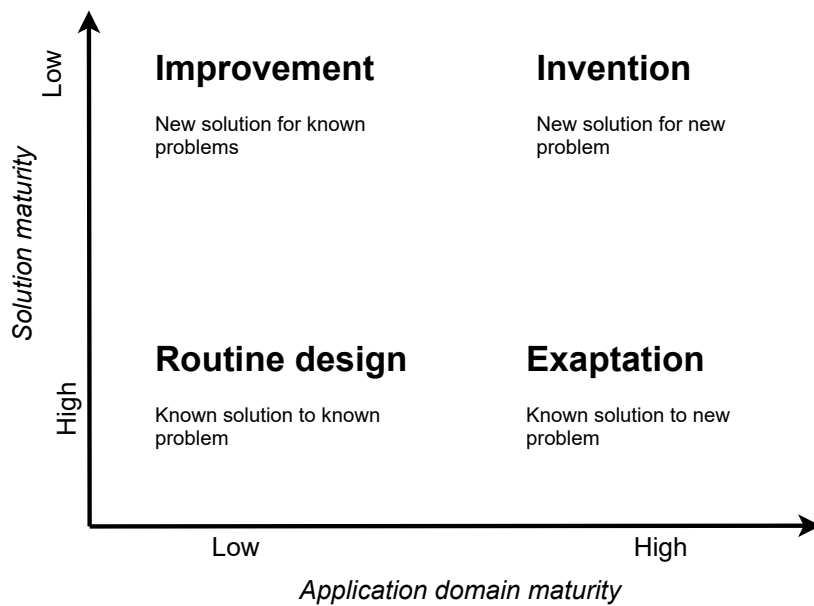
Artefacts can be seen as potential treatments to the identified problems. As shown by Gregor and Hevner [254], artefacts can be on different abstraction levels, ranging from grand theories to methods, models, design principles, and finally software products and implemented processes at the most specific level. As shown in Figure 6.1, artefacts are not always created from ground up, but can for instance be something existing applied to a new context.

The three main artefact types used in relation to the primary papers are described below.

### 6.2.1 Method artefact

*Method* and *technique* are terms often used indiscriminately, but one could argue that a technique is the more specific strategy to implement a method [255]. Within the primary papers, both terms are used, as well as *approach*, but for the sake of simplicity we only denote *method* as the overall term for “doing something to achieve something”. We also let *method* encompass *model creation*.

We have invented new methods to solve new problems in the area of cyber insurance, as well as extended known modelling techniques so that they can be applied for cyber security (exaptation). By using economic incentives to quantify risk we hope to achieve an improvement to a known problem. The methods include ways of gathering data, doing calculations/estimations and conducting threat modelling in practice.



**Figure 6.1:** Research context and potential contribution. Adapted from Gregor and Hevner [254]

### 6.2.2 Software tool artefact

Software tools are used as an aid to a method, and can be coded from scratch, modified from something existing or used “out of the box”. Depending on the study at hand, we have used all of the beforementioned strategies. For instance, the bow-tie modelling tool used in paper **F** and depicted in Figure 6.2 is a new solution for a known modelling technique, while the *Interactive Resource Cost Model* (IRCM) tool in paper **K** is a new solution to a new modelling method.

### 6.2.3 Test environment artefact

A test environment consists of software, hardware and/or network elements necessary to execute test cases [256, 257]. In some studies, we had to develop our own software and hardware setup in order to do benchmark evaluation as well as to gather data used to estimate attacker costs/profit. Hence, it can be seen both as a contribution by itself and something that support the other artefacts.

## 6.3 Evaluation methods

We have adopted the *DESMET* methodology by Kitchenham [258] to describe what kind of evaluation that has been conducted on the artefacts. Though *DESMED* is originally designed for software engineering methods and tools, security engineering is so much similar to (and often

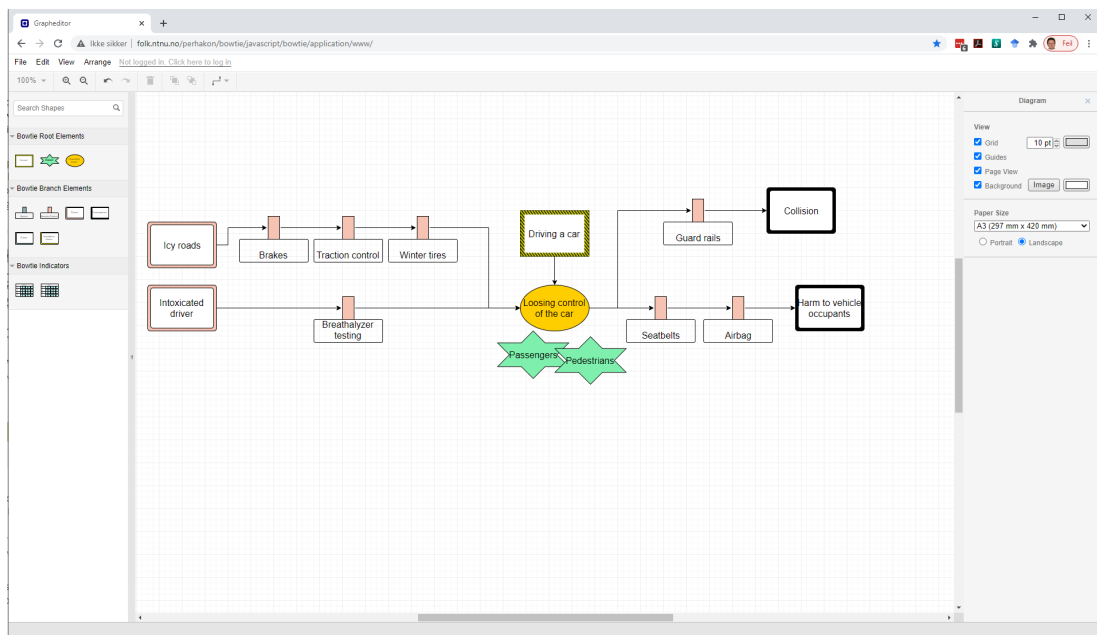


Figure 6.2: Screenshot from the bow-tie modelling software artefact.

part of) software engineering that the same principles apply. Kitchenham distinguishes between nine different evaluation types, but below we only include the ones that are relevant.

### 6.3.1 Quantitative formal experiment

Kitchenham refers to Pfleeger's [259] principles for *formal experiments*, which are basically that there is a high degree of control of the behavioural events and variables during the study. For instance, when investigating the effects of a method, the investigator has control over who uses the method, when and where it is used. A high level of control also makes experiment replication easy, but also requires a high degree of planning in advance.

We have a *quantitative* formal experiment when there are measurable properties that are expected to change as a result of the artefact. Some of the paper contributions have involved different types of users to experiment with artefacts such as threat modelling methodologies and prototype tools. The produced results from these experiments have been the basis of quantitative analysis.

### 6.3.2 Qualitative experiment

A *qualitative experiment* gathers feedback from a set of potential users after they try out the methods/tools on typical tasks in a controlled setting.

In our case, we have applied this method to compare different modelling techniques and tools (subjective benchmarking), as well to compare increments of the same tool. Opportunities from teaching courses and involvement in development projects gave us access to students, security experts and domain specialists.

Evaluation results have mainly consisted of feedbacks from interviews, evaluation forms and observations made by the investigators.

### 6.3.3 Qualitative case study

Compared to formal experiments, *case studies* are preferable when there is less control over variables and relevant behaviours cannot be manipulated [259]. This kind evaluation is done collecting *qualitative* feedback from someone using the artefact in a real project. Variable values should be typical for that kind of project.

### 6.3.4 Qualitative effects analysis

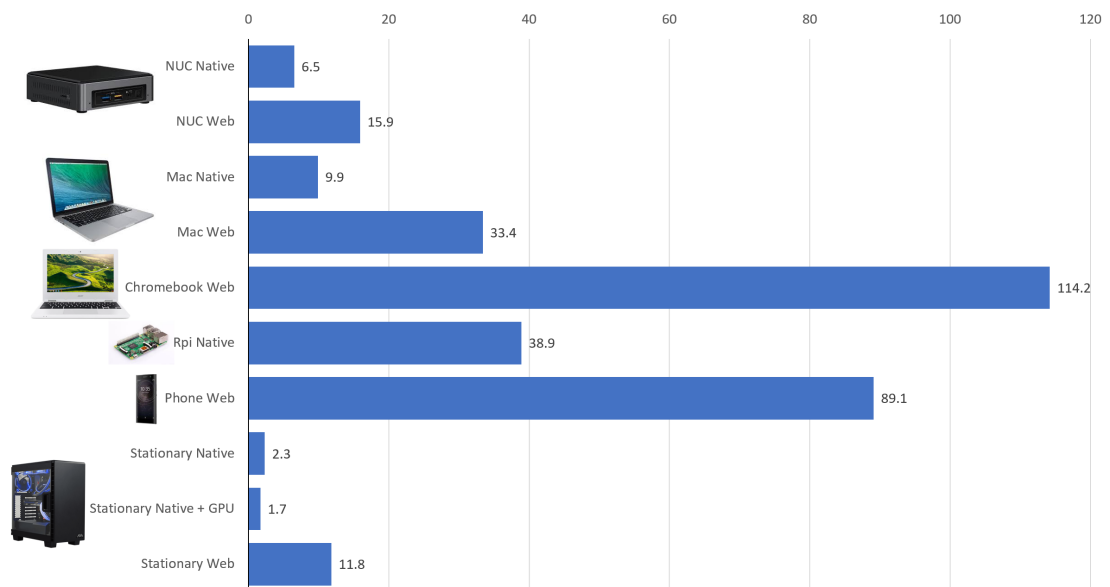
*Qualitative effects analysis* is used when an expert makes a subjective assessment on the quantitative effects of the artefact. This is regarded as a *hybrid evaluation method* by Kitchenham [258].

For instance, in paper H, we made subjective measurements on the performance decay of machines infected by a particular type of malware. We assume that these results are valid for other non-expert users as well.

### 6.3.5 Benchmarking

*Benchmarking* can be used to test alternative tools/methods and assess the relative performance between them. This is typically done objectively without user involvement, for instance as a *trial experiment* [260] measuring the performance of a system using predefined variables in a laboratory setting. Just as the method above, this is regarded as a hybrid evaluation method.

Specifically, in paper H, we have tried to estimate potential profit an attacker would hope to achieve using a specific exploit. This in turn can give an indication on the lucrativeness of the exploit and its likelihood. Different computer hardware and software were used as variables, and different measuring techniques gave us objective values for performance, and power consumption and profit. Figure 6.3 shows one of the benchmarking results from this study.



**Figure 6.3:** Benchmarking results of the time (years) it takes to mine a single Monero coin on different systems

## 6.4 Mapping of research methods

Table 6.3 shows how the research methods above are mapped to each of the primary papers. In many cases, more than one research method from *problem investigation*, *artefact creation* and *evaluation* have been applied. Note that *Literature study* is partially assigned to all papers except L. The reason for this is that they provide related work, while L is a dedicated literature study paper.

**Table 6.3:** Mapping between research methods and primary papers

Method ↓ Paper →	A	B	C	D	E	F	G	H	I	J	K	L	M
Literature study	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	✓	(✓)
Interview				✓			✓						
Netnography									✓	✓			
Method artefact	✓	✓	✓		✓	✓		✓	✓	✓	✓		✓
Software tool artefact					✓	✓						✓	✓
Test environment artefact								✓					
Quantitative formal experiment						✓							
Qualitative experiment	✓											✓	
Qualitative case study												✓	✓
Qualitative effects analysis		✓	✓		✓			✓	✓	✓			
Benchmarking								✓					

## CONTRIBUTIONS

*The best contribution one can make to  
humanity is to improve oneself*

---

Frank Herbert

This chapter highlights what the most important results from the individual papers in the context of the PhD theme. As already depicted in Figure 1.2 in Chapter 1, the contributions are meant to benefit the general area of concern (theory) as well as the more specific problems (practice). The papers are ordered chronologically within the categories primary and secondary papers, as well as poster contributions.

### 7.1 The use of the Contributor Roles Taxonomy

I have had the privilege to collaborate with many other researchers in almost all of my papers. In order to give duly credit to the contribution of each co-author, I have applied the *Contributor Roles Taxonomy* (CRediT) [261]. This taxonomy is adopted by many academic publishers, such as Springer, Elsevier and Sage, as well as a number of academic institutions. It consists of the following 14 standardised role descriptions representing contributions to scientific scholarly output:

- **Conceptualization** – Ideas; formulation or evolution of overarching research goals and aims.



- **Data curation** – Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.
- **Formal analysis** – Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data.
- **Funding acquisition** - Acquisition of the financial support for the project leading to this publication.
- **Investigation** – Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.
- **Methodology** – Development or design of methodology; creation of models.
- **Project administration** – Management and coordination responsibility for the research activity planning and execution.
- **Resources** – Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.
- **Software** – Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.
- **Supervision** – Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.
- **Validation** – Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.
- **Visualization** – Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.
- **Writing – original draft** – Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
- **Writing – review & editing** – Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.

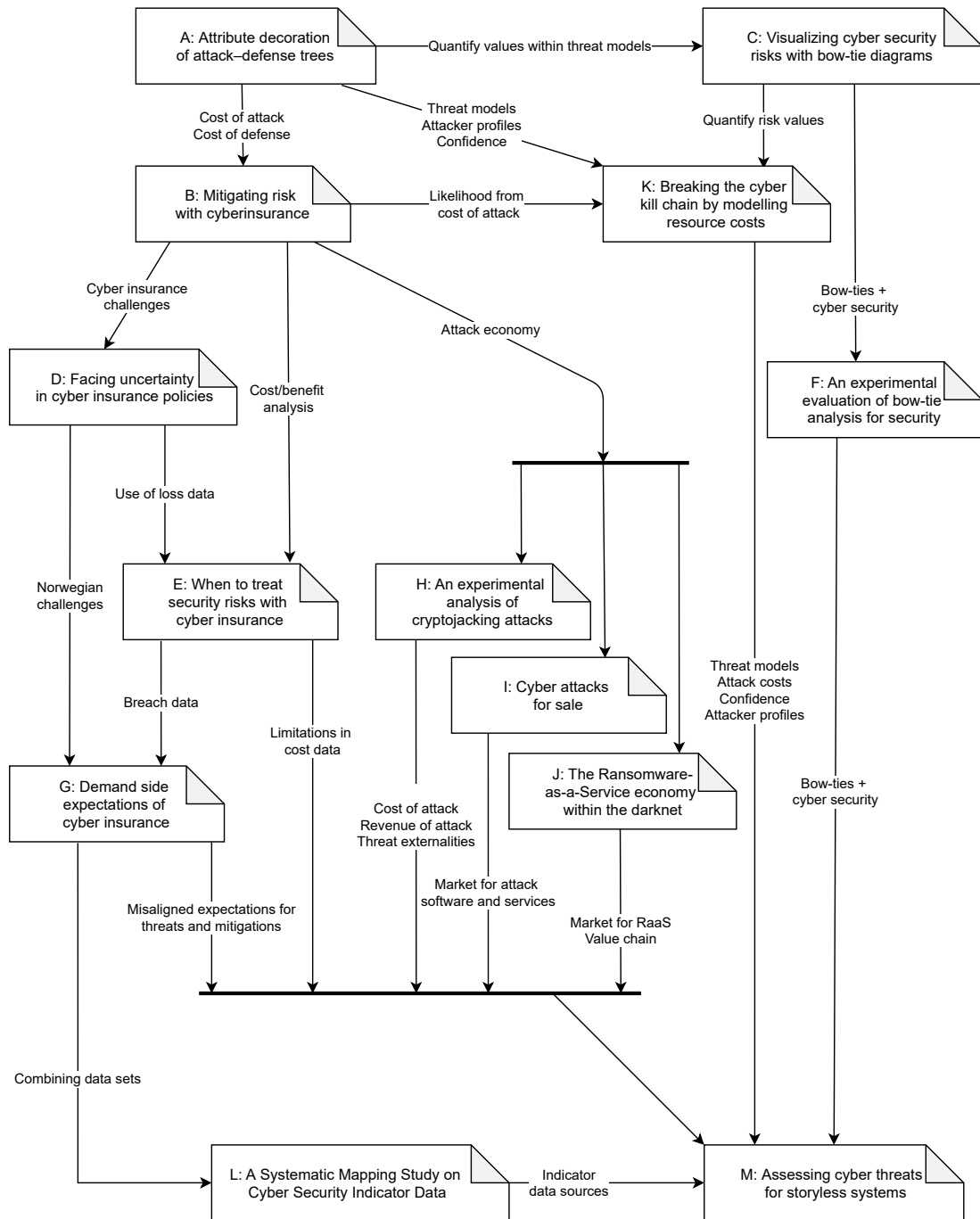
For each of the primary papers there is a dedicated section showing these role descriptions for all of the authors. My co-authors have approved and signed these descriptions through consent forms.

## 7.2 Primary papers contribution

The included primary papers do not follow a single evolutionary path. Instead, they can be thought upon a set of interrelated tracks in the landscape of the area of concern. They occasionally

cross paths, as well as provide parallel alternatives along the way, and the well-known proverb; “the journey matters more than the destination” has indeed been true for many of the studies.

Figure 7.1 illustrates a map of the primary papers in this landscape, showing how results from earlier papers have influenced the succeeding and how tracks are forked and merged. The following sections explain the terrain and main points of interest of the primary papers.



**Figure 7.1:** A high-level overview of the relationships between the primary papers contribution

### 7.2.1 A: ‘Attribute decoration of attack–defense trees’

*Attack-defense trees* (ADTrees) belong to a graphical modelling formalism used to express both attack methods and how they can be mitigated. The paper shows how the trees can be “decorated” with attributes that increase the expressiveness of such models. The main contents of the paper were:

- An explanation of the ADTree modelling elements (root, attack, defense nodes and their relationships).
- An overview of related work for both attack trees and ADTrees.
- A summary of which attributes have been used in related work.
- A case study where an ADTree is created for RFID-based goods management system, which was decorated with attributes and given values using a game-based approach.
- Results related to the perceived difficulties of assigning attribute values, modelling, interpretation and calculation.
- A discussion of the results and practical recommendations for the methodology.

The paper received the *Journal paper of the year* award from IGI Global (see Appendix C).

#### 7.2.1.1 Contribution to thesis

This paper was written in the aftermath of the EU-project SHIELDS (Grant Agreement No 215995) [262] and before the start of my PhD period. Still, it has been included as a primary paper since the use of modelling attributes for cost of attack and cost of defence were so inspirational to my continued research and PhD theme. It provides a scientific baseline for related work up to the publication date.

In this paper, we did not operate with specific cost values, but applied a linear order of relative values (*cheap < average < expensive < infinite*) and showed that this can be aggregated to e.g. the minimal expected attack cost. We also gained experience with *confidence* when assigning values, and how this could impact the reliability of the results. Another significant result was the possibility to generate attacker profiles based on attributes such as required *skill*, *attack cost* and *insider access*.

#### 7.2.1.2 CRediT authorship contribution statement

- **Alessandra Bagnato:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Resources, Validation – Verification, Visualization – Preparation, Writing – original draft, Writing – review & editing.
- **Barbara Fila (Kordy):** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation – Verification, Visualization – Preparation, Writing – original draft, Writing – review & editing.

- **Per Håkon Meland:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Validation – Verification, Visualization – Preparation, Writing – original draft, Writing – review & editing.
- **Patrick Schweitzer:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation – Verification, Visualization – Preparation, Writing – original draft, Writing – review & editing.

### 7.2.2 B: ‘Mitigating risk with cyberinsurance’

Cyber insurance was around 2015 emerging as a new risk management approach of transferring risk to a third party, and this paper highlighted challenges related to:

- The lack of technical experience and actuarial data on the insurer side.
- Problems companies are having documenting their security measures and interpreting insurance policies.
- The complexity in service chains and effects from externalities.
- Reporting security flaws and breaches.

Decisions to buy cyber insurance should be based on cost-benefit trade-offs. Making proper trade-offs is difficult especially for smaller businesses, which is also the place where the majority of cyber breaches occur.

#### 7.2.2.1 Contribution to thesis

This paper was written just before the start of my PhD period based on investigations from the *InSecurance* project [228] funded by SINTEF. I learned that the area of cyber insurance was particularly in need for improved ways of risk quantification, and therefore it became natural to extend this into the PhD work. This is regarded as primary paper since it:

- Surveyed existing academic and grey literature on the cyber insurance topic, which was relatively sparse at that time.
- Presented an approach to modelling and reasoning about treatment cost and benefit. Here, we argued that *annual loss expectancy* (ALE) should be a central factor in risk assessments.
- Exemplified how the different costs on both the defending and attacking side could influence the likelihood of attacks given that both parties are interested in making a profit (or avoiding loss). To properly do this, we advocated for an up-to-date overview of costs for performing attacks, as well as standardising indicators and metrics for predicting information security risks.

### 7.2.2.2 CRediT authorship contribution statement

- **Per Håkon Meland:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Visualization – Preparation, Writing – original draft, Writing – review & editing.
- **Inger Anne Tøndel:** Conceptualization, Investigation, Methodology, Visualization – Preparation, Writing – original draft, Writing – review & editing.
- **Bjørnar Solhaug:** Conceptualization, Investigation, Methodology, Visualization – Preparation, Writing – original draft, Writing – review & editing.

### 7.2.3 C: ‘Visualizing cyber security risks with bow-tie diagrams’

This paper presents a methodology for visualizing and assessing security risks by means of bow-tie diagrams. This particular notation was chosen due to its well-known familiarity from safety assessments among high-risk industries, such as oil and gas, mining, aviation, maritime and public health services. Using design science as a research methodology, the following research questions were addressed:

- How can bow-tie diagrams be extended to include security considerations in addition to safety considerations?
- How can the likelihood of cause and severity of cyber attacks be visualized in bow-tie diagrams?

Evaluation was done through analysis of descriptive, constructed use cases for maritime service scenarios. Our main conclusion is that adding security concepts to the bow-ties is a promising approach, since this is a notation that high-risk industries are already familiar with. However, their advantage as easy-to-grasp visual models should be maintained, hence complexity needs to be kept low.

#### 7.2.3.1 Contribution to thesis

This paper was rooted in the CySiMS project [223], which was developing a security solution for maritime communication systems. A special task in this project was to develop a risk assessment methodology for the maritime domain, which I was able to combine with my PhD research. This gave me access to maritime stakeholders and the opportunity to take part of the construction of new artefacts. The main contributions to the thesis were:

- Overview of related work for risk analysis, security modelling, safety modelling and considering safety and security in combination.
- Explanation of the real-world problems the maritime stakeholders are facing, such as low cyber security awareness and limited connectivity whilst sailing.

- A detailed explanation of the bow-tie notation and our proposed security extensions.
- A methodology for quantifying and visualising risk values for bow-ties.

### 7.2.3.2 CRediT authorship contribution statement

- **Karin Bernsmed:** Conceptualization, Investigation, Methodology, Validation – Verification, Writing – original, Writing – review & editing.
- **Christian Frøystad:** Conceptualization, Investigation, Methodology, Validation – Verification, Writing – original, Writing – review & editing.
- **Per Håkon Meland:** Conceptualization, Funding acquisition, Investigation, Methodology, Validation – Verification, Writing – original, Writing – review & editing.
- **Dag Atle Nesheim:** Resources, Validation – Verification, Writing – review & editing.
- **Ørnulf Jan Rødseth:** Funding acquisition, Resources, Validation – Verification, Writing – review & editing.

### 7.2.4 D: ‘Facing uncertainty in cyber insurance policies’

The paper presents the results from a qualitative interview study of ten Norwegian organisations, identifying how they perceive cyber insurance and in particular reasons for not buying. The following research questions were addressed:

- 1) What are the main uncertainty factors in the consideration phase as perceived by the demand side?
- 2) How can these uncertainties be reduced?

From the interviews we grouped uncertainties related to:

- The products themselves (including terminology, coverage, limit, premium).
- The process (how to assess threats and use cyber insurance as a risk transfer option, lack of cyber security knowledge).
- The support insurance companies would provide during and after an incident.

Our recommendation for reducing these uncertainties was to improve the awareness of coverage gaps, exclusions and loss during the negotiations/selection of policies.

#### 7.2.4.1 Contribution to thesis

This paper was a more specialised continuation of contribution B, digging more into the real-world problems that impede the use of cyber insurance. The following results were important for the thesis:

- An overview of related work on challenges for the demand side of cyber insurance. Evidence was gathered from Norway and compared with global observations to detect discrepancies.
- We got clear statements that risk managers and people with similar roles find it difficult to perform cost/benefit analysis for cyber security, and to have a good and dynamic overview of the relevant threats.
- We had obtained a dataset from Advisen dated November 2016 that contained 33 023 world-wide cyber loss events. We used this data to show that expected loss for certain cyber threats were not always reflected in cyber insurance coverage.
- We also showed that there are different types of loss to consider, such as response costs, economic loss, litigation cases and penalties/fines, and that these have different proportions.

#### 7.2.4.2 CRediT authorship contribution statement

- **Per Håkon Meland:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Visualization – Preparation, Writing – original draft, Writing – review & editing.
- **Inger Anne Tøndel:** Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing.
- **Marie Moe:** Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing.
- **Fredrik Seehusen:** Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing.

#### 7.2.5 E: ‘When to treat security risks with cyber insurance’

This paper presents a lightweight, data-driven approach for organisations to evaluate their own need for cyber insurance. A generic risk model, populated with available industry averages, is used as a starting point. Individual organisations can instantiate this model to obtain a risk profile for themselves related to relevant cyber threats. The risk profile is then used together with a cyber insurance profile to estimate the benefit and as a basis for comparing offers from different insurance providers.

In this approach, we have made use of real, available quantitative data concerning threats. This includes:

- A threat categorisation linked to costs/loss data (Advisen).
- Likelihood of threat incidents (UK Government).
- Distribution of threat incidents among industries/domain (Advisen).



- Frequency of breaches for different company sizes (UK Government)
- The size of different industry sectors (Fidelity Investments).

This paper is an extended journal version of paper P, which received the best paper award at the International Workshop of Cyber Insurance and Risk Controls (see Appendix C). The extension includes more details on the approach itself and calculation examples.

#### 7.2.5.1 Contribution to thesis

The development of this approach was motivated by the current practices and needs for cyber insurance decision making we described in contribution D. We aimed to address the problem of performing risk quantification for risk managers with limited expertise in cyber security and show how to take advantage of available information sources.

We highlight the need for better and updated data sets to enrich the generic risk model and create more accurate risk profiles. We analysed and compared several more data sources than we actually applied in our description, and observed that there very many deviations between the same measurements. Also, we argue for making available additional types of data that are not openly found, but would improve this and similar approaches. Among these are data on attacker costs.

#### 7.2.5.2 CRediT authorship contribution statement

- **Per Håkon Meland:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Resources, Software, Validation – Verification, Visualization – Preparation, Writing – original draft, Writing – review & editing.
- **Fredrik Seehusen:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Validation – Verification, Visualization – Preparation, Writing – original draft, Writing – review & editing.

#### 7.2.6 F: ‘An experimental evaluation of bow-tie analysis for security’

The objective of this paper was to gain empirical knowledge on the use of bow-tie analysis applied for cyber security. The central research hypothesis was that the bow-tie notation has a suitable expressiveness for security as well as safety. If we could falsify this for cyber security, then it would make no sense to pursue application with both safety and security in combination.

The study uses a formal, controlled quasi-experiment on two sample populations – security experts and security graduate students – working on the same case concerning digital exams. We compared the results with a similar experiment applying misuse case analysis on a Web

shop case. Misuse case analysis is an established technique for graphical information security modelling, but has also been used to assess safety hazards.

The results show that the collective group of graduate students, inexperienced in security modelling, perform similarly as security experts with a well-defined scope and familiar target system/situation. The bow-tie notation did not seem like an obstacle for expressing cyber security threats and consequences, which gives support to our hypothesis. Comparing results with misuse case modelling, there is no reason to believe that misuse case models outperform bow-tie diagrams in a security context if we consider content generation made by inexperienced users.

### 7.2.6.1 Contribution to thesis

To validate the methodology of paper C, we performed a controlled experiments with a large sample of NTNU students and a smaller sample of security experts. The focus was on the security modelling concepts, and the experiments did not go so far as to try and quantify values for them. Additional contributions were:

- Updated overview on the use of bow-ties, as well as models covering safety and security. We also explained misuse cases, their history and compared them to bow-ties.
- Identified future modelling tool features and research directions.
- The experiment results advocate for a combination of people involved when creating security models. This is to ensure both technical depth and broadness of threats.

This paper was an extended journal version of paper S, adding data from previous years (more students) and a comparison with misuse case analysis to make the results more valid.

### 7.2.6.2 CRediT authorship contribution statement

- **Per Håkon Meland:** Conceptualization, Data curation, Investigation, Methodology, Software, Validation – Verification, Visualization – Preparation, Writing – original draft, Writing – review & editing.
- **Karin Bernsmed:** Validation – Verification, Writing – review & editing.
- **Christian Frøystad:** Validation – Verification, Writing – review & editing.
- **Jingyue Li:** Resources, Supervision, Writing – review & editing.
- **Guttorm Sindre:** Supervision, Writing – review & editing.

### 7.2.7 G: ‘Demand side expectations of cyber insurance’

The purpose of this study was to examine the expectations that early and prospective customers have towards cyber insurance, and see if these are in line with contemporary incidents and

claims. Using qualitative interviews with Norwegian and Swedish organisations, we sought to identify misaligned expectations and discrepancies between industry domains. The expectations expressed in the interviews are compared with reports describing recent incident claims, claims statistics from 2018, as well as data breach statistics for different domains and a few cyber insurance loss scenarios. To guide the study, we defined the following research questions:

- 1) Are there different expectations in different business domains?
- 2) Are there discrepancies between coverage expectations and the costs of prevalent incidents as seen in incident data?
- 3) Are there discrepancies between coverage expectations and the costs of prevalent incidents as seen in scenarios?

The results show no obvious pattern of discrepancies between different domains. However, informant expectations on business interruption coverage are much greater than one would expect from its share of claims. This skewed expectation could be explained by the influence of prominent scenarios found in a number of recently published reports.

#### 7.2.7.1 Contribution to thesis

This work can be seen as a continuation of paper D, where the same interview guide was used for a number of Swedish organisations. We also combined the work with real incident data that we obtained from the insurance agent Willis Towers Watson, as well as the dataset provided by Advisen (also used in paper E). This contributes to how data from different sources can be combined to give indications on threat and attack trends.

#### 7.2.7.2 CRediT authorship contribution statement

- **Ulrik Franke:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Validation – Verification, Visualization – Preparation, Writing – original draft, Writing – review & editing.
- **Per Håkon Meland:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Validation – Verification, Visualization – Preparation, Writing – original draft, Writing – review & editing.

#### 7.2.8 H: ‘An Experimental Analysis of Cryptojacking Attacks’

This paper presents an experimental analysis of how different types of cryptojacking attacks impact a selection of consumer-grade devices, and the perceived annoyance by the user. We tried to address the following research questions:

- 1) How is performance on different types of devices affected by cryptojacking measured objectively and perceived subjectively?
- 2) What are the expected revenues and costs for the attacker based on the targeted devices?

Around 2018, cryptojacking was the new “big thing” among malware threats, however this trend dropped quite suddenly shortly after. The results from our experiment can be used to explain this decline and why the previously much feared cryptojacking threat is now practically gone. We show that even though the cost of attack is quite low, the revenue of this particular way of exploiting devices is just not worth it. The attacker is more likely to use more invasive methods such as ransomware based on the principle of opportunity costs. The market failure of cryptojacking is mostly related to externalities – namely the dramatic drop in the general cryptocurrency market and changes in the mining algorithm used by the monero currency.

#### **7.2.8.1 Contribution to thesis**

This paper focuses on one particular type of threat which is highly motivated by profit on the attacker side. Firstly, we show that cryptocurrency mining experiments can be used to calculate potential revenue based on the type and number of infected devices. This is valuable information when assessing the ways systems are likely to be exploited. Secondly, we prove that the concept of externalities is just as important on the attacker side as the defender side.

#### **7.2.8.2 CRediT authorship contribution statement**

- **Per Håkon Meland:** Conceptualization, Formal analysis, Investigation, Validation – Verification, Validation – Verification, Writing – original draft, Writing – review & editing.
- **Bent Heier Johansen:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation – Verification, Validation – Verification, Writing – original draft, Writing – review & editing.
- **Guttorm Sindre:** Supervision, Writing – review & editing.

#### **7.2.9 I: ‘Cyber Attacks for Sale’**

This study presents an online netnography study of eleven contemporary darknet marketplaces, addressing the following research questions:

- 1) What kind of cyber attack items are available on the darknet marketplaces?
- 2) What are the most profitable items for the vendors?

The results have been used to create a detailed categorization of items, showing a distribution based on item type and availability. This has been compared to the number of sold items and

revenue from four of the marketplaces, and we discuss these different views. Aided by related studies, we have identified trending cyber threats such as phone hacking, information theft and Bitcoin stealing.

#### 7.2.9.1 Contribution to thesis

Where there is a demand there will be a supply, and darknet marketplaces do supply software and services to threat agents in need of a cyber arsenal. Insights into the cybercrime economy can give an indication of the type and capabilities of attackers, what assets they are targeting and which vulnerabilities they are likely to exploit. We provide an inventory of available items and map these towards availability and popularity. Compared to previous work, our study provides an up-to-date analysis and more detailed categorisation of items. We also take scam items into account, and separate between indicators related to availability and sales.

#### 7.2.9.2 CRediT authorship contribution statement

- **Per Håkon Meland:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Validation – Verification, Visualization – Preparation, Writing – original draft, Writing – review & editing.
- **Guttorm Sindre:** Supervision, Writing – review & editing.

#### 7.2.10 J: ‘The Ransomware-as-a-Service economy within the darknet’

Ransomware has been recognized as one of the fastest growing cybercrimes in recent history. On the darknet markets, *Ransomware-as-a-Service* (RaaS) is being offered as a franchise model that allows people without programming skills to become active attackers and take part in the ransomware economy.

We have studied contemporary darknet markets and forums over a period of two years using a netnographic research approach. Our observations have been complemented with historical data found in archives and published interviews with stakeholders involved in darknet operations. The two main research questions we have tried to address using our findings are:

- 1) How severe is the RaaS threat?
- 2) What are the value chains related to this market?

Regarding the former, we show that the RaaS threat currently seems more modest than indicated in the media and reports from security companies. There are now relatively few RaaS items offered for sale in the most popular darknet marketplaces, and the number of successful sales does not indicate a large economy.

In order to answer the latter research questions, we have used marketplace observations, forums posts, available interviews and literature to map stakeholders and value chains in the ransomware economy. Due to anonymity and ample amounts of fraud, this is an volatile environment that lacks trust relationships. However, market strategies and mechanisms are constantly changing, driven by profit opportunities.

#### **7.2.10.1 Contribution to thesis**

This paper is a continuation of the secondary paper Q, expanding the study timeline, marketplaces and analysis of archive data. This paper proves that observations from the cybercrime economy are relevant when assessing this kind of threat. We combine a quantitative analysis on availability and sales, and a more qualitative analysis on the advertisements, reviews and discussions. The value chain can be useful when trying to break the underground economy behind ransomware and subsequently mitigate this cyber threat. Additional contributions include:

- Background information about darknet, dark web, marketplace, forums and ransomware.
- A survey of related work on marketplace and forum research, value chains and the economics of ransomware.
- Evidence of vendor resilience and survivability despite of numerous law enforcement take-downs.
- The trend and market size of RaaS compared to other (more profitable) threats.
- Evidence of fraud between cyber criminals.

#### **7.2.10.2 CRediT authorship contribution statement**

- **Per Håkon Meland:** Conceptualization, Investigation, Writing - original draft, Visualization, Resources, Data curation.
- **Yara Fareed Fahmy Bayoumy:** Conceptualization, Methodology, Investigation, Data curation.
- **Guttorm Sindre:** Supervision, Writing - review & editing.

#### **7.2.11 K: ‘Breaking the cyber kill chain by modelling resource costs’**

All cyber attacks require resources before they become a reality. In this paper, we hypothesize that during threat analysis, it is possible to reduce the complexity of the resource requirement to a monetary concern, complemented by a limited set of attacker characteristics. This will allow us to identify the potential offenders and come up with technical and non-technical mitigations that will significantly increase the attacker costs.

Using design science, we have developed a modelling approach and supportive tool that maps resource costs to different stages of a cyberattack. This tool is able to show calculations interactively and extract potential offenders based on a built-in library from available cybercriminal profile literature.

Evaluation results from security researchers, security industry experts and maritime domain specialists show that breaking down costs gives a higher confidence of the total costs, but this requires a limited set of attack paths in the threat model itself. Furthermore, our approach improves understanding of attacks and how they can be mitigated.

#### 7.2.11.1 Contribution to thesis

This paper builds on a number of results from previous contributions, and shows one realisation of threat quantification built around a cost model instead of historical incident data. The paper also contributes with:

- An overview of literature related to cyber attack chains or stages.
- An overview of the use cost data in attack trees.
- Related work on cybercriminal profiling.

Secondary paper contributions related to maritime communication and e-navigation have been vital to build real-world threat models that could be used for evaluation. The results have subsequently been used to assess the security of the maritime solution, which again has an impact of the safety related to ships and the surrounding environment.

#### 7.2.11.2 CRediT authorship contribution statement

- **Kristian Haga:** Conceptualization, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.
- **Per Håkon Meland:** Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing.
- **Guttorm Sindre:** Supervision, Writing – review & editing.

#### 7.2.12 L: ‘A Systematic Mapping Study on Cyber Security Indicator Data’

A security indicator is a sign that shows us what something is like or how a situation is changing and can aid us in making informed estimations on cyber risks. There are many different breeds of security indicators, but unfortunately, they are not always easy to apply due to a lack of available or credible sources of data. In this paper, we undertake a systematic mapping study on the academic literature related to cyber security indicator data. Systematic mapping is a methodology that is concerned with structuring a research area in order to give a broad overview,

showing concentrations of effort and revealing areas that need more attention. After a thorough search and screening procedure, 117 primary studies from the past five years were identified as relevant to answer our research questions:

- 1) What is the nature of the research using security indicators?
- 2) What is the intended use of the data?
- 3) What is the origin of the data for the indicators?
- 4) What types of the data are being used?
- 5) What is the data content of the indicators?

The primary studies have been classified according to a set of categories related to research type, domain, data openness, usage, source, type and content. Our results show that the research community is eagerly developing new methods and techniques that use indicators to support security decisions. There is still a need to take many of these from the conceptual plane, through empirical evaluation increase maturity and make them practical enough for real-world application. Indicators that are rather technical in nature can give valuable information about the contemporary cyber risk, while the increasing usage of unconventional data sources and threat intelligence feeds of more strategic and tactical nature represent a more forward-looking trend.

#### **7.2.12.1 Contribution to thesis**

This work was performed in relation to the H2020 project CyberSec4Europe (Grant Agreement No. 830929) [225]. The mapping study identifies existing work related to security indicators that could be used within threat models and identifies potential gaps. The main contributions in relation to the thesis work were:

- An overview of related survey work on security indicators.
- A classification scheme for security indicator data.
- A mapping of academic literature towards the scheme.
- A discussion on the trends and actual application of security indicator data.

#### **7.2.12.2 CRediT authorship contribution statement**

- **Per Håkon Meland:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Supervision, Validation – Verification, Visualization – Preparation, Writing - original draft, Writing – review & editing.
- **Shukun Tokas:** Data curation, Formal analysis, Investigation, Resources, Validation – Verification, Visualization – Preparation, Writing - original draft, Writing – review & editing



- **Gencer Erdogan:** Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation – Verification, Visualization – Preparation, Writing - original draft, Writing – review & editing
- **Karin Bernsmed:** Conceptualization, Data curation, Investigation, Writing - original draft, Writing – review & editing
- **Aida Omerovic:** Conceptualization, Data curation, Investigation, Methodology, Project administration, Writing - original draft, Writing – review & editing

### 7.2.13 M: ‘Assessing cyber threats for storyless systems’

This paper presents a systematic approach for assessing threats for storyless systems. The goal has been to develop something that can be readily applied in real-life projects, being efficient in terms of resource usage and flexible enough to be adjusted to the best data available. Quantifiable conditions are determined from the environment in which the system will reside and operate within, that is the availability of potential threat actors, their opportunities of performing attacks, the required means that are needed for the attack to succeed, and motivation factors. Through a case study performed in relation to a maritime system development project, we have sought answers to the following research questions:

- 1) How can we estimate threat likelihood for a new design?
- 2) What are the perceived advantages and disadvantages of such an approach?

The results show that representative participants from the cyber security and maritime community gave positive and consistent scores on the features, and regarded time usage, traceability of the threat assessment and the ability to indicate underlying uncertainty to be very appropriate. The approach has been proven useful for this domain and should be applicable to others as well, but the template requires up-front investments in gathering knowledge that is relevant and reusable in additional context situations.

#### 7.2.13.1 Contribution to thesis

This paper combines and extends different results from other primary papers, such as resource cost modelling (K), bow-tie analysis (F,C), indicator data (L) and attacker costs (H,I,J), to form a practical approach for cyber threat assessment. Thus, the contribution to the thesis can be seen as a defragmentation of results, application to one of the problem areas (maritime) and an evaluation with real end users. The work was done in relation to the CySiMS-SE project [223] and H2020 project CyberSec4Europe (Grant Agreement No. 830929) [225].

**7.2.13.2 CRediT authorship contribution statement**

- **Per Håkon Meland:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Funding acquisition.
- **Karin Bernsmed:** Conceptualization, Methodology, Writing - Review & Editing, Funding acquisition.
- **Dag Atle Nesheim:** Validation, Investigation, Resources, Writing - Review & Editing, Project administration, Funding acquisition.
- **Guttorm Sindre:** Writing - Review & Editing, Supervision.

**7.3 Secondary papers contribution**

As mentioned in Chapter 1, the secondary papers have been instrumental for situational inquiry and gaining practical domain knowledge. Figure 7.2 illustrates which problem areas (described in Chapter 3) they belong to, their grouping and proximity to the “core” of the PhD thesis work. The papers are also listed in Table 7.1, with a brief description of their relevance and contribution.

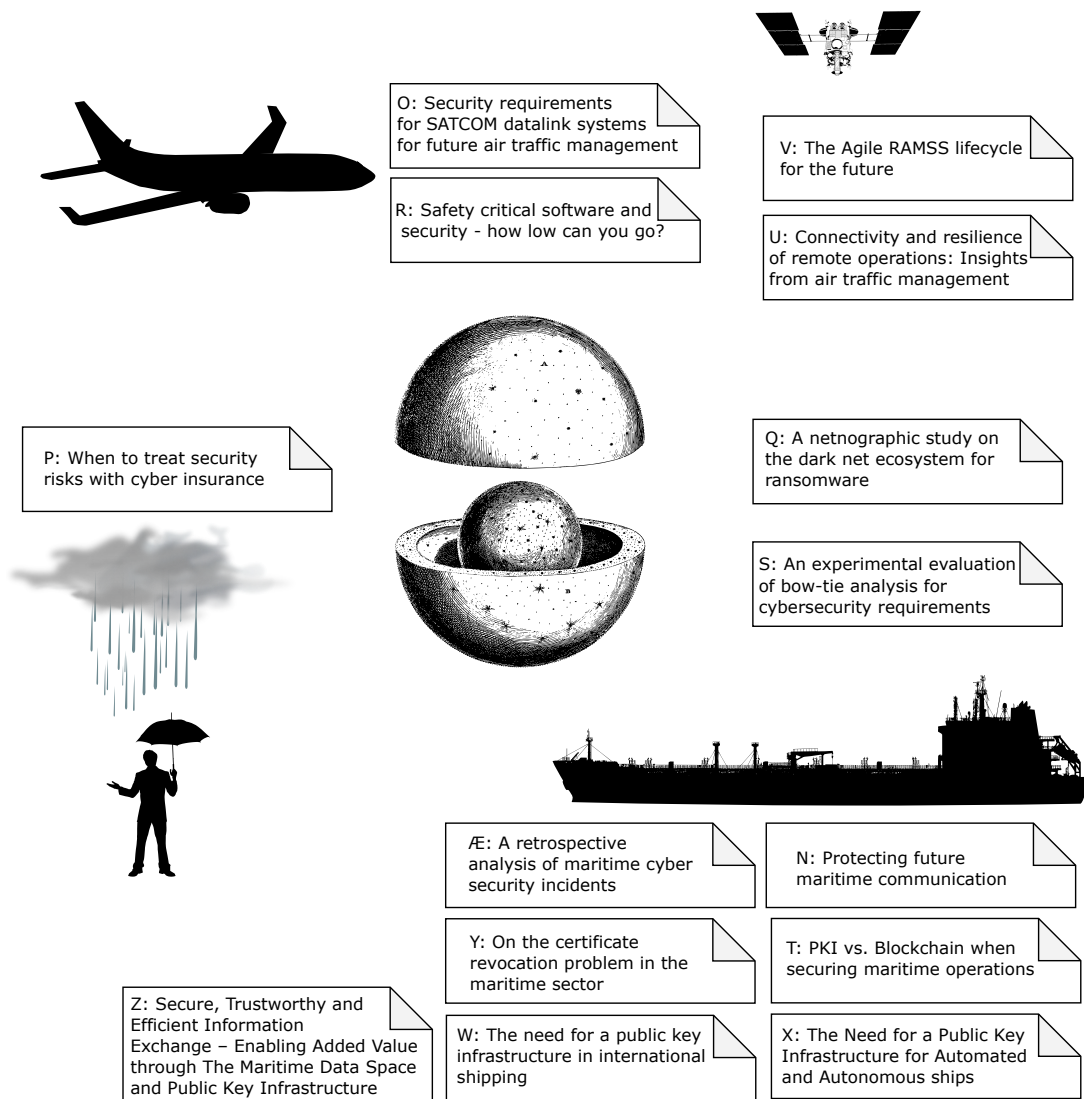


Figure 7.2: A high-level overview of the secondary papers contribution

**Table 7.1:** Summary of secondary papers

ID: Title	Relevance and contribution
N: ‘Protecting Future Maritime Communication’	Explains security needs related to new digital services for maritime communication. The suggested solution emerged from our risk assessment in the MAROFF RCN project CySiMS. This is a supportive paper since the work gained valuable insight into the security and safety challenges for the maritime domain.
O: ‘Security requirements for SATCOM datalink systems for future air traffic management’	Similar contribution as the paper above but set in the aviation domain. The main results originate from a <i>European Space Agency</i> (ESA) project on the development of a secure communication channel between the cockpit and air traffic controllers. We were involved with multiple users and got to know practical concerns and challenges.
P: ‘When to Treat Security Risks with Cyber Insurance’	The initial version of paper E which presents an approach where company risk profiles can be created using available dataset related to monetary loss, company size and industry.
Q: ‘A netnographic study on the dark net ecosystem for ransomware’	The paper includes our first experience with netnography as a research method. We gathered qualitative data from darknet marketplaces and forums to gain an understanding of the market and anonymised social interactions. The continuation of this study eventually lead to paper J.
R: ‘Safety Critical Software and Security - How Low Can You Go?’	Another supportive paper from the aviation domain, focusing on strength and weaknesses of developing software for a high reliability system.
S: ‘An experimental evaluation of bow-tie analysis for cybersecurity requirements’	This is the initial version of paper F, validating the bow-tie approach presented in paper C with a large sample of NTNU students and a smaller sample of security experts.

*continues on next page*

- 
- T:** ‘PKI vs. Blockchain when securing maritime operations’ A supportive paper with situational inquiry within the maritime domain. We compared the application of PKI and blockchain technologies given the limitations at sea. As comparative examples we used nautical safety information, port state reporting and ship certificates.
- U:** ‘Connectivity and resilience of remote operations: insights from air traffic management’ Situational inquiry on cyber threats related to the emerging remote tower concept for aviation. This work was also an opportunity to learn more about the resilience domain, especially the NIST publications on developing cyber resilient systems [47].
- V:** ‘The Agile RAMSS life-cycle for the future’ The work involved situational inquiry at the intersection between safety and security standards when developing safety-critical systems.
- W:** ‘The Need for a Public Key Infrastructure in International Shipping’ The paper presents maritime service use cases and their security needs. The work provided insight into the maritime business constraints, and defines cyber threats aimed at maritime communication and what resulting unwanted events could be. The results were used as a source for maritime threat modelling in papers **K** and **M**.
- X:** ‘The Need for a Public Key Infrastructure for Automated and Autonomous ships’ The paper describes thirteen use cases for maritime services for autonomous and automated ships and analyse how a PKI system can provide security barriers to mitigate relevant cyber threats and possible consequences of unwanted events. Similarly to the paper above this work has been useful for maritime threat modelling.
- 

*continues on next page*

---

---

<p><b>Y:</b> ‘On the Certificate Revocation Problem in the Maritime Sector’</p>	<p>The paper provides an analysis of certificate revocation techniques based on how they fulfil fundamental maritime requirements and simulated usage over time. The work provided insight into real-world issues with security solutions.</p>
<p><b>Z:</b> ‘Secure, Trustworthy and Efficient Information Exchange – Enabling Added Value through The Maritime Data Space and Public Key Infrastructure’</p>	<p>The paper presents an ecosystem for secure, trustworthy and efficient data transfer and information exchange between maritime stakeholders. It combines the results from the <i>Maritime Data Space</i> (MDS) and Cyber Security in Merchant Shipping (CySiMS) projects and provides valuable domain knowledge.</p>
<p><b>Æ:</b> ‘A Retrospective Analysis of Maritime Cyber Security Incidents’</p>	<p>The paper analyses and gives an overview of 46 maritime cyber security incidents from the last decade (2010-2020). The characteristics have been used to create a Top-10 list of maritime cyber threats. The results show that the maritime sector typically has incidents with low frequency and high impact, which makes them hard to predict and prepare for based on historical data. We also infer that different types of attackers use a variety of attack points and techniques, hence there is no single solution to this problem.</p>

---

## 7.4 Posters contribution

The primary purpose of the posters have been to present the early ideas of the research in a visually appealing way. The posters have thus facilitated conversations with peers attending these events. Table 7.2 gives a brief overview of the relevance and contribution of the posters, which both can be found in Appendix B.

**Table 7.2:** Summary of posters

ID: Title	Relevance and contribution
Ø: 'Combining threat models with security economics'	The poster explains the rationale of the initial version of the main research question. An extended abstract has been published as a part of the conference proceedings. The contents were based on my research plan.
Å: 'Resilient cyber security through cybercrime market analysis'	This work proposes how data from the cybercrime economy can be applied within the context of cyber resiliency strategies (anticipate, withstand, recover, evolve). An extended abstract has been published as a part of the conference proceedings. The poster and abstract are contributions to risk management and the main research question.

## DISCUSSION

*I know that I know nothing*

---

Socrates

The purpose of this PhD study has been to investigate new methods for managing cyber security risks without too much reliance on historical events. This chapter includes a synthesis of the results from the primary papers and discuss these contributions as a whole. It is organised according to the research questions defined in Section 4, and contains interpretation of the major findings, their implications, limitations and relationship to the literature and recent events. There is a dedicated section on ethical issues, followed by opportunities for future work.

### 8.1 Addressing the research questions

The main research question has been stated as:

**Main RQ:** How can modelling threats and economic incentives improve cyber risk management?

It is not a question that can be answered in a binary or quantitative way, but seeks solutions forged from an approach based on design science and practice research principles as explained in Section 5. A high-level summary of the answer is that the primary papers describe techniques and tools that:

- Help identify threats and the likelihood of their occurrence, and consequently the way we assess risks.

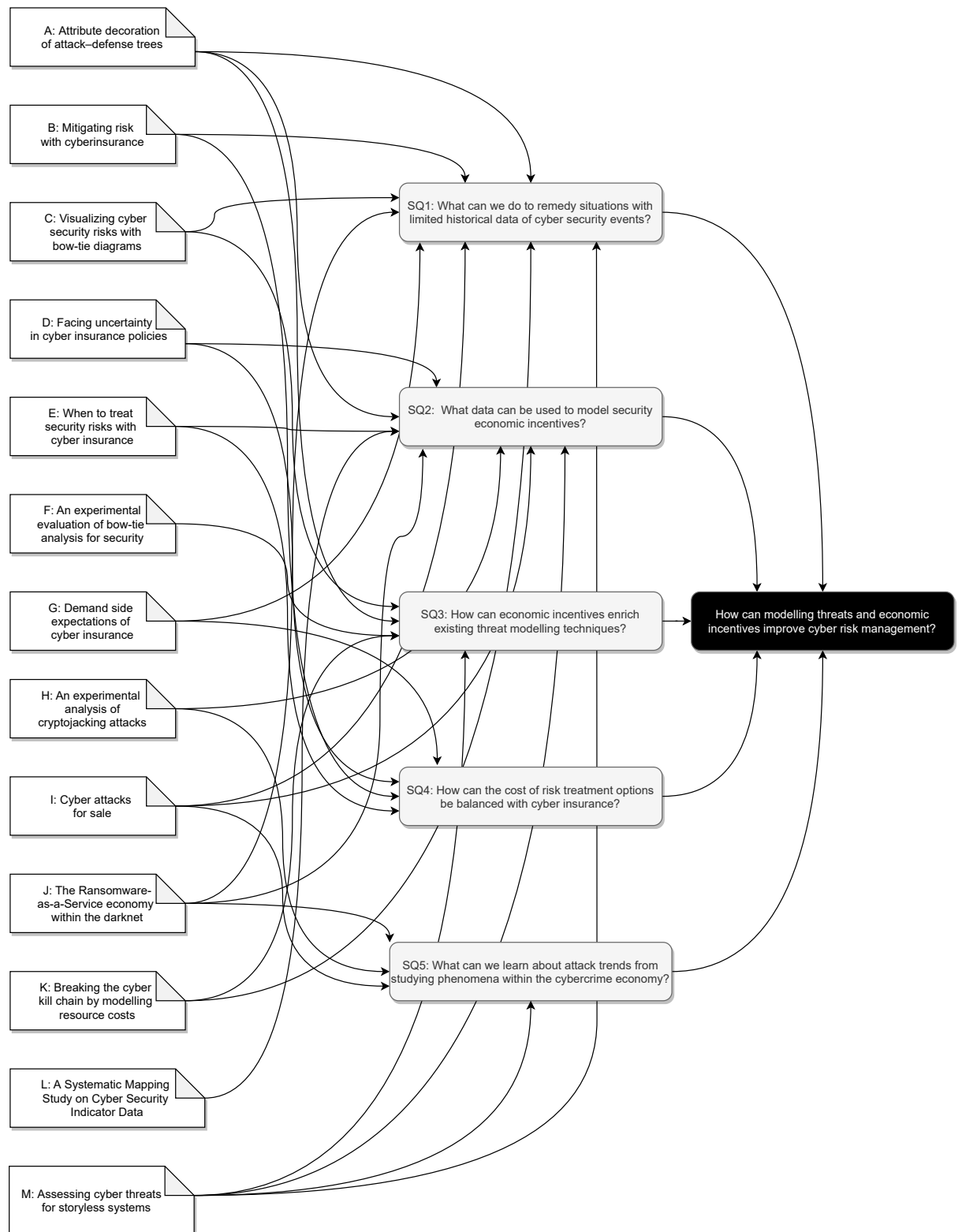


- Help choose among risk treatment options, especially balancing off with cyber insurance to cover residual risk.

In both cases, we have used the combined knowledge of security experts and domain specialists, and shown how they can be supported by new and different types of data sources. There is not really a one-size-fits all solution here, and this implies that in a real-world setting, the strategy should be adjusted according to available human resources, existing risk traditions within the sector and what data that actually exists and is relevant for the given system. As Husák et al. [101] have already observed in their literature survey on cyber attack projection, prediction, and forecasting; “popular datasets are old, unreliable, and created for other purposes”, and thus potential pitfalls.

During the course of this PhD study, there have been other researchers on the same quest that have taken similar paths. For instance, Hubbard [263] has proposed the HTMA approach (*how to measure anything*) for cyber security risks, which heavily relies on subjective expert opinions. In 2019, Santini et al. [264] extended upon this, adding more objective data from several sources to progressively improve the risk model. These *key risk indicators* (KRIs) were mainly based on measurements internal to the organisation, such as malware, vulnerabilities, data breaches and deep web exposure. In 2020, Figueira et al. [265] proposed a mixed qualitative-quantitative risk analysis approach, where they use regression models instead of data about the past to compute future threat probability. Similarly to Santini et al., they base their estimations on currently known system vulnerabilities. Kissoon [266] also applies regression models to measure the effectiveness of current implemented cyber security measures in organisations. She uses internal variables such as risk appetite, security budget and loss after security breach obtained from surveys and interviews. Paté-Cornell et al. [7] have through a set of case studies presented several ways to gather and use the information available to quantify cyber risk. For extreme events without data, they used probabilistic analysis of potential scenarios where the limits of statistical data are completed by expert opinions. Among the data were potential points of access, vulnerabilities, software update time and the costs/loss after successful attacks. In 2021, Al-Hadhrami et al. [267] proposed to use subjective logic within attack trees to compensate for the lack of accurate, probabilistic data. The subjective opinions were based on two criteria; a vulnerability level and technical difficulty of an attack.

More precise answers are given for the associated sub-questions we defined on page 37, and furthermore detailed for the paper-specific questions. Figure 8.1 shows the actual instantiation of Figure 4.1, showing the hierarchy between the different levels of research questions. The term “spaghetti-diagram” can be used to describe it. The individual arrows are not that important (and difficult to follow), instead primary purpose of the diagram is to show that there are several links from each paper to different sub-questions.



**Figure 8.1:** A spaghetti-diagram showing how the primary papers relate to the research questions

The answers to the sub-questions are discussed in the sections below. These should be seen in the context of the main research question and are somewhat interdependent, meaning that one answer sets the premises of another question.

### 8.1.1 SQ1: What can we do to remedy situations with limited historical data of cyber security events?

This question represents the *Holy Grail* of cyber risk analysis, where the ultimate answer is sought by many and yet to be found. Cyber security is uncertain by nature, especially when it is storyless, i.e. dealing with new technology development and/or industry sectors that are undergoing rapid digitalisation. The type of historical data we would like to have tied to the target organisation would typically be *attack frequency*, *attack type distribution*, *number of successful attacks*, *number of prevented attacks* and *loss per attack* (both successful and failed). But even with good records of this, the past is not always a good source to predict the future. Rational attackers, especially *advanced persistent threats* (APTs), would also use history to their advantage and not waste their efforts on attempts that are likely to fail. Storyless systems, such as the ones we have been working with within aviation and maritime, may therefore have a narrow time window of peace. After all, the number of incidents in both of these sectors has been relatively low so far. However, all good things must come to an end, and we have to expect that these systems will be increasingly targeted.

The primary objective of threat modelling is to anticipate attacks, often taking the perspective of the attacker. In the case of paper A, we studied how attack-defence trees could be extended with different types of data. The tree nodes contained a combination of attributes not tied to past events. The tree itself and attribute values were assigned in a game-based way by experts and specialists playing the role as either attacker or defender, and we show how to combine and aggregate the values in a bottom-up approach. Based on experiences from a case study, we encouraged the involvement of domain specialists in order to give accurate estimates, and the use of consensus meetings when discrepancies arise. Still, we encountered what we defined as conflicting modelling goals that have practical implications on the quality of the risk analysis, namely:

- **Time** - Creating models and involving experts is time consuming, and the most suitable people are not always available.
- **Reusability** - Though the model structures can be reused between projects and thus save time and resources, actual attribute values are probably less reusable as they are more context dependent.
- **Accuracy** - Fine grades of values give more accurate results than coarse ones, but require more work to assign and specific expertise.

- **Simplicity** - The modelling methodology should be easily understandable by non-experts so that there are more people that can potentially contribute to the value estimations.

Most of these goals are in line with the wider survey on graphical security models Hong et al. [91] published in 2017, which pointed to common practical challenges related to scalability when creating complex models, reusability and tool availability.

Similar strategies of using experts and specialist to identify threats and making estimations were used in papers C, K and M, however applied to other types of threat modelling techniques (see SQ3) and working specifically with risk scenarios from the maritime sector.

There will always be limitations associated to subjective opinions. With paper A, we experienced challenges related to expert calibration and coming to an agreement on metric values, while we experienced the opposite in the experiments identifying threats, controls and consequences we conducted in paper F. Holm et al. [268] have highlighted the uncertainty in data quality when expert judgment is used, and in their experiments they could for instance see both a significant negative correlation and a strong positive correlation between experience and calibration, suggesting that additional years' experience can both decrease and increase the calibration.

In paper F, we were able to provide some evidence that a large group non-experts perform similarly as a smaller group of security experts in identifying threats when there is a well-defined scope and familiar target system/situation. However, the teams of non-experts were significantly larger than the number of experts. There was also a difference between the type of threats, showing the positive effects of combining the results of people with different background. However, it is difficult to say what the optimal composition should be. Both the experts and non-experts had a technology background. According to Falco et al. [269], the cross-disciplinary nature of cyber risk also implies that the needed expertise should not be limited to computer security/science, but include fields such as behavioural science, economics, law, management science and political science. Of course, this should be determined as part of the context establishment activity as defined in ISO/IEC 27005 [57] (see page 11). If the scope is limited to a technical system activity, the required diversity of people involved may be more relaxed.

It was interesting to see that our performance measurements on threat identification are similar to how professionals and students perform in software engineering (e.g. [270–272]). In contrast, the findings of Hallberg et al. [273] show generally low consensus values both among cyber security experts and non-experts, but that was related to ratings of probability and severity of incidents, and supports the findings of Holm et al. [268] already mentioned above.

Our experts worked on the models individually, while the students collaborated in small teams. This configuration choice can be adjusted, but that was not something we experimented with.

Though teams can provide more diverse thinking, there is always the danger of *groupthink*<sup>a</sup> or domineering that could have the opposite effect.

To limit the subjectiveness of threat estimations, we have sought for data-driven methods that can be used to support quantification. Within security economics, it is practices related to economic threat modelling (see page 17) that has been our main focus. This requires a way of determining if and who would benefit from an attack and making that unattractive (low utility). A similar path can be seen in a series of papers by Knez et al. [276], Llansó et al. [277], McNeil et al. [278], that describe a *capability-based approach* to cyber risk management for space missions. They criticize the required amount of labour that is needed to describe attack paths and give estimation on likelihood and impact, emphasizing that estimations are too subjective and do not scale well for complex systems. Instead, they suggest that mitigations (or modifications using ISO-terminology) should be based on representations of presumed offensive capabilities of attackers and the defensive capabilities. Still, this and our approach requires some knowledge about the attacker, which is further detailed under **SQ2** and **SQ5**.

After a risk assessment, we still need to decide upon risk treatment options. Here, risk modification is the main priority when there are unacceptable risks. With both attack-defence trees and bow-tie models we found it useful to include both preventive and reactive controls, and not follow a strict stepwise process when creating them. However, a storyless system makes it difficult to determine what the right mixture should be. In both aviation and maritime, the primary goal is to preserve the mission even during periods of distress. Shutting down a plane mid-air is not really an option. In these situations, it might be better to turn to more a reactive cyber security strategy, and follow principles of cyber resiliency (as presented in poster Å). This is in accordance with what Anderson et al. [279] concluded in 2012, and once more in 2019, that it is often economically rational to spend less in anticipation of cybercrime and more on response. Tundis et al. [200] have a similar opinion that leans towards risk retention, stating that “it is simply not practical to implement counter-measures in a timely and economical manner for all possible attacks”. Falco et al. [269] argue that cyber security can be characterized as a public good and this can lead to under-investments for individual organisations. With **SQ4** we look more into investing in cyber insurance for risk sharing.

### 8.1.2 SQ2: What data can be used to model security economic incentives?

As shown in Section 2.2, there are many approaches to modelling security economic incentives found in the literature. Whilst models for optimal security investments are more tied to risk

---

<sup>a</sup>The term was established by Janis [274], based on the dystopian novel *Nineteen Eighty-Four* by George Orwell [275]. It is used when a group of people set aside their personal beliefs or uncritically adopt the opinion of the rest of the group to reach a consensus.

treatment options, the ones belonging to econometrics of wickedness are more useful for risk assessments following the rational attacker's paradigm. Table 8.1 gives an overview of the incentives that have been applied in our primary papers. We separate between data type and source, where the former defines the metrics and the latter the source of the actual values of the data. The terminology is not consistent throughout the different papers (nor in the general literature), hence there is a description the meaning of the data.

**Table 8.1:** Economic incentive data used in primary papers

Data type (paper)	Description	Data source (paper)
Defender investment (A, B, D, E)	Up-front monetary requirement for preventive types of risk modification, e.g., equipment, software, insurance premium and awareness training. Independent of successful attacks.	Expert opinion (A)
Defender reactive cost (B, D, G)	Monetary requirement for reactive/detective risk modification and recovery, e.g., extra personnel, consultancy, forensics investigation and backup restoration. Dependent on attack attempts, and not whether the attacker succeeds or not.	Cyber loss events (D, G)
Defender loss (A, B, C, E, G)	The economic severity or consequence from the system owner's point of view given successful attacks, including defender reactive costs and further examples such as damages, deaths, litigation cases, notification costs and business disruption.	Expert opinion (A), Public statistics (C), Cyber loss events database (D, E, G), Public scenario descriptions (G), Incident claims reports (G)
Defender reimbursement (B, D, E, G)	An economic compensation that a defender could receive given a successful attack, e.g. insurance payout.	Coverage estimation (D, E), Incident claims reports (G), Incident claims statistics (G)
Attacker investment (A, B, H, I, J, K, M)	Up-front monetary requirement (attacker/attack cost) needed to finance the attack, e.g., equipment or software costs, outsourcing, development or size of a bribe. This investment may fail before the attack has started.	Expert opinion (A, K, M), (Retail) Price lists (H, K, M), Service documentation (H, M), Darknet markets (H, I, J, M), Attacker profile database (K, M)
Attacker penalty (A, B, H, J)	The consequences for the attacker given that the attack fails, e.g., fines or loss of equipment.	Expert opinion (A), Darknet markets (J)
Attacker profit (A, B, H, J, M)	The economic profit or gain the attacker will receive should the attack succeed, e.g., from ransom, fraud or scam money.	Expert opinion (A, M), Cryptocurrency market (H), Simulation (H), Darknet markets (J)
Attacker supplier profit (H, I, J, M)	The economic profit or gain for someone supplying the attackers with tools, services or information, e.g., zombie network rent or sale of malware. Does not necessarily depend on the success of the subsequent attacks, but could be based on commission.	Cryptocurrency market (H), Darknet markets (H, I, J)
Attacker opportunity cost (H, I, J)	A type of loss for the attacker due to poor choices, e.g., missed profit, wasted time or over-investment in an attack. Also includes frauds by attacker suppliers.	Darknet markets (I, J), Simulation (H)

The table is not meant to show an exhaustive list of data types. There are additional possible metrics, as well as more detailed ones, e.g., different types of loss and investments shown by Hoo [189], Brecht and Nowey [280] and Wang et al. [173]. Huang et al. [164] provide a literature survey that includes pricing mechanisms for cyber criminal services. The important thing is that the application of data should be based on needs and availability. Including data types in a model where there are no relevant sources will have little practical usage, and would at best describe a data gap. Paper L set out to map current practices of using security indicator data, and it is safe to say that currently cost related data are scarcely used compared to many of the other types of content.

Where there are data sources available, one should be aware of the strengths and limitations of these. For instance, when do data become too historical and unreliable? After one month, two years or ten year? As Hoo [189] points out; “past data are still relevant to new security incidents and that despite the fact that the road ahead may bend with human whim and technological advance, ...it does not appear to bend too sharply too often”. This implies that we should follow the recommendations of Almkaynizi et al. [193], making sure that we have transparency of the data and can make a judgment on its relevance.

Our work has not had a focus on data for *defender investments*, as this area already has a lot of attention in the research community. Hence, we have not worked with actual sources besides from *expert opinions*. It is worth mentioning that the size of the security budget of an organisation might be useful information for an attacker. These figures can be available in open annual reports, and give an indication on the limits of the risk modifiers in place. As a consequence, this could possibly attract or scare away rational attackers.

We have gained more experience with *defender reactive cost* and *loss*. This has mostly been in line with others in the field, such that they are often biased by region and sector. For instance, the *cyber loss events* we obtained from Advisen are mostly based on incidents that have taken place in the U.S., where costs following data breaches are relatively high. This dataset is also incrementally growing, and it is difficult to assess the age and relevance. During discussions with stakeholders from the insurance industry, it became clear that they are aware of these limitations, but it is still regarded as the best there is out there. We would argue that such data sources with content that is more than 2-3 years should not be considered instrumental for risk management decisions. Furthermore, there is a potential bias related to the source provider. Many security companies and police agencies publish annual reports on cyber loss, but as Anderson points out [1]; “most of the relevant publications come from organisations with an incentive to talk up the losses”. Woods and Böhme [8] similarly claim that security vendors tend to provide answers that are of self-interest using shaky methodologies, and this limits their credibility. This is in perfect accordance with what we observed related to external estimations for Ransomware-as-a-Service



in papers J and Q, further discussed under SQ5. We found *incident claims reports* and *statistics* to be advantageous in terms of precision, detail and relevance, but these have limited openness and availability. There have been initiatives to collect and share this kind of information, such as the *Cyber Incident Data Exchange and Repository* (CIDER) [281] by the Geneva Association. However, organisations are not readily willing to share sensitive information of this nature, and anonymised data have been causing issues related to duplicate entries. Furthermore, Nurse et al. [282] argue that insurers have invested in building up their own datasets, and sharing these would negatively impact their efforts by lowering the barriers for competitors to enter the market.

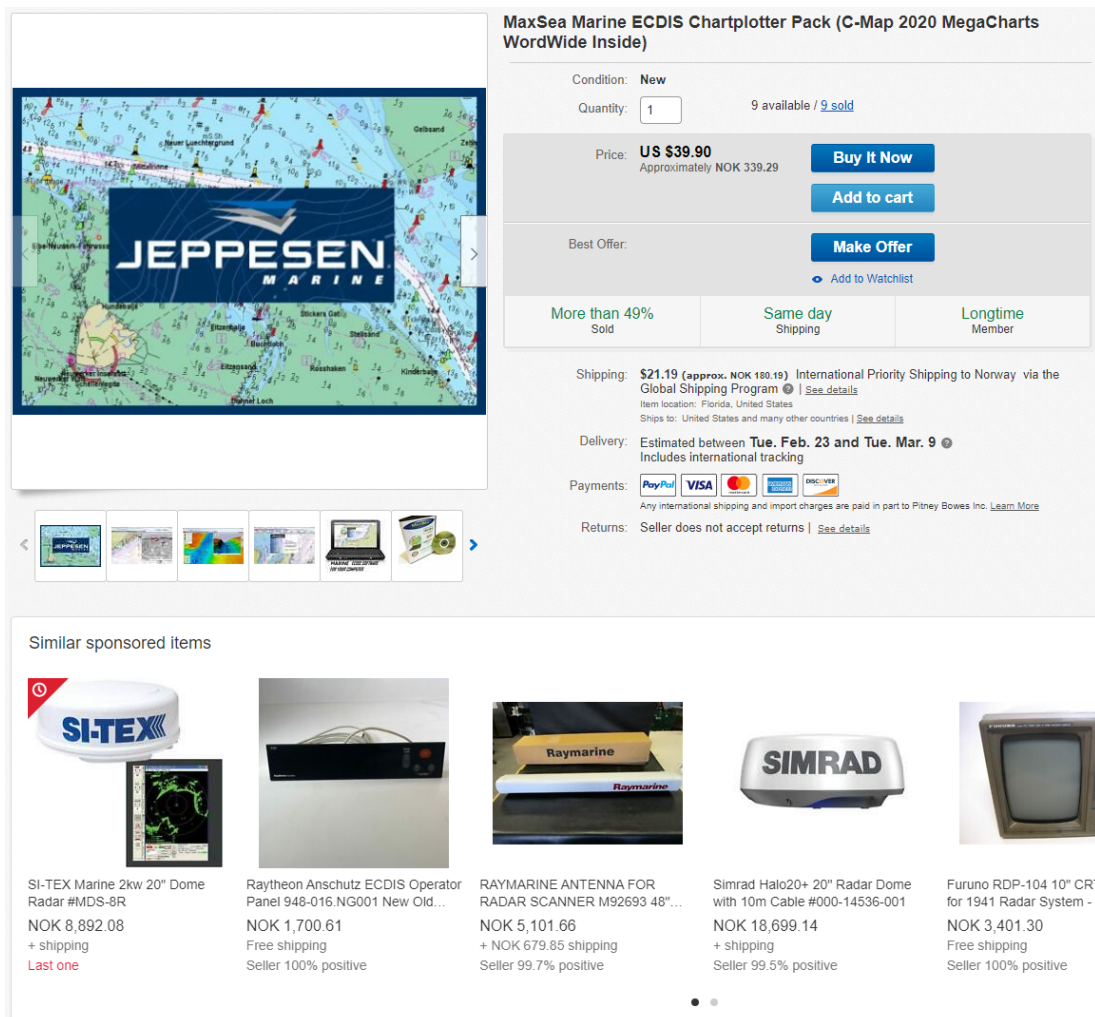
*Defender reimbursement* is a data type primary related to cyber insurance payout, but we also saw that the direct monetary compensations were not always the primary interests of the insurees.

Switching to the attacker's point of view, we find what we believe are the most novel contributions to both theory and practice. *Attacker investment* is basically a hidden data type, and even though many operate with such a metric within different threat modelling approaches, there are not that many making use of actual data sources besides from expert opinions. For instance, Wang et al.'s publication from 2018 [173] use the popular *Lincoln laboratory attacks classification* [283] to quantify attacker investments, but this dates back to 2006 and there has been a significant development in the cyber crime economy since then. Wortman and Chandy [284] recently claimed that possibly the most difficult value to verify when establishing security risks is the costs of performing an attack. Finding up-to-date and reliable data is therefore a significant challenge, whereas we have used a combination of sources. First of all, experts, and specifically system owners that know the weaknesses and vulnerabilities of their own product, were involved in cost estimations (see paper M). Secondly, we saw that many of the costs can be broken down to non-hidden items. For instance, an attacker would typically during the reconnaissance phase setup a copy of the target system in order to develop and test the attack technique. The required hardware and documentation could of course be stolen, but in many cases it seemed simpler just purchasing these items legally. In the cost estimation example in paper K, we used *price lists* for estimating a setup for the *Electronic Chart Display and Information System* (ECDIS) target. This included various buying options of hardware, documentation and software. Figure 8.2 shows an example of an ECDIS software available for purchase on eBay.

Similarly in paper H, the start-up fee and provision of using a particular mining software were found in public *service documentation*. Though the particular software and service we were using was abandoned and shut down during the course of the study<sup>b</sup>, it still gives an indication of market price. Furthermore, we analysed the code complexity of similar software, and made

---

<sup>b</sup>Possibly due to an exit scam.



**MaxSea Marine ECDIS Chartplotter Pack (C-Map 2020 MegaCharts WordWide Inside)**

Condition: **New**  
 Quantity:  9 available / 9 sold

Price: **US \$39.90**  
 Approximately NOK 339.29

[Buy It Now](#)  
[Add to cart](#)

Best Offer: [Make Offer](#)  
[Add to Watchlist](#)

More than 49% Sold | Same day Shipping | Longtime Member

Shipping: **\$21.19 (approx. NOK 180.19)** International Priority Shipping to Norway via the Global Shipping Program | [See details](#)  
 Item location: Florida, United States  
 Ships to: United States and many other countries | [See details](#)






Delivery: Estimated between **Tue. Feb. 23** and **Tue. Mar. 9**   
 Includes international tracking

Payments:

Any international shipping and import charges are paid in part to Pitney Bowes Inc. [Learn More](#)

Returns: Seller does not accept returns | [See details](#)

Similar sponsored items

				
SI-TEX Marine 2kw 20" Dome Radar #MDS-8R NOK 8,892.08 + shipping <b>Last one</b>	Raytheon Anschutz ECDIS Operator Panel 948-016.NG001 New Old... NOK 1,700.61 Free shipping Seller 100% positive	RAYMARINE ANTENNA FOR RADAR SCANNER M92693 48"..." NOK 5,101.66 + NOK 679.85 shipping Seller 99.7% positive	Simrad Halo20+ 20" Radar Dome with 10m Cable #000-14536-001 NOK 18,699.14 + shipping Seller 99.5% positive	Furuno RDP-104 10" CRT for 1941 Radar System - T NOK 3,401.30 Free shipping Seller 100% positive

**Figure 8.2:** Second-hand price for ECDIS software obtained from eBay.com

estimations on what it would take to develop such software in-house.

Not all attackers do or have the abilities to develop their own attacks, and in several studies, we have used *darknet markets* (or cryptomarkets) and associated forums to gain an indication of what it would cost to purchase ready-made/tailored attack software or hire someone to perform the attack entirely. However, we discovered severe questions about the legitimacy of these figures, which we discuss more under **SQ5**.

There are a number of *attacker profile databases* that give some kind of estimate of what kind of resources different types of attackers have at their disposal. This information can be used in two ways; either by estimating who the potential attackers are given the required attacker investment, or if the attacker is known, indicate what likely attack vectors could be. Papers **A** and **K** mainly dealt with the former approach, but also argued that costs alone should not be the

exclusion criteria.

*Attacker penalty* is an external risk modifier, which is highly dependent on the likelihood of getting caught. These values will vary depending on where the target system resides and under what jurisdiction. Since cyber attacks are global phenomena, it is challenging to generally quantify such values. With papers A and J, we have taken simple approaches, using scale values based on expert opinion and qualitative data from darknet forums, respectively. This can certainly be improved, and there are others that have done a more thorough job here to obtain quantifiable data. Based on the work of Fultz and Grossklags [163], Konradt et al. [285] have examined public information about previous convictions and sentence guidelines. Their estimations were specifically related to *phishing* attacks and U.S. law, and it is difficult to say how transferable these numbers are to other attacks and regions. Since court convictions tend to be based on case law, penalty is actually a type of historical data that seem relatively stable within the same jurisdiction.

*Attacker profit* is the main driver for rational attackers, and we have tried to identify data sources that can reveal such information. Much of the *loot* from cyber crime are in some form of cryptocurrency, and the general *cryptocurrency market* gives an indication on the profitability of attacks producing that. With paper H, we specifically examined the coin mining threat *cryptojacking*, and hypothesized that its rise and fall was significantly related to the Monero market value. We concluded in 2019 that “if the cryptocurrency markets should resurge, it is likely that cryptojacking will follow suit”. At that point in time, there had been a severe drop for almost all cryptocurrencies, and cryptojacking had gone from one of the most feared threats to almost becoming extinct. Presently, the cryptocurrency market is booming again (especially for Bitcoin), and the McAfee Labs Threats Report for November 2020 showed that coin miner malware increased over 25% over the previous quarter and becoming once again one of the dominating malware threats. Through *simulation*, we were able to determine what potential profit would be based attacker investments, cryptocurrency market value and parameters such mining malware type (native or script), infected device type (CPU, threads, memory) and number of infected devices. Parallel to us, Saad et al. [286] and Papadopoulos et al. [287] also analysed the potential profit from in-browser cryptojacking on different devices, coming to the same conclusion that it is not a good source of income.

In paper J, we examined typical profit from ransomware, particularly *ransomware-as-a-service* (RaaS), by using data from darknet markets. However, there are several weaknesses with such sources as we discuss more under **SQ5**.

As was shown in paper J and its related work, attackers seldom operate alone, but are part of more complex cybercrime ecosystems that involve different roles and stakeholders. Figure 8.3 illustrates the value chains (blue arrows) originating from the victim and then going to the

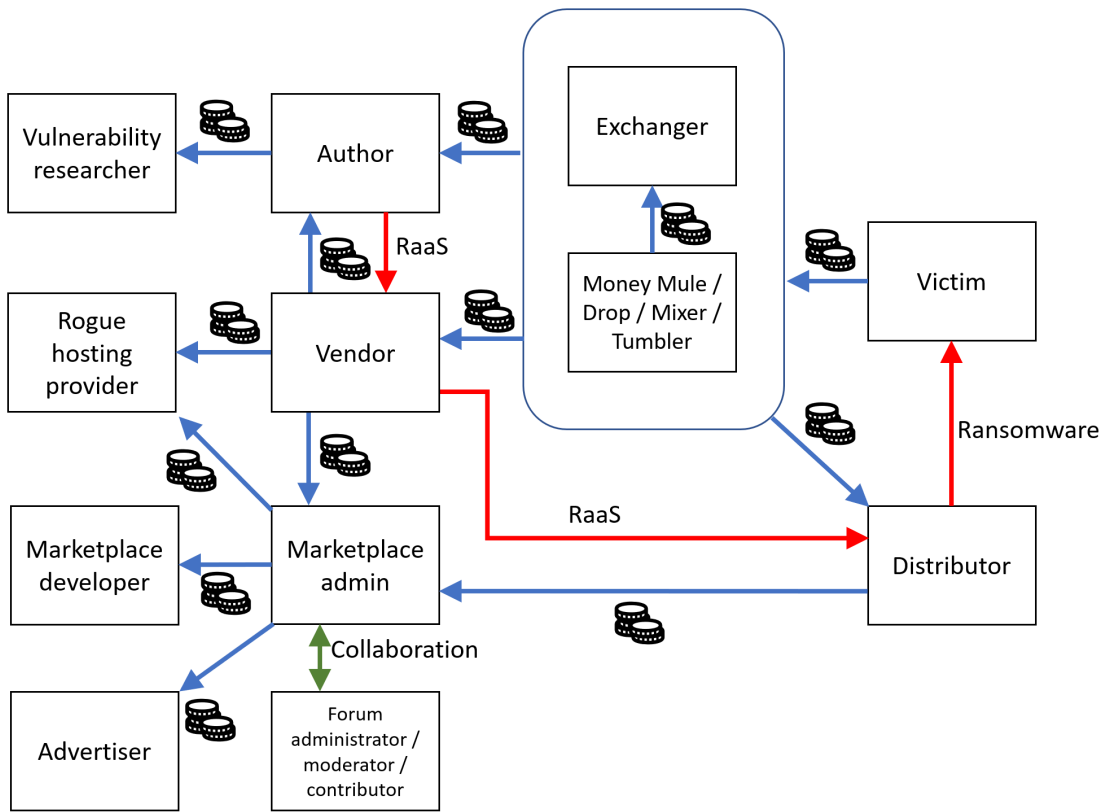
various middlemen in a RaaS ecosystem. We were able to find data about *attacker supplier profit* by examining darknet markets in papers I and J, and combining that with cryptocurrency market price indexes. Where there is a high supply and volume of sales, e.g. of a particular type of malware, we can assume that there is a high demand. We have argued that this demand is an indication of active threats. Providing malware is not necessarily a crime by itself, as the suppliers argue that it should only be used for educational purposes. Hence, prices are openly published and can be collected for various software and services. In the literature survey by Huang et al. [164], there are many examples of existing, evolving and emerging cyber crime services, their pricing models (licences, subscription, commission, pay-per-install, etc.) and estimated prices. In a study from 2021, Lee and Choi [288] draw on *routine activity theory* (RAT) and *cyber-routine activity theory* (C-RAT) to explore possible links among Bitcoin, ransomware, and terrorist activities. The findings indicated unidirectional ties between the prevalence of ransomware and Bitcoin as well as ties between the prevalence of ransomware and terrorist activities. However, as the study did not use direct measurements of ransomware frequencies, but trending search queries as a proxy measurement (Google Trends data), the relationships could not be properly measured. As future work, it would be interesting to test their hypotheses with more accurate data.

*Attacker opportunity costs* can be seen as a loss component of attacker profit, and we suggest that it can be used to measure the irrationality of an attacker. For instance, given that an attacker has successfully infected a target system, if the concrete exploit has a high opportunity cost, one can suspect that there are other motives behind the attack. Alternatively, this could be an indication of mere incompetence. Through simulation we could estimate what wrong choices related to cryptojacking could result in from a monetary point of view. From darknet markets, we could observe concrete data that can be tied to opportunity cost. For instance, there were lots of malware source code for sale in many of the marketplaces, which can be found for free elsewhere on public web sites. The same goes for software and e-books typically found in various hack-packs.

### **8.1.3 SQ3: How can economic incentives enrich existing threat modelling techniques?**

As already mentioned in Section 2.1, there are many ways of representing threat models, there is no single ideal, and combinations can be used to illuminate different problems. Table 8.2 gives an overview of the assemblage from the primary papers and which economic incentives that have been used in relation to them.

As can be seen from this table, different types of incentives were applied to *attack-defence trees* in paper A. These were costs seen both from both the attacker's and defender's point of



**Figure 8.3:** A value chain for the RaaS economy as presented in paper J

view, while the previous literature were mainly focused on attacker costs in attack trees. It is the scenarios with a low *minimal costs of a successful attack* that we want to identify and consequently treat. Following this publications, there have been other researchers that used the suggested cost attributes and built on our warehouse case study, such as [115, 289–294]. In 2020, Buldas et al. [295] presented a quantitative attribute approach for attack trees that deals with incomplete information. This could be applied when there are some historical data and some domain knowledge available to the model. Even more recently, ter Beek et al. [296] have developed the *RisQFLan* framework for quantitative security risk modelling and analysis based on attack-defence diagrams. Here, the cost of an attack (both successful and failed) are calculated and used as a constraint.

*Bow-tie diagrams* are meant to provide a more high-level view of risk scenarios, and are centred around causes and consequences for unwanted events. Following the threat modeling manifesto [74] recommendation of reaching for practical and relevant solutions, bow-ties are already used in practice in safety-critical sectors, hence it made sense to try to include cyber risk considerations here as well. Parallel to us, Abdo et al. [191] have also taken the approach of combining safety and cyber security in bow-ties for industrial risk analysis. We initially only

**Table 8.2:** Economic incentives and threat modelling

Paper	Threat modelling technique(s)	Economic incentive(s)
A	Attack-defence tree	Defender investment, defender loss, attacker investment, attacker penalty, attacker profit
B	Risk and treatment modelling	Defender investment, defender reactive cost, defender loss, defender reimbursement, attacker investments, attacker profit, attacker penalty
C	Bow-tie diagram	Defender loss
E	Part of a generic risk model	Defender investment, defender loss, defender reimbursement
H	Attacker mining model	Attacker investments, attacker profit, attacker supplier profit, attacker opportunity costs
I	Market model for cyber attacks	Attacker investments, attacker supplier profit, attacker opportunity costs
J	Market model for RaaS	Attacker investments, attacker profit, attacker supplier profit, attacker penalty, attacker opportunity costs
K	Cyber kill chain and resource cost modelling	Attacker investments
M	Bow-tie diagram, OWASP risk rating, cyber kill chain and resource cost modelling	Attacker investments, attacker profit, attacker supplier profit

used defender loss as a quantifiable indicator on the consequence side of the diagrams as shown in paper C. This was an important component for estimating risk values for the unwanted events, and used in practice in the development of a solution for securing maritime communication (further addressed in secondary papers N, T, W, X, Y and Z). In order to have a better foundation for estimating the left side of the bow-tie diagrams (likelihood of causes), we wanted to make more use of incentives for attackers, specifically attacker investments. Hence, in paper K we used the existing *cyber kill chain* approach to break down attacks in various investment phases and developed the *resource cost modelling* notation based on attack trees. This approach does not consider profit-related incentives for the attackers, which makes the models less complex and easier to populate with values. Instead, there is a simple assumption that *cheap attack paths allow for many types of potential attackers*, and we want to break their assumed budget before they get to any profit stage. In addition to our own evaluation of this approach, Walde and Hanus [297] have applied it for modelling necessary investments for performing maritime *AIS spoofing* attacks, and verified this through real purchase of the necessary equipment and performing an attack demonstration.

In paper M, we tied bow-tie diagrams and underlying models together to show how we

can use different techniques and data for managing maritime cyber risk. In relation to this work, we also experimented with defender loss and defender investment as parameters related to preventive and reactive security barriers found in bow-ties, wanting to identify cost-benefit trade-offs. However, this work was not part of the evaluation, and thus we regarded it as out of scope for the paper. A lesson learned was that it might have been beneficial to apply underlying models such as Bayesian networks to combine different sources of knowledge, especially when there are small and incomplete datasets [99]. However, Falco et al. [269] point to pitfalls related to cross-disciplinary nature of cyber risk. As an example, a statistician might apply Bayesian modelling to predict future cyber events, even though it is not entirely clear what bearing the input data have on future ones.

Unlike the threat modelling techniques above, *risk and treatment modelling* described in paper B does not use a graphical representation, but describes which attacker and defender incentives considered in relation to each other, and where cyber insurance fits in. Similarly, the *generic risk model* in paper E consists of equations that support defender decision-making based on investments, loss and reimbursement. The data sources are real, but we consider them more as illustrative examples of what can be done with the best data available. We also suggest to extend the generic risk model by using attacker costs as possible baseline data type. The *attacker mining model*, the *market model for RaaS* and the wider *market model for cyber attacks* as presented in papers H, J and I contain attacker-specific incentives and can be used as concrete input for such baseline data. This implies that these attacks techniques must be relevant for the threat model. To cover further threats, there is a need to develop additional cyber attack market models, like Konradt et al. [285] have done for phishing, Vasek and Moore [298] have done for *Bitcoin Ponzi schemes*, Tais [299] has done for *DDoS attacks* (“botconomics”) and Stone-Gross et al. [300] have done for *spam*. Furthermore, the models must have sound and up-to-date input data, preferably adjusted to the local context. Jamil et al. [301] have recently studied practices for threat modelling, and interviews show that key challenges are related to models not being updated and that threat modelling knowledge cannot be transferred from one domain to another.

#### **8.1.4 SQ4: How can the cost of risk treatment options be balanced with cyber insurance?**

It is obvious that storyless systems tend to have an unclear threat picture, and it is hard to determine the amount and type of preventive controls. As pointed out by Grobman and Cerra [9]; “the defense-in-depth approach is only as effective as its architects are in anticipating new threats and identifying them when they emerge”. In these situations of uncertainty, it might be better to acknowledge incident are bound to occur, and it becomes a question of minimizing the potential loss. Cyber insurance has emerged as a beneficial risk treatment option, and with

paper **B** we were early to show how a cost-benefit analysis can be used for such decision making. However, applying such a method is not so straight forward due to the uncertainty factors and expectations we empirically investigated among actual and potential insurees in papers **D** and **G**. As a contribution to practice, we developed the generic risk model already mentioned under **SQ3**, and showed in paper **E** how to populate this with available data sources and create a cyber insurance profile that can be used for risk transfer/sharing decision making. A limitation to this approach is the lack of open results from qualitative case study involving real users. Though we collaborated with both insurers and insuree agents, this step stranded due to the following reasons:

- Methods and techniques to support underwriting and classify customers were under development and not ready for any comparative analysis, and
- additional data that could be used to populate the generic risk model were considered business sensitive by the insurance companies, hence could not be shared nor evaluated.

There is a general lack of non-fictitious evaluation in most papers related to cyber insurance models, and the reasons above could be a common explanation to that. Xiang et al. [302] criticize many existing studies, claiming that they have limited practicality and remain conceptual as long as they neglect the highly uncertain nature of losses incurred by cyber incidents. We also acknowledge that we have created a simplistic model for selecting and comparing insurance policies, and this could be extended with more advanced techniques. Parallel to our work, Bodin et al. [303] have developed a model for selecting the optimal set of cyber security insurance policies (an insurance ladder) given a finite number of policies being offered by one or more insurance companies. Another example is from Xiang et al. [302], who suggest including a Bonus-Malus system for cyber insurance, which is a cost-reducing mechanism frequently used in vehicle insurance products. This will benefit organisations that are able to have a good track record and reduce moral hazard. Similarly, Wang [146] proposes a new type of innovative cyber insurance covering with a stronger focus on partnership between the insurer and insuree, and where the premiums are adjusted according to security benchmarks. As shown by Mazzoccoli and Naldi [304], moral hazard is indeed an issue with organisations that have either low or high vulnerability values. In these cases, the optimal strategy may be not to invest in security at all, but to rely on the protection provided by insurance alone.

The challenges and research topics we pointed to in our papers are still very much relevant today, though the availability and actual uptake of insurance products have increased. However, Norway is far from being an early adopter, and a study by Bahşi et al. [305] from 2019 showed that the uptake has been lower compared to the other Nordic countries. They were also able to observe a general scepticism about cyber insurance in the IT departments based on the perception that such products expressed a lack of confidence in them. Such psychological factors are seldom



considered in a risk management process, and could create tension within an organisation. Nurse et al. [282] have in 2020 conducted qualitative study with UK cyber insurance professionals in a focus group, and conclude that “cyber insurance is still a field in its infancy” and a number of open questions remain. They also found evidence of reluctance to data sharing due to potential loss of competitive advantages. This same year, Wrede et al. [306] have interviewed experts from German insurance industry, and found that the demand is still hindered by;

- the lack of transparency in and the complexity of insurance terms and conditions in the cyber insurance policies, and
- companies are insecure towards the cyber coverage.

Furthermore, we can see barriers related to quantifying risk and lack of solid data reappearing in another study from 2020 by Zeller and Scaerer [190].

### **8.1.5 SQ5: What can we learn about attack trends from studying phenomena within the cybercrime economy?**

Predictions about the *normal* economy is already said to be notoriously unreliable due to the complexity of the many economic influences, such as the human behaviour [182]. Likewise, it is difficult to get a clear insight of how the cybercrime economy works today and even more so in the future. What we do know is that it is (partly) hidden, growing world-wide, competitive, constantly evolving, increasing in complexity and of course unregulated. This research question was framed in order to investigate whether the cybercrime economy can be used as source for cyber threats despite its erratic nature. Such data typically stem from darknet marketplaces or forums, and we showed in paper L that such unconventional data sources are being increasingly used in the security research literature. This kind of data crawling/extraction is known as *scraping*, and is popularly used by commercial threat intelligence services as well. To overcome the problem with overwhelming data [9], AI-based techniques are commonly used to make sense of it. For instance, Marin et al. [307] and Deb et al. [308] use AI-based tools to look for vulnerability mentions among thousands of darknet forum posts and correlate them with cyber incidents to predict cyber-attacks.

As already mentioned in **SQ2**, we specifically looked at the market availability and supply chain of RaaS in paper J, which extends paper Q. High availability of an inexhaustible product would intuitively indicate a high demand, and subsequently a high risk of ransomware attacks. This line of argumentation is typically found in threat reports from security vendors as we have mentioned under **SQ2**.

We investigated RaaS more closely and were able to debunk this myth based on the actual number of successful sales of such items, as well as feedback from unhappy customers claiming that the products were scams. Though RaaS represents a small economy on the most popular

open darknet markets, which are mainly dominated by drugs, the real RaaS deal is more likely to be found in closed forums that are difficult to scrape. Examples of successful RaaS products sold through such channels are *Cerber*, *Conti*, *CTB-Locker*, *FONIX* and *GandCrab* [309]. In 2020, the threat intelligence company Intel 471 tracked over 25 different RaaS providers over a year [310]. They confirm that most of the RaaS are sold through private Russian-language forum groups on the surface web (e.g., *XSS* and *Exploit.in*), and it is difficult to verify sales volumes beyond what is claimed by the providers. The most utilized variants (*DoppelPaymer*, *Egregor*, *Netwalker*, *REvil* and *Ryuk*) have allegedly pulled in hundreds of millions in ransoms.

We question the use of marketplace inventories as a threat indicator, especially if it is done without a qualitative inspection of the items. For instance, we found evidence that many of the items are deliberately written or tagged in such ways that they would fool machines. This hinders the applicability of automated scraping and analysis of such data. A number of practical limitations to scraping dark net markets have also been identified by Lawrence et al. [311] Ball et al. [312] and Hayes et al. [313], such as significant downtime, user account timeouts, captchas, banning and non-standardised implementations.

With paper I, we looked at the more wider inventory of cyber attack software and services found in darknet markets and forums. With a greater awareness on the numerous fraudulent items being offered, we compared availability and number of actual sales (popularity). We also identified what the most profitable items were for the vendors, and looked at which items were new and trending, e.g., related to phone hacking, information theft and Bitcoin stealing. We conclude that actual sales is a more reliable indicator than availability, however, fewer marketplaces provide that kind of information. A similar approach has been taken by Soska and Christin [314] regarding darknet drug sales, however their sales volume estimation is based on user feedback, which we found to be questionable as well. Though some vendors have a “Trusted” status, we saw that this is a stamp that can be bought, and not really earned. Looking for new items appearing in marketplace inventories is also an approach taken by e.g., Nunes et al. [315], Robertson et al. [316], Lawrence et al. [311] and Dong et al. [317], and can provide better information about instant threat than sales. A high price could additionally be an indicator of legitimacy, but this could be speculative from the vendor side as well.

Despite that scraping from darknet marketplaces and forums can be deceitful, there are also many opportunities once you know what to look for and the limitations of that data. This opens up to further research avenues. For instance, the security company Positive Technologies use advertisement analysis on dark net forums to measure demand [318]. They found that 90% of the users are searching for a hacker who can provide them with a particular resource or who can download a user database, and that there has been a growing demand followed by the increased internet usage of organisations and individuals since the start of COVID-19. Their research

show that the theft of a custom client database can cost up to \$20 000.

Thanks to a number of external shocks during our study period, we were able to observe concrete examples of vendor resilience after law-enforcement take-downs. This includes transfer of vendor reputation despite that accounts are anonymously tied to pseudonyms. We believe that this indicates collaboration between marketplaces operators and vendors, and that access to attack software and services is not significantly impaired by such events. We also saw an external shock related to cryptojacking in our study leading to paper H, as the group behind the Coinhive JavaScript suddenly shut down their support in 2019. The crash in the cryptocurrency market and the required mining effort simply made the whole business unprofitable. This script accounted for 70-75% of the market in the year before, and browser-based mining (both illegal and legit) has yet to recover. Tekiner et al. [319] have in their recent study showed that the trend is more towards host-based attacks on hardware with more processing power than personal computers, e.g. enterprise cloud infrastructure servers. Such targets are fewer, but represents more profit in lesser time.

Complementary to our work, Spagnoletti et al. [320] have studied the forces underlying dark net markets, showing their evolution and evidence of resilience. Furthermore, Abeer et al. [321] have recently analysed 24 separate episodes of unexpected marketplace closures. They have showed that despite marketplaces might appear fragile, user migration is done swiftly and sales volumes recover quickly, and the overall economy has a systemic resilience. Collier et al. [322] have taken a qualitative sociological approach with interviews of darknet infrastructure providers and analysis of scraped data from forums and chats. They suggest that the boring nature of operating this kind of infrastructure leads to burnout and withdrawal of services. Hence, it should be possible to smash the cybercrime economy by taking the fun and profit out of it.

## 8.2 Ethical issues

As shown by Burstein [323], Christen et al. [324] and Macnish and van der Ham [325], there is a number of potential ethical issues when conducting research related to cyber security. For instance, obtaining and sharing data might expose individuals or put them in harm's way, experiments with malware may cause third party damages, and researchers may find themselves in a position where they have the possibility to disrupt or mitigate attacks.

Where applicable, the details of such considerations are described in the individual papers, whereas this section discusses the more general considerations of the overall work with the PhD thesis. This is structured according to the set of ethical principles for guiding information and communication technology research defined in the *Menlo Report* [326] published by the U.S. Department of Homeland Security.

### **8.2.1 Stakeholder perspectives and considerations**

First, this principle is about how information in the research could identify individuals, and balancing the risks and benefits of the multiple stakeholders involved. Regarding participants in interviews, experiments and evaluations, all data have been anonymised and cannot be traced back to individuals. All research authors have been involved in the writing process and given their consent to exposing their names and results. Activities related to malicious actors within the darknet or surface web can be traced to pseudonyms, which in theory could be used to identify individuals. Online traces such as archive data can be accessed by anyone, so this information should technically be considered as published open content. Pseudonyms tied to cybercrime item offerings, sales and feedback were not recorded, and can thus not be used for prosecution.

Second, there could be stakeholders that rely on information and systems that are involved in the research and could be harmed by that. This was not a major concern since the systems related to our research were new designs and not in operational use yet. With all the hardware used to benchmark cryptomining, we used our own devices dedicated for that purpose.

### **8.2.2 Respect for persons**

This principle includes consideration of the computer systems and data that directly interface, integrate with, or otherwise impact persons who are typically not research subjects themselves. All participants in interviews, experiments and evaluations were informed about the research and gave their consent. They were free to withdraw at any time without losing any benefits. Though we have pointed to differences between experts and students, this has not been to discredit any of them. Similarly, we believe we have not put informants from interviews regarding cyber insurance in negative light when they have talked about obstacles and challenges.

### **8.2.3 Beneficence**

This principle reflects the concept of appropriately balancing probable harm and likelihood of enhanced welfare resulting from the research. The overall objective with cyber security research is to provide benefits to non-criminal stakeholders and the society as a whole. However, one has to be aware of potential misuse of results, and that publishing could be a double-edged sword. Our results involve threat models with ways of exploiting systems and making profit, and this could be used for planning attacks by threat agents. Furthermore, Zheng et al. [327] argue that data might reveal to adversaries what is known about their activities, inadvertently assisting them in their criminal activities.

To avoid this concern, we have not made public any detailed threat assessment related to aviation nor maritime. Instead, we have focused on principles and practical usage of threat modelling techniques combined with economic incentives to aid cyber risk management. This comes at some cost for transparency and openness, but has been necessary to protect the security of the systems in development and the business models of the involved stakeholders. At the same time, we are happy to see that specifications necessary for secure VDES interoperability has become part of standards such as the technical guidelines from the *International Association of Marine Aids to Navigation and Lighthouse Authorities* (IALA) [328].

Regarding the informants and data providers related to cyber insurance, they have been given access to reports and publications stemming from this work, and have had the opportunity to make use of this as they see fit. As we have seen, there is a general agreement that more research is needed in this field to make the products more viable, and we hope that our results can be an enabler for further work.

When it comes to the experiments involving students, we designed these so that they would benefit from the experience as threat modelling is commonly used for exam assignments. It was entirely voluntarily to submit the models they created to us without losing this benefit.

#### **8.2.4 Justice: Fairness and equity**

This principle concerns a fair selection of research subjects, and an equitable distribution of the burdens and benefits of research according to individual need, effort, societal contribution and merit. Though research on human subjects has not been the primary objective of this PhD study, several types of stakeholders have been involved as already explained. This selection has mainly come naturally from their existing involvement in the ongoing research projects presented in Chapter 3, and can be seen as a purposeful inclusion based on their ability to understand the problem setting and possible remedies. Other stakeholders have voluntarily participated based on self-interest in the research topics or the benefits mentioned above. Though further inclusion could have improved the research results, we do not believe we have made any exclusions unrelated to the purpose of the research.

#### **8.2.5 Respect for law and public interest**

This is a separate principle with two applications; 1. *compliance* – where the researchers should engage in due diligence to identify and respect applicable laws, regulations, contracts, and other private agreements, and 2. *transparency and accountability* - which is about clearly communicating the purposes of research and that research methodology, ethical evaluations,

data collected, and results generated should be documented and made available responsibly in accordance with balancing risks and benefits.

For all studies related to this PhD thesis, we made a deliberate choice of not supporting or sponsoring any illegal activities to gain access to empirical data or software. This can be seen as a limitation to the research results, as we for instance would have been able to better verify the authenticity of malicious items and services by engaging in active trade. We consider buying malware or paying for access to closed darknet forums to be unacceptable for our kind of research, though we are aware that this is done to some extent by some law enforcement agencies, security companies and state intelligence bureaus. Furthermore, we have not tried to deceive, intimidate, provoke or confuse people operating in these spaces, but rather collect already available digital traces.

Any organisation that has shared datasets with us has done so willingly and we have established agreements on the usage of the data. We have made references to the origin of the data, so that others have the possibility to make similar agreements with the source. Datasets that we have created ourselves, such as from paper **F** and **L**, have been published as open research data identified by a DOI.

We have disseminated our methods and results through a number of scientific publications and research reports, and chosen to make this open access when possible. As mentioned above, there are cases where the details of threat analysis on actual systems have been kept restricted. This PhD thesis will be made publicly available by NTNU, and we have obtained written permission to openly include all the primary papers. Finally, support tools for bow-tie modelling and risk estimation have been released under open source licences.

### **8.3 Future opportunities and recommendations**

Several areas for future research on cyber threats and economic incentives could add to the findings in this study. Some of these were simply too comprehensive to be included in this scope from the beginning, others were identified during the research as a possible follow-up or based on coincidental findings. The following bullets outline our main recommendations:

- As mentioned under **SQ3**, there is a need to develop further market models for different types of attacks than the ones we and existing literature have covered. As new attack techniques and ways of making profit emerges, new market models should be developed as well. Breaking the economy behind the front-line risk-takers is more beneficial and economically sustainable than defending all possible attack points.
- The accuracy in both defender and attacker investment estimations could be improved if we had a database of baseline costs (not just loss). This would easily fall under the

category *unreliable and misleading historical data* if not kept up to date. It is therefore of interest to investigate whether the information could be collected and maintained in a purposefully way. Perhaps a collaboration between academic researcher, the insurance industry and security vendors could result in a less-biased way of presenting data for security investments.

- If there are domain specific considerations, these should be identified and added to the baseline data as well. For instance, systems related to services/hospitality seem to be more exposed or are easier to breach than financial systems (see paper E). The number of possible victims (size of the domain) and e.g. willingness to pay ransom could be useful to include in an attacker reward model. As reported by the security company CheckPoint in the beginning of 2021 [329], healthcare organisations seem to be lucrative targets as the COVID-19 cases have risen again.
- More credibility could be given to the resource cost model and similar approaches if we modelled attacker investments of known incidents, and compared the results with models based on theoretical estimations. This would require access to data about what the attackers actually spent in the real cases. Interviewing perpetrators is something that has been done related to *Nigeria scams* (a.k.a. *419 fraud*, *advance fee scam* or *Spanish prisoner scam*) [330, 331], which gave insight into success rate, typical reward and resources invested into each victim. This could be done with other cyber criminals and their techniques as well, but does of course impose ethical considerations.
- Threat modelling tools can improve collaboration between different stakeholders. With both the tool for bow-tie modelling and research cost modelling we added support for model sharing and collaborative editing. We are already working on better support for calculating costs as part of the user interface, and functionality for automatically collecting empirical data from the models. We have also played around with features that allow users to anonymously make votes for threat rankings to avoid the aforementioned groupthink phenomenon, and a simple experiment with a class of students showed that this could be done in a fun and useful way.
- Finally, more comprehensive studies may go beyond the stated limitations (see page 20) of this research, also encompassing non-rational actors. For instance, *hybrid threats* can have a complex mix of conventional and unconventional offensive methods, such where diplomacy, military, economy and technology are used by state or non-state actors to undermine fundamental democratic values and liberties [332].

## CONCLUSION

*We shall not cease from exploration and the end of  
all of our exploring, will be to arrive where we  
started and know the place for the first time*

---

T.S. Eliot, *Little Gidding*

Cyber risk management is about identifying, assessing and reducing risk to an acceptable level. There is a cost to reaching this level, and this study has sought methods for quantifying and balancing risks in order to make informed decisions about security investments. For more than two decades, threat modelling, which involves taking the mindset of a potential attacker, has been a technique to support this process and can be found as a vital component in many prominent cyber security frameworks. However, the literature also shows that threat modelling is costly and difficult to perform in practice, and tends to fail due to uncertainty of input data.

The main research question of this thesis has been *How can modelling threats and economic incentives improve cyber risk management?* To address this, the overall approach of the study leans towards practice research, where interventions and designs contribute to local practices as well as generalized knowledge. We have shown ways of using threat models based on the context, available expertise and input data at hand. In situations with limited historical data to rely on, the subjective opinions of security experts and domain specialists can be supported by economic incentives that are used to help identify threats, the likelihood of their occurrence and ways of making them more costly.



Without advocating for particular threat modelling techniques, we have shown that graph models make it intuitively easier to break down high-level costs into more tangible costs, hence improving accuracy on estimations. Instead of modelling specific attack paths, we can spend less time and obtain a higher degree of reusability by focusing on the capabilities that are needed for the different attack stages. Similarly, starting from a generic risk model and tailoring it to fit individual organisations based on their characteristics is a relatively simple way of evaluating risk treatment options such as cyber insurance.

Just as there is no one-solution-fits-all for threat modelling, we cannot use data types and sources for economic incentives uncritically. It is therefore important to study and know their strengths and weaknesses, and if it makes sense to include or combine them in a model. Through focused studies, we have used the data types; investment, reactive cost, loss and reimbursement as seen from the defender's viewpoint, and; investment, penalty, profit and opportunity cost as seen from the attacker's viewpoint. Our sources of actual values span from expert/specialist opinions, single incident claims, datasets for incidents tied to sector and industry size, retail price lists, darknet markets and profit simulations on different types of infected systems.

We have seen that unconventional data sources, such as social media, darknet market listings and cyber crime ads, are becoming increasingly popular for threat intelligence. However, we question the legitimacy of estimating attacker investments based on such sources due to what we consider to be a significant portion of fraudulent items. On the other hand, these sources can be useful indications of trending attack techniques, discovery of vulnerabilities and victims. Furthermore, observable externalities such as changes in the cryptocurrency market or mining algorithms can have a significant impact on attacker's choice of target and method of exploitation.

We have argued that with a more extensive set of cyber attack market models, we can improve our threat prediction and find ways of reducing attacker incentives. From a macroeconomic standpoint, joint efforts targeting the complex cybercrime value chains are considered to be more rational than investing in protection everywhere. Unfortunately, international cooperation on policing is complicated and time consuming [333], therefore it is still very much up to individual organisations to spend a sufficient amount of resources on security, both for their own sake and the society as a whole.

The solutions supporting risk management in relation to this PhD study share a number of common traits, and we have tried to define a set of generalised recommendations for modelling threat with economic incentives. If and how they should be applied should be determined on a case-by-case basis, taking the following the steps and questions into account:

- *First*, we have to determine what we know about the system at hand<sup>a</sup>. Is it a completely storyless system, or can we learn from similar ones or other domains? When assessing

---

<sup>a</sup>This is in accordance with the first question from the Threat Modeling Manifesto [74] (*What are you building?*)

threats, we should determine whether or not expertise is available or exists at all, and if subjective judgments alone will be satisfactory. With a team of experts and domain specialists, how should they collaboratively work together, e.g. to avoid groupthink or domineering? Which threat modelling representations are suitable for the system and stakeholders involved?

- *Second*, what kind of supportive data, such as economic incentives, are actually available? Are those data relevant, reliable and feasible to consume as a part of the threat model? With several sources of data, how should they be combined in a meaningful manner? Are there conflicting data, and which should take precedence?
- *Third*, reflect on how confident we are in the result<sup>b</sup>. Over-spending resources does not necessarily result in a more secure system, and security decisions tend to only get attention when they are wrong.

Keeping these steps in mind will hopefully help a wisely spending of security resources and avoid future Maginot lines.

---

<sup>b</sup>This is in accordance with the fourth question from the Threat Modeling Manifesto [74] (*Did you do a decent job of analysis?*)



## BIBLIOGRAPHY

- [1] R. Anderson, *Security engineering, 3rd edition*. John Wiley & Sons, 2020.
- [2] R. Soodalter, *The maginot mentality*, Jul. 2019. [Online]. Available: <https://www.historynet.com/the-maginot-mentality.htm> (visited on 01/03/2021).
- [3] Merriam-Webster.com, *Maginot line*. [Online]. Available: <https://www.merriam-webster.com/dictionary/Maginot%20Line> (visited on 02/03/2021).
- [4] Wikipedia, *Maginot line*. [Online]. Available: [https://en.wikipedia.org/wiki/Maginot\\_Line](https://en.wikipedia.org/wiki/Maginot_Line) (visited on 02/03/2021).
- [5] Böcker, *Bild 101i-382-0248-33a*, CC-BY-SA 3.0, May 1940. [Online]. Available: <https://www.bild.bundesarchiv.de/dba/de/search/?query=Bild+101I-382-0248-33A> (visited on 08/03/2021).
- [6] B. Filkins, ‘Quantifying risk: Closing the chasm between cybersecurity and cyber insurance,’ SANS Institute InfoSec Reading Room, Tech. Rep., Mar, Tech. Rep., 2016. [Online]. Available: <https://www.sans.org/reading-room/whitepapers/leadership/paper/36770>.
- [7] M.-E. Paté-Cornell, M. Kuypers, M. Smith and P. Keller, ‘Cyber risk management for critical infrastructure: A risk analysis model and three case studies,’ *Risk Analysis*, vol. 38, no. 2, pp. 226–241, 2018.
- [8] D. W. Woods and R. Böhme, ‘Systematization of knowledge: Quantifying cyber risk,’ [Online]. Available: [https://informationsecurity.uibk.ac.at/pdfs/WB2020\\_sok\\_cyberrisk\\_snp.pdf](https://informationsecurity.uibk.ac.at/pdfs/WB2020_sok_cyberrisk_snp.pdf) (visited on 14/01/2021).
- [9] S. Grobman and A. Cerra, *The Second Economy: The Race for Trust, Treasure and Time in the Cybersecurity War*. Apress, 2016.
- [10] C. Beek, S. Chandana, T. Dunton, S. Grobman, R. Gupta, T. Holden, T. Hux, K. McGrath, D. McKee, L. Munson, K. Narayan, J. Olowo, C. Pak, C. Palm, T. Polzer, S. R. Ryu, R. Samani, S. Sarukkai and C. Schmugar, ‘McAfee Labs Threats Report, November 2020,’

- Tech. Rep., Nov. 2020. [Online]. Available: <https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-nov-2020.pdf>.
- [11] S. Berinato, 'Active defense and 'hacking back': A primer,' *Harvard Business Review*, May 2018. [Online]. Available: <https://hbr.org/2018/05/active-defense-and-hacking-back-a-primer>.
- [12] Merriam-Webster, *Definition of storyless*, 2020. [Online]. Available: <https://www.merriam-webster.com/dictionary/storyless> (visited on 18/12/2020).
- [13] C. Costley and J. Fulton, *Methodologies for Practice Research: Approaches for Professional Doctorates*. SAGE Publications Limited, 2018.
- [14] A. H. Van de Ven, *Engaged scholarship: A guide for organizational and social research*. Oxford University Press on Demand, 2007.
- [15] L. Mathiassen, 'Designing engaged scholarship: From real-world problems to research publications,' *Engaged Management Review*, vol. 1, no. 1, p. 2, 2017.
- [16] P. Checkland, 'From framework through experience to learning: The essential nature of action research,' *Information System Research*, 1991.
- [17] A. Bagnato, B. Kordy, P. H. Meland and P. Schweitzer, 'Attribute decoration of attack-defense trees,' *International Journal of Secure Software Engineering (IJSSE)*, vol. 3, no. 2, pp. 1–35, 2012. DOI: <https://doi.org/10.4018/jsse.2012040101>.
- [18] P. H. Meland, I. A. Tondel and B. Solhaug, 'Mitigating risk with cyberinsurance,' *IEEE Security & Privacy*, vol. 13, no. 6, pp. 38–43, 2015. DOI: <https://doi.org/10.1109/MSP.2015.137>.
- [19] K. Bernsmed, C. Frøystad, P. H. Meland, D. A. Nesheim and Ø. J. Rødseth, 'Visualizing cyber security risks with bow-tie diagrams,' in *International Workshop on Graphical Models for Security (GraMSec)*, Springer, 2017, pp. 38–56. DOI: [https://doi.org/10.1007/978-3-319-74860-3\\_3](https://doi.org/10.1007/978-3-319-74860-3_3).
- [20] P. H. Meland, I. A. Tøndel, M. Moe and F. Seehusen, 'Facing uncertainty in cyber insurance policies,' in *International Workshop on Security and Trust Management*, Springer, 2017, pp. 89–100. DOI: [https://doi.org/10.1007/978-3-319-68063-7\\_6](https://doi.org/10.1007/978-3-319-68063-7_6).
- [21] P. H. Meland and F. Seehusen, 'When to treat security risks with cyber insurance,' *International Journal on Cyber Situational Awareness*, vol. 3, no. 1, pp. 39–60, 2018. DOI: <https://doi.org/10.22619/ijcsa.2018.100119>.

- [22] P. H. Meland, K. Bernsmed, C. Frøystad, J. Li and G. Sindre, 'An experimental evaluation of bow-tie analysis for security,' *Information & Computer Security*, vol. 27, no. 4, pp. 536–561, 2019. DOI: <https://doi.org/10.1108/ICS-11-2018-0132>.
- [23] U. Franke and P. H. Meland, 'Demand side expectations of cyber insurance,' in *2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, IEEE, Jun. 2019, pp. 1–8. DOI: <https://doi.org/10.1109/CyberSA.2019.8899685>.
- [24] P. H. Meland, B. H. Johansen and G. Sindre, 'An experimental analysis of cryptojacking attacks,' in *Nordic Conference on Secure IT Systems (NordSec)*, Springer, 2019, pp. 155–170. DOI: [https://doi.org/10.1007/978-3-030-35055-0\\_10](https://doi.org/10.1007/978-3-030-35055-0_10).
- [25] P. H. Meland and G. Sindre, 'Cyber attacks for sale,' in *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, IEEE, 2019, pp. 54–59. DOI: <https://doi.org/10.1109/CSCI49370.2019.00016>.
- [26] P. H. Meland, Y. F. F. Bayoumy and G. Sindre, 'The Ransomware-as-a-Service economy within the darknet,' *Computers & Security*, vol. 92, no. May 2020, 2020. DOI: <https://doi.org/10.1016/j.cose.2020.101762>.
- [27] K. Haga, P. H. Meland and G. Sindre, 'Breaking the cyber kill chain by modelling resource costs,' in *International Workshop on Graphical Models for Security (GramSec)*, Springer, 2020, pp. 111–126, ISBN: 978-3-030-62230-5. DOI: [https://doi.org/10.1007/978-3-030-62230-5\\_6](https://doi.org/10.1007/978-3-030-62230-5_6).
- [28] P. H. Meland, S. Tokas, G. Erdogan, K. Bernsmed and A. Omerovic, 'A systematic mapping study on cyber security indicator data,' *Electronics*, vol. 10, no. 9, p. 1092, 2021. DOI: <https://doi.org/10.3390/electronics10091092>.
- [29] P. H. Meland, D. A. Nesheim, K. Bernsmed and G. Sindre, 'Assessing cyber threats for storyless systems,' *Submitted to Information Security and Applications*, 2021, ISSN: 2214-2126.
- [30] C. Frøystad, K. Bernsmed and P. H. Meland, 'Protecting future maritime communication,' in *Proceedings of the 12th International Conference on Availability, Reliability and Security*, Association for Computing Machinery, 2017, pp. 1–10, ISBN: 9781450352574. DOI: [10.1145/3098954.3103169](https://doi.org/10.1145/3098954.3103169).
- [31] K. Bernsmed, C. Frøystad, P. H. Meland, T. A. Myrvoll *et al.*, 'Security requirements for SATCOM datalink systems for future air traffic management,' in *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, IEEE, 2017, pp. 1–10. DOI: <https://doi.org/10.1109/DASC.2017.8102083>.

- [32] P. H. Meland and F. Seehusen, 'When to treat security risks with cyber insurance,' in *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, IEEE, 2018, pp. 1–8, ISBN: 978-1-5386-4565-9. DOI: <https://doi.org/10.1109/CyberSA.2018.8551456>.
- [33] Y. F. F. Bayoumy, P. H. Meland and G. Sindre, 'A netnographic study on the dark net ecosystem for ransomware,' in *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, IEEE, 2018, pp. 1–8. DOI: <https://doi.org/10.1109/CyberSA.2018.8551424>.
- [34] K. Bernsmed, M. G. Jaatun and P. H. Meland, 'Safety critical software and security - how low can you go?' In *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*, IEEE, 2018, pp. 1–6, ISBN: 978-1-5386-4113-2. DOI: <https://doi.org/10.1109/DASC.2018.8569579>.
- [35] P. H. Meland, K. Bernsmed, C. Frøystad, J. Li and G. Sindre, 'An experimental evaluation of bow-tie analysis for cybersecurity requirements,' in *ESORICS 2018 International Workshops, CyberICPS 2018 and SECPRE 2018*, Springer, 2018, pp. 173–191, ISBN: 978-3-030-12786-2. DOI: [https://doi.org/10.1007/978-3-030-12786-2\\_11](https://doi.org/10.1007/978-3-030-12786-2_11).
- [36] Ø. J. Rødseth, P. H. Meland, C. Frøystad and O. V. Drugan, 'PKI vs. Blockchain when securing maritime operations,' *European Journal of Navigation*, vol. 18, no. 3, pp. 4–11, 2018, ISSN: 1571-473-X. [Online]. Available: <http://hdl.handle.net/11250/2612306>.
- [37] M. Branlat, P. H. Meland, T. E. Evjemo and A. Smoker, 'Connectivity and resilience of remote operations: Insights from air traffic management,' in *REA Symposium on Resilience Engineering Embracing Resilience*, 2019, ISBN: 978-91-88898-95-1. DOI: <https://doi.org/10.15626/rea8.15>.
- [38] T. Myklebust, P. H. Meland, T. Stålhane and G. K. Hanssen, 'The Agile RAMSS lifecycle for the future,' in *Proceedings of the 29th European Safety and Reliability Conference (ESREL)*, 2019. DOI: [https://doi.org/10.3850/978-981-11-2724-3\\_0170-cd](https://doi.org/10.3850/978-981-11-2724-3_0170-cd).
- [39] Ø. J. Rødseth, C. Frøystad, P. H. Meland and K. Bernsmed, 'The need for a public key infrastructure in international shipping,' in *International Maritime and Port Technology and Development Conference (MTEC)*, 2019.
- [40] Ø. J. Rødseth, C. Frøystad, P. H. Meland, K. Bernsmed and D. A. Nesheim, 'The Need for a Public Key Infrastructure for Automated and Autonomous ships,' in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, vol. 929, 2020. DOI: <http://dx.doi.org/10.1088/1757-899X/929/1/012017>.

- [41] G. Bour, K. Bernsmed, R. Borgaonkar and P. H. Meland, 'On the certificate revocation problem in the maritime sector,' in *Nordic Conference on Secure IT Systems (NordSec)*, Cham: Springer International Publishing, 2021, pp. 142–157, ISBN: 978-3-030-70852-8. DOI: [https://doi.org/10.1007/978-3-030-70852-8\\_9](https://doi.org/10.1007/978-3-030-70852-8_9).
- [42] D. A. Nesheim, Ø. J. Rødseth, B. M. v. Zernichow, P. H. Meland and K. Bernsmed, 'Secure, Trustworthy and Efficient Information Exchange – Enabling Added Value through The Maritime Data Space and Public Key Infrastructure,' in *the 20th Conference on Computer Applications and Information Technology in the Maritime Industries (COMPIT'21)*, 2021.
- [43] P. H. Meland, K. Bernsmed, E. Wille, Ø. J. Rødseth and D. A. Nesheim, 'A retrospective analysis of maritime cyber security incidents,' *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation*, vol. 15, no. 3, pp. 519–530, 2021, ISSN: 2083-6473. DOI: <https://doi.org/10.12716/1001.15.03.04>.
- [44] P. H. Meland, 'Combining threat models with security economics,' in *The 11th Norwegian Information Security Conference (NISK)*, IEEE, 2018. [Online]. Available: <https://ojs.bibsys.no/index.php/NISK/article/view/570/486>.
- [45] P. H. Meland, 'Resilient cyber security through cybercrime market analysis,' in *REA Symposium on Resilience Engineering Embracing Resilience*, 2019, ISBN: 978-91-88898-41-8. [Online]. Available: <https://open.lnu.se/index.php/rea/article/view/1975/1695>.
- [46] R. Von Solms and J. Van Niekerk, 'From information security to cyber security,' *Computers & Security*, vol. 38, pp. 97–102, 2013.
- [47] R. Ross, V. Pillitteri, R. Graubart, D. Bodeau and R. McQuaid, 'Sp 800-160 vol. 2 developing cyber resilient systems: A systems security engineering approach,' National Institute of Standards and Technology, Tech. Rep., Nov. 2019. [Online]. Available: <https://csrc.nist.gov/publications/detail/sp/800-160/vol-2/final>.
- [48] W. Gibson, 'Burning chrome,' *Omni Magazine*, Jul. 1982.
- [49] BBC, *The vocabularist: How we use the word cyber*, 2016. [Online]. Available: <https://www.bbc.com/news/magazine-35765276> (visited on 15/03/2016).
- [50] N. Wiener, 'Cybernetics: Or, control and communication in the animal and the machine,' 1958.
- [51] L. Williams, *The dr who site*, 2021. [Online]. Available: <https://thedoctorwhosite.co.uk/> (visited on 09/01/2021).



- [52] DHS, *Cyber physical systems security*, 2021. [Online]. Available: <https://www.dhs.gov/science-and-technology/cpssec> (visited on 11/01/2021).
- [53] C. W. Churchman and M. Verhulst, *Management Sciences: Models and Techniques: Proceedings of the Sixth International Meeting of the Institute of Management Sciences: Conservatoire National Des Arts & Métiers, Paris, 7-11 September 1959*. Pergamon Press, 1960, vol. 2.
- [54] ISO, 'ISO/IEC 27000 Information technology - Security techniques - Information security management systems - Overview and vocabulary,' International Organization for Standardization, Standard, Feb. 2018. [Online]. Available: <https://www.iso.org/standard/73906.html>.
- [55] É. Dubois, P. Heymans, N. Mayer and R. Matulevičius, 'A systematic approach to define the domain of information system security risk management,' in *Intentional Perspectives on Information Systems Engineering*, Springer, 2010, pp. 289–306.
- [56] G. Stoneburner, A. Goguen and A. Feringa, 'Risk management guide for information technology systems,' *NIST special publication*, vol. 800, no. 30, 2002.
- [57] ISO, 'ISO/IEC 27005 Information technology - Security techniques - Information security management systems - Information security risk management,' International Organization for Standardization, Standard, Jul. 2018. [Online]. Available: <https://www.iso.org/standard/75281.html>.
- [58] ISO, 'ISO 31000 Risk Management Guidelines,' International Organization for Standardization, Standard, Feb. 2018. [Online]. Available: <https://www.iso.org/iso-31000-risk-management.html>.
- [59] A. S. Huff, *Writing for scholarly publication*. Sage, 1999.
- [60] J. Bau and J. C. Mitchell, 'Security modeling and analysis,' *IEEE Security & Privacy*, vol. 9, no. 3, pp. 18–25, 2011.
- [61] B. Schneier, 'Threat modeling and risk assessment,' in *E-privacy*, Springer, 2000, pp. 214–229.
- [62] F. Swiderski, *Threat modeling*. Microsoft Press, 2004.
- [63] P. Torr, 'Demystifying the threat modeling process,' *IEEE Security & Privacy*, vol. 3, no. 5, pp. 66–70, 2005. DOI: [10.1109/MSP.2005.119](https://doi.org/10.1109/MSP.2005.119).
- [64] A. Shostack, *Threat modeling: Designing for security*. John Wiley & Sons, 2014.
- [65] J. Steven, 'Threat modeling - perhaps it's time,' *IEEE Security & Privacy*, vol. 8, no. 3, pp. 83–86, 2010. DOI: [10.1109/MSP.2010.110](https://doi.org/10.1109/MSP.2010.110).

- [66] D. J. Bodeau, C. D. McCollum and D. B. Fox, 'Cyber threat modeling: Survey, assessment, and representative framework,' Homeland Security Systems Engineering and Development Institute (HSSEDI), Tech. Rep., Apr. 2018. [Online]. Available: <https://apps.dtic.mil/sti/pdfs/AD1108051.pdf>.
- [67] A. Magar, 'State-of-the-art in cyber threat modeling models and methodologies,' Sphyrna Security, Tech. Rep., Mar. 2016. [Online]. Available: [https://cradpdf.drdc-rddc.gc.ca/PDFS/unc225/p803699\\_A1b.pdf](https://cradpdf.drdc-rddc.gc.ca/PDFS/unc225/p803699_A1b.pdf).
- [68] NIST, 'Framework for improving critical infrastructure cybersecurity,' Tech. Rep., Feb. 2014. [Online]. Available: <https://www.nist.gov/system/files/documents/cyberframework/cybersecurity-framework-021214.pdf>.
- [69] NIST, 'Special publication 800-30 risk management guide for information technology systems,' 2012. [Online]. Available: <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf>.
- [70] B. Rogers, 'Cbest intelligence-led testing,' Tech. Rep., 2016. [Online]. Available: <https://www.bankofengland.co.uk/-/media/boe/files/financial-stability/financial-sector-continuity/cbest-implementation-guide>.
- [71] ISACA, *Control objectives for information and related technologies (cobit)*, 2019. [Online]. Available: <https://www.isaca.org/resources/cobit> (visited on 11/01/2021).
- [72] D. Bodeau and R. Graubart, 'Cyber prep 2.0: Motivating organizational cyber strategies in terms of threat preparedness,' MITRE, Tech. Rep., 2016.
- [73] C. J. Alberts, S. G. Behrens, R. D. Pethia and W. R. Wilson, 'Operationally critical threat, asset, and vulnerability evaluation (OCTAVE) framework, version 1.0,' Carnegie Mellon Software Engineering Institute, Tech. Rep., 1999. DOI: <https://doi.org/10.1184/R1/6575906.v1>. [Online]. Available: <https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=13473>.
- [74] Z. Braiterman, A. Shostack, J. Marcil, S. d. Vries, I. Michlin, K. Wuyts, R. Hurlbut, B. S. Schoenfield, F. Scott, M. Coles, C. Romeo, A. Miller, I. Tarandach, A. Douglan and M. French, *Threat modeling manifesto*, 2020. [Online]. Available: <https://www.threatmodelingmanifesto.org/> (visited on 11/01/2021).
- [75] S. Myagmar, A. J. Lee and W. Yurcik, 'Threat modeling as a basis for security requirements,' in *Symposium on requirements engineering for information security (SREIS)*, University of Pittsburgh, vol. 2005, 2005, pp. 1–8.

- [76] I. A. Tøndel, M. G. Jaatun and P. H. Meland, ‘Security requirements for the rest of us: A survey,’ *IEEE software*, vol. 25, no. 1, pp. 20–27, 2008. doi: <https://doi.org/10.1109/MS.2008.19>.
- [77] B. Kordy, L. Piètre-Cambacédès and P. Schweitzer, ‘DAG-based attack and defense modeling: Don’t miss the forest for the attack trees,’ *Computer science review*, vol. 13, pp. 1–38, 2014.
- [78] IEC, *IEC 61025: 2006, Fault tree analysis (FTA)*, 2006.
- [79] IEC, ‘IEC 300–3–9 dependability management–part 3: Application guide–section 9: Risk analysis of technological systems,’ International Electrotechnical Commission, Tech. Rep., 1995.
- [80] B. Schneier, ‘Attack trees,’ *Dr. Dobbs’s journal*, vol. 24, no. 12, pp. 21–29, 1999.
- [81] S. Mauw and M. Oostdijk, ‘Foundations of attack trees,’ in *International Conference on Information Security and Cryptology*, Springer, 2005, pp. 186–198. doi: [https://doi.org/10.1007/11734727\\_17](https://doi.org/10.1007/11734727_17).
- [82] S. Bistarelli, F. Fioravanti and P. Peretti, ‘Defense trees for economic evaluation of security investments,’ in *First International Conference on Availability, Reliability and Security (ARES’06)*, 2006, 8 pp.–423. doi: [10.1109/ARES.2006.46](https://doi.org/10.1109/ARES.2006.46).
- [83] B. Kordy, S. Mauw, S. Radomirović and P. Schweitzer, ‘Foundations of attack–defense trees,’ in *International Workshop on Formal Aspects in Security and Trust*, Springer, 2010, pp. 80–95. doi: [https://doi.org/10.1007/978-3-642-19751-2\\_6](https://doi.org/10.1007/978-3-642-19751-2_6).
- [84] R. Kumar and M. Stoelinga, ‘Quantitative security and safety analysis with attack-fault trees,’ in *2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE)*, IEEE, 2017, pp. 25–32.
- [85] T. R. Ingoldsby, ‘Attack tree-based threat risk analysis,’ *Amenaza Technologies Limited*, pp. 3–9, 2010.
- [86] R. Jhawar, B. Kordy, S. Mauw, S. Radomirović and R. Trujillo-Rasua, ‘Attack trees with sequential conjunction,’ in *IFIP International Information Security and Privacy Conference*, Springer, 2015, pp. 339–353.
- [87] D. S. Nielsen, *The cause/consequence diagram method as a basis for quantitative accident analysis*. Risø National Laboratory, 1971.
- [88] S. Jha, O. Sheyner and J. Wing, ‘Two formal analyses of attack graphs,’ in *Proceedings 15th IEEE Computer Security Foundations Workshop. CSFW-15*, IEEE, 2002, pp. 49–63.

- [89] O. Sheyner, J. Haines, S. Jha, R. Lippmann and J. M. Wing, ‘Automated generation and analysis of attack graphs,’ in *Proceedings 2002 IEEE Symposium on Security and Privacy*, IEEE, 2002, pp. 273–284.
- [90] M. S. Lund, B. Solhaug and K. Stølen, *Model-driven risk analysis: the CORAS approach*. Springer Science & Business Media, 2010.
- [91] J. B. Hong, D. S. Kim, C.-J. Chung and D. Huang, ‘A survey on the usability and practical applications of graphical security models,’ *Computer Science Review*, vol. 26, pp. 1–16, 2017. DOI: <https://doi.org/10.1016/j.cosrev.2017.09.001>.
- [92] W. Wideł, M. Audinot, B. Fila and S. Pinchinat, ‘Beyond 2014: Formal Methods for Attack Tree-Based Security Modeling,’ *ACM Comput. Surv.*, vol. 52, no. 4, Aug. 2019, ISSN: 0360-0300. DOI: [10.1145/3331524](https://doi.org/10.1145/3331524). [Online]. Available: <https://doi.org/10.1145/3331524>.
- [93] R. Jhavar, K. Lounis and S. Mauw, ‘A stochastic framework for quantitative analysis of attack-defense trees,’ in *International Workshop on Security and Trust Management*, Springer, 2016, pp. 138–153.
- [94] L. Piètre-Cambacédès and M. Bouissou, ‘Beyond attack trees: Dynamic security modeling with boolean logic driven markov processes (bdmp),’ in *2010 European Dependable Computing Conference*, IEEE, 2010, pp. 199–208.
- [95] T. T. Teoh, Y. Y. Nguwi, Y. Elovici, N. M. Cheung and W. L. Ng, ‘Analyst intuition based hidden markov model on high speed, temporal cyber security big data,’ in *2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, 2017, pp. 2080–2083. DOI: [10.1109/FSKD.2017.8393092](https://doi.org/10.1109/FSKD.2017.8393092).
- [96] B. Kordy, M. Pouly and P. Schweitzer, ‘A probabilistic framework for security scenarios with dependent actions,’ in *International Conference on Integrated Formal Methods*, Springer, 2014, pp. 256–271.
- [97] B. Kordy, M. Pouly and P. Schweitzer, ‘Probabilistic reasoning with graphical security models,’ *Information sciences*, vol. 342, pp. 111–131, 2016.
- [98] L. Wang, T. Islam, T. Long, A. Singhal and S. Jajodia, ‘An attack graph-based probabilistic security metric,’ in *IFIP Annual Conference on Data and Applications Security and Privacy*, Springer, 2008, pp. 283–296.
- [99] S. Chockalingam, W. Pieters, A. Teixeira and P. van Gelder, ‘Bayesian network models in cyber security: A systematic review,’ in *Secure IT Systems*, H. Lipmaa, A. Mitrokotsa and R. Matulevičius, Eds., Cham: Springer International Publishing, 2017, pp. 105–122, ISBN: 978-3-319-70290-2.

- [100] T. M. Chen, J. C. Sanchez-Aarnoutse and J. Buford, 'Petri net modeling of cyber-physical attacks on smart grid,' *IEEE Transactions on smart grid*, vol. 2, no. 4, pp. 741–749, 2011.
- [101] M. Husák, J. Komárková, E. Bou-Harb and P. Čeleda, 'Survey of attack projection, prediction, and forecasting in cyber security,' *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 640–660, 2018.
- [102] M. Barrère, C. Hankin, N. Nicolaou, D. G. Eliades and T. Parisini, 'Measuring cyber-physical security in industrial control systems via minimum-effort attack strategies,' *Journal of Information Security and Applications*, vol. 52, p. 102 471, 2020, issn: 2214-2126. doi: <https://doi.org/10.1016/j.jisa.2020.102471>. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2214212619311342>.
- [103] G. Sindre and A. L. Opdahl, 'Eliciting security requirements by misuse cases,' in *Proceedings 37th International Conference on Technology of Object-Oriented Languages and Systems. TOOLS-Pacific 2000*, 2000, pp. 120–131.
- [104] I. Jacobson, *Object-oriented software engineering: a use case driven approach*. Pearson Education India, 1993.
- [105] J. McDermott and C. Fox, 'Using abuse case models for security requirements analysis,' in *Proceedings 15th Annual Computer Security Applications Conference (ACSAC'99)*, IEEE, 1999, pp. 55–64.
- [106] P. H. Meland and E. A. Gjære, 'Representing threats in BPMN 2.0,' in *2012 Seventh International Conference on Availability, Reliability and Security*, IEEE, 2012, pp. 542–550. doi: <https://doi.ieeecomputersociety.org/10.1109/ARES.2012.13>.
- [107] W. P. Stevens, G. J. Myers and L. L. Constantine, 'Structured design,' *IBM systems journal*, vol. 13, no. 2, pp. 115–139, 1974.
- [108] C.-W. Ten, C.-C. Liu and G. Manimaran, 'Vulnerability assessment of cybersecurity for scada systems,' *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1836–1846, 2008.
- [109] E. L. Lazarus, D. L. Dill, J. Epstein and J. L. Hall, 'Applying a reusable election threat model at the county level,' in *Proceedings of the 2011 Conference on Electronic Voting Technology/Workshop on Trustworthy Elections*, ser. EVT/WOTE'11, San Francisco, CA: USENIX Association, 2011, p. 12.
- [110] X. Lin, P. Zavorsky, R. Ruhl and D. Lindskog, 'Threat Modeling for CSRF Attacks,' in *2009 International Conference on Computational Science and Engineering*, IEEE, vol. 3, 2009, pp. 486–491. doi: <https://doi.org/10.1109/CSE.2009.372>.

- [111] S. Jajodia, S. Noel and B. O'berry, 'Topological analysis of network attack vulnerability,' in *Managing cyber threats*, Springer, 2005, pp. 247–266.
- [112] R. Jhawar, S. Mauw and I. Zakiuddin, 'Automating cyber defence responses using attack-defence trees and game theory,' in *European Conference on Cyber Warfare and Security*, Academic Conferences International Limited, 2016, p. 163.
- [113] P. H. Meland, I. A. Tøndel and J. Jensen, 'Idea: Reusability of Threat Models – Two Approaches with an Experimental Evaluation,' in *International Symposium on Engineering Secure Software and Systems*, Springer, 2010, pp. 114–122. DOI: [https://doi.org/10.1007/978-3-642-11747-3\\_9](https://doi.org/10.1007/978-3-642-11747-3_9).
- [114] O. Gadyatskaya, 'How to generate security cameras: Towards defence generation for socio-technical systems,' in *International Workshop on Graphical Models for Security*, Springer, 2015, pp. 50–65. DOI: [https://doi.org/10.1007/978-3-319-29968-6\\_4](https://doi.org/10.1007/978-3-319-29968-6_4).
- [115] O. Gadyatskaya, R. R. Hansen, K. G. Larsen, A. Legay, M. C. Olesen and D. B. Poulsen, 'Modelling attack-defense trees using timed automata,' in *International Conference on Formal Modeling and Analysis of Timed Systems*, Springer, 2016, pp. 35–50. DOI: [https://doi.org/10.1007/978-3-319-44878-7\\_3](https://doi.org/10.1007/978-3-319-44878-7_3).
- [116] F. Arnold, H. Hermanns, R. Pulungan and M. Stoelinga, 'Time-dependent analysis of attacks,' in *International Conference on Principles of Security and Trust*, Springer, 2014, pp. 285–305. DOI: [https://doi.org/10.1007/978-3-642-54792-8\\_16](https://doi.org/10.1007/978-3-642-54792-8_16).
- [117] R. Anderson, 'Why information security is hard-an economic perspective,' in *Seventeenth Annual Computer Security Applications Conference*, IEEE, 2001, pp. 358–365. DOI: <https://doi.org/10.1109/ACSAC.2001.991552>.
- [118] R. Anderson and T. Moore, 'Information security economics—and beyond,' in *Annual International Cryptology Conference*, Springer, 2007, pp. 68–91. DOI: [https://doi.org/10.1007/978-3-540-74143-5\\_5](https://doi.org/10.1007/978-3-540-74143-5_5).
- [119] T. Moore and R. Anderson, 'Economics and internet security: A survey of recent analytical, empirical, and behavioral research,' Tech. Rep., 2011. [Online]. Available: <https://dash.harvard.edu/handle/1/23574266>.
- [120] R. Anderson, 'Security economics: A personal perspective,' in *Proceedings of the 28th Annual Computer Security Applications Conference*, 2012, pp. 139–144.
- [121] R. Anderson and T. Moore, 'The economics of information security: A survey and open questions,' in *Fourth bi-annual Conference on the Economics of the Software and Internet Industries*, 2007, pp. 19–20.

- [122] R. Anderson and T. Moore, 'Information security: Where computer science, economics and psychology meet,' *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 367, no. 1898, pp. 2717–2727, 2009.
- [123] J. J. Cordes, 'An overview of the economics of cybersecurity and cybersecurity policy,' Tech. Rep., 2011, pp. 1–18. [Online]. Available: [http://infoseccon.net/workshop/downloads/2011/pdf/An\\_Overview\\_of\\_the\\_Economics\\_of\\_Cybersecurity\\_and\\_Cybersecurity\\_Policy.pdf](http://infoseccon.net/workshop/downloads/2011/pdf/An_Overview_of_the_Economics_of_Cybersecurity_and_Cybersecurity_Policy.pdf).
- [124] N. Jentzsch, 'State-of-the-art of the economics of cyber-security and privacy,' Tech. Rep., 2016. [Online]. Available: <http://dx.doi.org/10.2139/ssrn.2671291>.
- [125] S. L. Pfleeger and R. Rue, 'Cybersecurity economic issues: Clearing the path to good practice,' *IEEE software*, vol. 25, no. 1, pp. 35–42, 2008. doi: <https://doi.ieeecomputersociety.org/10.1109/MS.2008.4>.
- [126] G. Stoneburner, A. Goguen and A. Feringa, 'Sp800-30 risk management guide for information technology systems,' National Institute of Standards and Technology Special Publication, Tech. Rep., 2002, pp. 800–30.
- [127] M. D. Adler and E. A. Posner, 'Cost-benefit analysis: Legal, economic and philosophical perspectives,' 2000.
- [128] A. E. Boardman, D. H. Greenberg, A. R. Vining and D. L. Weimer, *Cost-benefit analysis: concepts and practice*. Cambridge University Press, 2017.
- [129] A. Sen, 'The discipline of cost-benefit analysis,' *The Journal of Legal Studies*, vol. 29, no. S2, pp. 931–952, 2000.
- [130] S. Noel, S. Jajodia, B. O'Berry and M. Jacobs, 'Efficient minimum-cost network hardening via exploit dependency graphs,' in *19th Annual Computer Security Applications Conference, 2003. Proceedings.*, IEEE, 2003, pp. 86–95.
- [131] L. A. Gordon and M. P. Loeb, 'The economics of information security investment,' *ACM Transactions on Information and System Security (TISSEC)*, vol. 5, no. 4, pp. 438–457, 2002.
- [132] ENISA, 'Introduction to return on security investment,' The European Network and Information Security Agency, Tech. Rep., Dec. 2012. [Online]. Available: <https://www.enisa.europa.eu/publications/introduction-to-return-on-security-investment>.
- [133] D. Schatz and R. Bashroush, 'Economic valuation for information security investment: A systematic literature review,' *Information Systems Frontiers*, vol. 19, no. 5, pp. 1205–1228, 2017. doi: <https://doi.org/10.1007/s10796-016-9648-8>.

- [134] M. D. Smith and M. E. Paté-Cornell, 'Cyber risk analysis for a smart grid: How smart is smart enough? a multiarmed bandit approach to cyber security investment,' *IEEE Transactions on Engineering Management*, vol. 65, no. 3, pp. 434–447, 2018. doi: [10.1109/TEM.2018.2798408](https://doi.org/10.1109/TEM.2018.2798408).
- [135] Uncredited, *Principles of economics*, 2011. [Online]. Available: <https://open.lib.umn.edu/principleseconomics/front-matter/publisher-information/> (visited on 12/01/2021).
- [136] R. Anderson, C. Barton, R. Böhme, R. Clayton, M. J. Van Eeten, M. Levi, T. Moore and S. Savage, 'Measuring the cost of cybercrime,' in *The economics of information security and privacy*, Springer, 2013, pp. 265–300.
- [137] W. Jansen, 'Research directions in security metrics,' *Journal of Information System Security*, vol. 7, no. 1, pp. 3–22, 2011.
- [138] R. Böhme, 'Security metrics and security investment models,' in *International Workshop on Security*, Springer, 2010, pp. 10–24.
- [139] R. M. Savola, 'Towards a taxonomy for information security metrics,' in *Proceedings of the 2007 ACM Workshop on Quality of Protection*, 2007, pp. 28–30.
- [140] T. Heyman, R. Scandariato, C. Huygens and W. Joosen, 'Using security patterns to combine security metrics,' in *2008 Third International Conference on Availability, Reliability and Security*, IEEE, 2008, pp. 1156–1163.
- [141] A. Jaquith, *Security metrics: replacing fear, uncertainty, and doubt*. Pearson Education, 2007.
- [142] R. Böhme and G. Schwartz, 'Modeling cyber-insurance: Towards a unifying framework,' in *The Ninth Workshop on the Economics of Information Security (WEIS 2010)*, 2010.
- [143] M. Eling and W. Schnell, 'What do we know about cyber risk and cyber risk insurance?' *The Journal of Risk Finance*, vol. 17, no. 5, pp. 474–491, 2016. doi: <https://doi.org/10.1108/JRF-09-2016-0122>.
- [144] A. Marotta, F. Martinelli, S. Nanni, A. Orlando and A. Yautsiukhin, 'Cyber-insurance survey,' *Computer Science Review*, vol. 24, pp. 35–61, 2017. doi: <https://doi.org/10.1016/j.cosrev.2017.01.001>.
- [145] D. Woods and T. Moore, 'Does insurance have a future in governing cybersecurity?' *IEEE Security and Privacy Magazine*, 2019. doi: <https://doi.org/10.1109/MSEC.2019.2935702>.



- [146] S. S. Wang, 'Integrated framework for information security investment and cyber insurance,' *Pacific-Basin Finance Journal*, vol. 57, p. 101-173, 2019, ISSN: 0927-538X. DOI: <https://doi.org/10.1016/j.pacfin.2019.101173>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0927538X19300794>.
- [147] L. A. Gordon, M. P. Loeb and W. Lucyshyn, 'Sharing information on computer systems security: An economic analysis,' *Journal of Accounting and Public Policy*, vol. 22, no. 6, pp. 461–485, 2003.
- [148] E. Gal-Or and A. Ghose, 'The economic incentives for sharing security information,' *Information Systems Research*, vol. 16, no. 2, pp. 186–208, 2005.
- [149] R. Anderson, R. Böhme, R. Clayton and T. Moor, 'Security economics and european policy,' in *ISSE 2008 Securing Electronic Business Processes*, Springer, 2009, pp. 57–76.
- [150] S. Laube and R. Böhme, 'Strategic aspects of cyber risk information sharing,' *ACM Computing Surveys (CSUR)*, vol. 50, no. 5, pp. 1–36, 2017.
- [151] A. Alexeev, E. Jardine and K. Krutilla, 'Optimal investment in cyber attack and resilience: A dynamic differential game,' [Online]. Available: <http://ceur-ws.org/Vol-2040/paper14.pdf> (visited on 12/01/2021).
- [152] T. Kissoon, 'Optimum spending on cybersecurity measures,' *Transforming Government: People, Process and Policy*, vol. 14, no. 3, pp. 417–431, 2020, ISSN: 1750-6166. DOI: <https://doi.org/10.1108/TG-11-2019-0112>.
- [153] L. A. Gordon, M. P. Loeb and L. Zhou, 'Information segmentation and investing in cybersecurity,' *Journal of Information Security*, vol. 12, no. 1, pp. 115–136, 2021. DOI: <https://doi.org/10.4236/jis.2021.121006>. [Online]. Available: <https://www.scirp.org/journal/paperinformation.aspx?paperid=106601>.
- [154] S. E. Schechter and M. D. Smith, 'How much security is enough to stop a thief?' In *Financial Cryptography*, R. N. Wright, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 122–137, ISBN: 978-3-540-45126-6.
- [155] D. Florêncio and C. Herley, 'Sex, lies and cyber-crime surveys,' in *Economics of information security and privacy III*, Springer, 2013, pp. 35–53.
- [156] C. Herley, 'The plight of the targeted attacker in a world of scale.,' in *The 9th Workshop on the Economics of Information Security (WEIS)*, 2010.

- [157] C. Herley, 'Why do Nigerian Scammers Say They are from Nigeria?' In *The Eleventh Workshop on the Economics of Information Security (WEIS)*, 2012. [Online]. Available: <https://www.microsoft.com/en-us/research/publication/why-do-nigerian-scammers-say-they-are-from-nigeria/>.
- [158] M. Goncharov, 'Russian underground 101,' Trend Micro Incorporated, Tech. Rep., 2012, p. 26. [Online]. Available: <https://www.trendmicro.de/cloud-content/us/pdfs/security-intelligence/white-papers/wp-russian-underground-101.pdf>.
- [159] A. K. Sood, R. Bansal and R. J. Enbody, 'Cybercrime: Dissecting the state of underground enterprise,' *IEEE internet computing*, vol. 17, no. 1, pp. 60–68, 2012.
- [160] R. Thomas and J. Martin, 'The underground economy: Priceless,' *login*, vol. 31, no. 6, 2006.
- [161] D. Florêncio and C. Herley, 'Where do all the attacks go?' In *Economics of information security and privacy III*, Springer, 2013, pp. 13–33.
- [162] N. Kshetri, 'Simple economics of cybercrime and the vicious circle,' in *The Global Cybercrime Industry: Economic, Institutional and Strategic Perspectives*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 35–55, ISBN: 978-3-642-11522-6. DOI: 10.1007/978-3-642-11522-6\_2. [Online]. Available: [https://doi.org/10.1007/978-3-642-11522-6\\_2](https://doi.org/10.1007/978-3-642-11522-6_2).
- [163] N. Fultz and J. Grossklags, 'Blue versus red: Towards a model of distributed security attacks,' in *International Conference on Financial Cryptography and Data Security*, Springer, 2009, pp. 167–183.
- [164] K. Huang, M. Siegel and S. Madnick, 'Systematically understanding the cyber attack business: A survey,' *ACM Computing Surveys (CSUR)*, vol. 51, no. 4, pp. 1–36, 2018. [Online]. Available: <https://doi.org/10.1145/3199674>.
- [165] T. Casey, P. Koeberl and C. Vishik, 'Defining threat agents: Towards a more complete threat analysis,' in *ISSE 2010 Securing Electronic Business Processes*, Springer, 2011, pp. 214–225.
- [166] R. Broadhurst, D. Lord, D. Maxim, H. Woodford-Smith, C. Johnston, H. W. Chung, S. Carroll, H. Trivedi and B. Sabol, 'Malware trends on 'darknet' crypto-markets: Research review,' Tech. Rep., 2018. [Online]. Available: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3226758](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3226758).
- [167] A. Buldas, P. Laud, J. Priisalu, M. Saarepera and J. Willemson, 'Rational choice of security measures via multi-parameter attack trees,' in *International Workshop on Critical Information Infrastructures Security*, Springer, 2006, pp. 235–248.

- [168] D. Geer, E. Jardine and E. Leverett, 'On market concentration and cybersecurity risk,' *Journal of Cyber Policy*, vol. 0, no. 0, pp. 1–21, 2020. DOI: 10.1080/23738871.2020.1728355. eprint: <https://doi.org/10.1080/23738871.2020.1728355>. [Online]. Available: <https://doi.org/10.1080/23738871.2020.1728355>.
- [169] R. Slayton, 'What is the cyber offense-defense balance? conceptions, causes, and assessment,' *International Security*, vol. 41, no. 3, pp. 72–109, 2017. DOI: 10.1162/ISEC\_a\_00267. eprint: [https://doi.org/10.1162/ISEC\\_a\\_00267](https://doi.org/10.1162/ISEC_a_00267). [Online]. Available: [https://doi.org/10.1162/ISEC\\_a\\_00267](https://doi.org/10.1162/ISEC_a_00267).
- [170] J. Clark and W. L. Davis, 'A human capital perspective on criminal careers,' *Journal of Applied Business Research (JABR)*, vol. 11, no. 3, pp. 58–64, 1995.
- [171] I. Ehrlich, 'Crime, punishment, and the market for offenses,' *Journal of economic perspectives*, vol. 10, no. 1, pp. 43–67, 1996.
- [172] J. Grossklags, N. Christin and J. Chuang, 'Secure or insure? a game-theoretic analysis of information security games,' in *Proceedings of the 17th international conference on World Wide Web*, 2008, pp. 209–218.
- [173] J. Wang, Y. Yi, H. Zhang and N. Cao, 'Network attack prediction method based on threat intelligence,' in *Cloud Computing and Security*, X. Sun, Z. Pan and E. Bertino, Eds., Cham: Springer International Publishing, 2018, pp. 151–160, ISBN: 978-3-030-00012-7.
- [174] S.K., *What is the nash equilibrium and why does it matter?* Sep. 2016. [Online]. Available: <https://www.economist.com/the-economist-explains/2016/09/06/what-is-the-nash-equilibrium-and-why-does-it-matter> (visited on 14/01/2021).
- [175] M. Kassner, *How game theory and nash equilibrium can help decide cybersecurity responses*, May 2017. [Online]. Available: <https://www.techrepublic.com/article/how-game-theory-and-nash-equilibrium-can-help-decide-cybersecurity-responses/> (visited on 14/01/2021).
- [176] S. Shiva, S. Roy and D. Dasgupta, 'Game theory for cyber security,' in *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research*, 2010, pp. 1–4.
- [177] S. Roy, C. Ellis, S. Shiva, D. Dasgupta, V. Shandilya and Q. Wu, 'A survey of game theory as applied to network security,' in *2010 43rd Hawaii International Conference on System Sciences*, IEEE, 2010, pp. 1–10.
- [178] Y. Wang, Y. Wang, J. Liu, Z. Huang and P. Xie, 'A survey of game theoretic methods for cyber security,' in *2016 IEEE First International Conference on Data Science in Cyberspace (DSC)*, 2016, pp. 631–636. DOI: 10.1109/DSC.2016.90.

- [179] C. T. Do, N. H. Tran, C. Hong, C. A. Kamhoua, K. A. Kwiat, E. Blasch, S. Ren, N. Pissinou and S. S. Iyengar, ‘Game theory for cyber security and privacy,’ *ACM Computing Surveys (CSUR)*, vol. 50, no. 2, pp. 1–37, 2017.
- [180] J. Pawlick, E. Colbert and Q. Zhu, ‘A game-theoretic taxonomy and survey of defensive deception for cybersecurity and privacy,’ *ACM Comput. Surv.*, vol. 52, no. 4, Aug. 2019, ISSN: 0360-0300. DOI: <https://doi.org/10.1145/3337772>.
- [181] P. Dasgupta and J. B. Collins, ‘A survey of game theoretic approaches for adversarial machine learning in cybersecurity tasks,’ *AI Mag.*, vol. 40, pp. 31–43, 2019. DOI: <https://doi.org/10.1609/aimag.v40i2.2847>.
- [182] A. K. Sen, A. G. M. Last and R. Quirk, ‘Prediction and economic theory [and discussion],’ *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 407, no. 1832, pp. 3–23, 1986, ISSN: 00804630. [Online]. Available: <http://www.jstor.org/stable/2397778>.
- [183] R. Clayton, T. Moore and N. Christin, ‘Concentrating correctly on cybercrime concentration.,’ in *The 2015 Workshop on the Economics of Information Security (WEIS)*, 2015.
- [184] S. Morgan, *Cybercrime to cost the world \$10.5 trillion annually by 2025*, 2020. [Online]. Available: <https://cybersecurityventures.com/cybercrime-damages-6-trillion-by-2021/> (visited on 13/11/2020).
- [185] K. Bissell, R. M. Lasalle and P. D. Cin, ‘State of cybersecurity report 2020,’ Accenture, Tech. Rep., 2020. [Online]. Available: [https://www.accenture.com/\\_acnmedia/PDF-116/Accenture-Cybersecurity-Report-2020.pdf](https://www.accenture.com/_acnmedia/PDF-116/Accenture-Cybersecurity-Report-2020.pdf).
- [186] C. Biener, M. Eling and J. H. Wirfs, ‘Insurability of cyber risk: An empirical analysis,’ *The Geneva Papers on Risk and Insurance-Issues and Practice*, vol. 40, no. 1, pp. 131–158, 2015. DOI: <https://doi.org/10.1057/gpp.2014.19>.
- [187] A. Marotta, F. Martinelli, S. Nanni and A. Yautsiukhin, ‘A survey on cyber-insurance,’ *Technical Rep. IIT TR-17/2015. Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche, Pisa*, 2015.
- [188] R. Böhme, S. Laube and M. Riek, ‘A fundamental approach to cyber risk analysis,’ *Variance*, vol. 12, no. 2, pp. 161–185, 2019.
- [189] K. J. S. Hoo, ‘How much is enough? a risk management approach to computer security,’ Ph.D. dissertation, Stanford University Stanford, 2000.

- [190] G. Zeller and M. Scherer, ‘A comprehensive model for cyber risk based on marked point processes and its application to insurance,’ Jul. 2020. [Online]. Available: <https://dx.doi.org/10.2139/ssrn.3668228>.
- [191] H. Abdo, M. Kaouk, J.-M. Flaus and F. Masse, ‘A safety/security risk analysis approach of industrial control systems: A cyber bowtie—combining new version of attack tree with bowtie analysis,’ *Computers & Security*, vol. 72, pp. 175–195, 2018. DOI: <https://doi.org/10.1016/j.cose.2017.09.004>.
- [192] J. M. Ahrend and M. Jirotko, ‘Anticipation in cyber-security,’ in *Handbook of Anticipation: Theoretical and Applied Aspects of the Use of Future in Decision Making*, R. Poli, Ed. Cham: Springer International Publishing, 2017, pp. 1–28, ISBN: 978-3-319-31737-3. DOI: [10.1007/978-3-319-31737-3\\_26-1](https://doi.org/10.1007/978-3-319-31737-3_26-1). [Online]. Available: [https://doi.org/10.1007/978-3-319-31737-3\\_26-1](https://doi.org/10.1007/978-3-319-31737-3_26-1).
- [193] M. Almukaynizi, E. Marin, M. Shah, E. Nunes, G. I. Simari and P. Shakarian, ‘A logic programming approach to predict enterprise-targeted cyberattacks,’ in *Data Science in Cybersecurity and Cyberthreat Intelligence*, L. F. Sikos and K.-K. R. Choo, Eds. Cham: Springer International Publishing, 2020, pp. 13–32, ISBN: 978-3-030-38788-4. DOI: [https://doi.org/10.1007/978-3-030-38788-4\\_2](https://doi.org/10.1007/978-3-030-38788-4_2).
- [194] E. Doynikova and I. Kotenko, ‘Enhancement of probabilistic attack graphs for accurate cyber security monitoring,’ in *2017 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computed, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCCom/IOP/SCI)*, 2017, pp. 1–6. DOI: [10.1109/UIC-ATC.2017.8397618](https://doi.org/10.1109/UIC-ATC.2017.8397618).
- [195] S. Brown, J. Gommers and O. Serrano, ‘From cyber security information sharing to threat management,’ in *Proceedings of the 2nd ACM Workshop on Information Sharing and Collaborative Security*, ser. WISCS ’15, Denver, Colorado, USA: Association for Computing Machinery, 2015, pp. 43–49, ISBN: 9781450338226. DOI: [10.1145/2808128.2808133](https://doi.org/10.1145/2808128.2808133). [Online]. Available: <https://doi.org/10.1145/2808128.2808133>.
- [196] R. McMillan, *Definition: Threat intelligence*, G00249251, May 2013. [Online]. Available: <https://www.gartner.com/en/documents/2487216/definition-threat-intelligence>.
- [197] W. Tounsi and H. Rais, ‘A survey on technical threat intelligence in the age of sophisticated cyber attacks,’ *Computers & Security*, vol. 72, pp. 212–233, 2018, ISSN:

- 0167-4048. DOI: <https://doi.org/10.1016/j.cose.2017.09.001>. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167404817301839>.
- [198] D. Chismon and M. Ruks, 'Threat intelligence: Collecting, analysing, evaluating,' MWR InfoSecurity Ltd, Tech. Rep., 2015.
- [199] T. D. Wagner, E. Palomar, K. Mahbub and A. E. Abdallah, 'Towards an anonymity supported platform for shared cyber threat intelligence,' in *Risks and Security of Internet and Systems*, Cham: Springer International Publishing, 2018, pp. 175–183, ISBN: 978-3-319-76687-4. DOI: [https://doi.org/10.1007/978-3-319-76687-4\\_12](https://doi.org/10.1007/978-3-319-76687-4_12).
- [200] A. Tundis, S. Ruppert and M. Mühlhäuser, 'On the automated assessment of open-source cyber threat intelligence sources,' in *International Conference on Computational Science*, Springer, 2020, pp. 453–467. DOI: [https://doi.org/10.1007/978-3-030-50417-5\\_34](https://doi.org/10.1007/978-3-030-50417-5_34).
- [201] H. Griffioen, T. Booij and C. Doerr, 'Quality evaluation of cyber threat intelligence feeds,' in *Applied Cryptography and Network Security*, Cham: Springer International Publishing, 2020, pp. 277–296, ISBN: 978-3-030-57878-7. DOI: [https://doi.org/10.1007/978-3-030-57878-7\\_14](https://doi.org/10.1007/978-3-030-57878-7_14).
- [202] Wikipedia, *Aviation*, 2020. [Online]. Available: <https://en.wikipedia.org/wiki/Aviation> (visited on 21/12/2020).
- [203] History.com, *September 11 attacks*, 2020. [Online]. Available: <https://www.history.com/topics/21st-century/9-11-attacks> (visited on 16/10/2020).
- [204] J. Mullen, *Report: Mh17 hit by burst of high-energy objects from outside*, 2014. [Online]. Available: <https://edition.cnn.com/2014/09/09/world/europe/netherlands-ukraine-mh17-report/index.html> (visited on 09/09/2014).
- [205] H. Marud, *Somali plane bomb: What happened?* 2016. [Online]. Available: <https://www.bbc.com/news/world-africa-35521646> (visited on 13/02/2016).
- [206] K. Sampigethaya and R. Poovendran, 'Aviation cyber–physical systems: Foundations for future aircraft and air transport,' *Proceedings of the IEEE*, vol. 101, no. 8, pp. 1834–1855, 2013.
- [207] L. H. Mutuel, P. Neri and E. Paricaud, 'Initial 4d trajectory management concept evaluation,' in *Tenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2013)*, 2013.
- [208] SESARJU, *Sesar joint undertaking*, 2020. [Online]. Available: <https://www.sesarju.eu/> (visited on 21/12/2020).

- [209] FAA, *Next generation air transportation system*, 2020. [Online]. Available: <https://www.faa.gov/nextgen/> (visited on 21/12/2020).
- [210] N. Collins, 'Cyber terrorism is biggest threat to aircraft,' *The Telegraph*, 27th Dec. 2013. [Online]. Available: <https://www.telegraph.co.uk/finance/newsbysector/transport/10526620/Cyber-terrorism-is-biggest-threat-to-aircraft.html> (visited on 20/02/2020).
- [211] ESA, *Iris precursor - iris precursor service development*, 2020. [Online]. Available: <https://artes.esa.int/projects/iris-precursor> (visited on 21/12/2020).
- [212] T. Myklebust and T. Stålhane, *The agile safety case*. Springer, 2018.
- [213] ESA, *Iris service evolution*, 2020. [Online]. Available: <https://artes.esa.int/projects/iris-service-evolution> (visited on 21/12/2020).
- [214] RemoteTower, *Project pj05 remote tower for multiple airports*, 2020. [Online]. Available: <https://www.remote-tower.eu/> (visited on 21/12/2020).
- [215] DHS, 'National maritime domain awareness plan for national strategy for maritime security,' Homeland Security Digital Library, Tech. Rep., Dec. 2013.
- [216] C. A. Kontovas and H. N. Psaraftis, 'Formal safety assessment: A critical review,' *Marine technology*, vol. 46, no. 1, p. 45, 2009.
- [217] D. Cimpean, J. Meire, V. Bouckaert, S. Vande Castele, A. Pelle and L. Hellebooge, 'Analysis of cyber security aspects in the maritime sector,' ENISA, Tech. Rep., 2011. [Online]. Available: [https://www.enisa.europa.eu/publications/cyber-security-aspects-in-the-maritime-sector-1/at\\_download/fullReport](https://www.enisa.europa.eu/publications/cyber-security-aspects-in-the-maritime-sector-1/at_download/fullReport).
- [218] G. Dobie, H. Kidston, T. Chamberlain and C. Fields, 'Safety and shipping review 2015: An annual review of trends and developments in shipping losses and safety,' Tech. Rep., 2015, p. 36. [Online]. Available: <https://www.agcs.allianz.com/content/dam/onemarketing/agcs/agcs/reports/AGCS-Safety-Shipping-Review-2015.pdf>.
- [219] L. A. LaRocco, *Suez canal blockage is delaying an estimated \$400 million an hour in goods*, Mar. 2021. [Online]. Available: <https://www.cnbc.com/2021/03/25/suez-canal-blockage-is-delaying-an-estimated-400-million-an-hour-in-goods.html> (visited on 25/03/2021).
- [220] Seatrade, 'Shipping next 'playground for hackers' warns imb,' *Maritime News*, 21st Aug. 2014. [Online]. Available: <https://www.seatrade-maritime.com/americas/shipping-next-playground-hackers-warns-imb> (visited on 20/02/2020).

- [221] P. H. Meland, K. Bernsmed, Ø. J. Rødseth and D. A. Nesheim, 'Trusselvurdering i forbindelse med strategi for maritim digital sikkerhet,' SINTEF, Tech. Rep., Jan. 2021. [Online]. Available: <https://www.sdir.no/contentassets/174739a55adb44098b05bcb8ef3b2f65/trusselvurdering-i-forbindelse-med-strategi-for-maritim-digital-sikkerhet.pdf?t=1610546311360>.
- [222] A. Greenberg, 'The untold story of NotPetya, the most devastating cyberattack in history,' *Wired*, Aug. 2018. [Online]. Available: <https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/>.
- [223] CySiMS, *Cyber security in merchant shipping*, 2020. [Online]. Available: <http://cysims.no/> (visited on 10/12/2020).
- [224] IALA, *VDES - VHF Data Exchange System*, 2020. [Online]. Available: <https://www.iala-aism.org/technical/connectivity/vdes-vhf-data-exchange-system/> (visited on 21/12/2020).
- [225] CyberSec4Europe, *Cyber security for europe, grant agreement no. 830929*, 2021. [Online]. Available: <https://cybersec4europe.eu/> (visited on 01/04/2021).
- [226] Kystverket, *Vessel traffic service*, 2020. [Online]. Available: [https://www.kystverket.no/en/EN\\_Maritime-Services/Vessel-Traffic-Service/](https://www.kystverket.no/en/EN_Maritime-Services/Vessel-Traffic-Service/) (visited on 21/12/2020).
- [227] E. Baranoff, P. Brockett and Y. Kahane, *Risk Management for Enterprises and Individuals, v. 1.0 [electronic publication]*. Irvington, NY: Flat World Knowledge. Saylor Academy, 2012. [Online]. Available: [https://saylordotorg.github.io/text\\_risk-management-for-enterprises-and-individuals/index.html](https://saylordotorg.github.io/text_risk-management-for-enterprises-and-individuals/index.html) (visited on 23/12/2020).
- [228] SINTEF, *Insecurance*, 2020. [Online]. Available: <https://www.sintef.no/en/digital/software-and-service-innovation/secure-iot-software/insecurance/> (visited on 01/12/2020).
- [229] I. A. Tøndel, P. H. Meland, A. Omerovic, E. A. Gjære and B. Solhaug, 'Using cyber-insurance as a risk management strategy: Knowledge gaps and recommendations for further research,' SINTEF, Tech. Rep., Nov. 2015. [Online]. Available: <https://sintef.brage.unit.no/sintef-xmlui/handle/11250/2379189>.
- [230] N. Dewan, *Indian Life and Health Insurance Industry: A Marketing Approach*. Springer Science & Business Media, 2008. doi: <https://doi.org/10.1007/978-3-8349-9788-3>.



- [231] S. Romanosky, L. Ablon, A. Kuehn and T. Jones, 'Content analysis of cyber insurance policies: How do carriers price cyber risk?' *Journal of Cybersecurity*, vol. 5, no. 1, tyz002, 2019. DOI: <https://doi.org/10.1093/cybsec/tyz002>.
- [232] C. Toregas and N. Zahn, 'Insurance for cyber attacks: The issue of setting premiums in context,' George Washington University, Tech. Rep., 2014. [Online]. Available: [https://cspri.seas.gwu.edu/sites/g/files/zaxdzs1446/f/downloads/cyberinsurance\\_paper\\_pdf\\_0.pdf](https://cspri.seas.gwu.edu/sites/g/files/zaxdzs1446/f/downloads/cyberinsurance_paper_pdf_0.pdf).
- [233] T. Bandyopadhyay, V. S. Mookerjee and R. C. Rao, 'Why it managers don't go for cyber-insurance products,' *Communications of the ACM*, vol. 52, no. 11, pp. 68–73, 2009.
- [234] N. Robinson, 'Incentives and barriers of the cyber insurance market in europe,' European Network, Information Security Agency (ENISA) and RAND Europe, Tech. Rep., 2012. [Online]. Available: <https://www.enisa.europa.eu/publications/incentives-and-barriers-of-the-cyber-insurance-market-in-europe>.
- [235] M. F. Carfora, F. Martinelli, F. Mercaldo, A. Orlando and A. Yautsiukhin, 'Cyber Risk Management: A New Challenge for Actuarial Mathematics,' in *Mathematical and Statistical Methods for Actuarial Sciences and Finance: MAF 2018*. Cham: Springer International Publishing, 2018, pp. 199–202, ISBN: 978-3-319-89824-7. DOI: [https://doi.org/10.1007/978-3-319-89824-7\\_36](https://doi.org/10.1007/978-3-319-89824-7_36).
- [236] J. Dillon, 'The classification of research questions,' *Review of Educational Research*, vol. 54, no. 3, pp. 327–361, 1984.
- [237] BMS, *How to formulate a research question*, 2020. [Online]. Available: <https://www.utwente.nl/en/bms/m-store/manuals/toolbox/ResearchQuestion/> (visited on 02/10/2020).
- [238] H. A. Simon, *The sciences of the artificial*. MIT Press, 1996.
- [239] S. L. McGregor and J. A. Murnane, 'Paradigm, methodology and method: Intellectual integrity in consumer scholarship,' *International journal of consumer studies*, vol. 34, no. 4, pp. 419–427, 2010.
- [240] S. Weber, 'Design science research: Paradigm or approach?' In *AMCIS*, 2010, p. 214. [Online]. Available: <https://aisel.aisnet.org/amcis2010/214>.
- [241] A. R. Hevner, S. T. March, J. Park and S. Ram, 'Design science in information systems research,' *MIS quarterly*, pp. 75–105, 2004.
- [242] G. Goldkuhl, 'Pragmatism vs interpretivism in qualitative information systems research,' *European journal of information systems*, vol. 21, no. 2, pp. 135–146, 2012.

- [243] A. O. Lovejoy, 'The thirteen pragmatisms. i,' *The Journal of Philosophy, Psychology and Scientific Methods*, vol. 5, no. 1, pp. 5–12, 1908.
- [244] I. Walsh, J. A. Holton, L. Bailyn, W. Fernandez, N. Levina and B. Glaser, 'What grounded theory is. . . a critically reflective conversation among scholars,' *Organizational Research Methods*, vol. 18, no. 4, pp. 581–599, 2015.
- [245] UiO, *INF9571 Action Research Workshop*, 2017. [Online]. Available: <https://www.uio.no/studier/emner/matnat/ifi/INF9571/index-eng.html>.
- [246] G. Goldkuhl, 'Action research vs. design research: Using practice research as a lens for comparison and integration,' in *The 2nd international SIG Prag workshop on IT Artefact Design & Workpractice Improvement (ADWI-2013), 5 June, 2013, Tilburg, the Netherlands*, 2013.
- [247] J. Iivari and J. R. Venable, 'Action research and design science research-seemingly similar but decisively dissimilar,' in *ECIS*, 2009. [Online]. Available: <https://aisel.aisnet.org/ecis2009/73>.
- [248] G. Goldkuhl, 'The research practice of practice research: Theorizing and situational inquiry,' *Systems, Signs & Actions*, vol. 5, no. 1, pp. 7–29, 2011.
- [249] D. Plowright, 'To what extent do postgraduate students understand the principles of mixed methods in educational research?' *International Journal of Multiple Research Approaches*, vol. 7, no. 1, pp. 66–82, 2013.
- [250] J. W. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches*, 4th ed. Sage publications, 2014.
- [251] P. Brereton, B. A. Kitchenham, D. Budgen, M. Turner and M. Khalil, 'Lessons from applying the systematic literature review process within the software engineering domain,' *Journal of systems and software*, vol. 80, no. 4, pp. 571–583, 2007.
- [252] S. Harvey-Jordan and S. Long, 'The process and the pitfalls of semi-structured interviews,' *Community Practitioner*, vol. 74, no. 6, p. 219, 2001.
- [253] R. V. Kozinets, *Netnography: The essential guide to qualitative social media research*. SAGE Publications Limited, 2019.
- [254] S. Gregor and A. R. Hevner, 'Positioning and presenting design science research for maximum impact,' *MIS quarterly*, pp. 337–355, 2013.
- [255] D. B. Hofler, 'Approach, method, technique a clarification,' *Reading World*, vol. 23, no. 1, pp. 71–72, 1983. DOI: <https://doi.org/10.1080/19388078309557742>.

- [256] ScienceDirect, *Test environment*, 2020. [Online]. Available: <https://www.sciencedirect.com/topics/computer-science/test-environment> (visited on 30/11/2020).
- [257] Guru99, *Test environment for software testing*, 2020. [Online]. Available: <https://www.guru99.com/test-environment-software-testing.html> (visited on 30/11/2020).
- [258] B. A. Kitchenham, 'Evaluating software engineering methods and tool part 1: The evaluation context and evaluation methods,' *ACM SIGSOFT Software Engineering Notes*, vol. 21, no. 1, pp. 11–14, 1996.
- [259] S. L. Pfleeger, 'Experimental design and analysis in software engineering,' *Annals of Software Engineering*, vol. 1, no. 1, pp. 219–253, 1995.
- [260] M. Tedre and N. Moisseinen, 'Experiments in computing: A survey,' *The Scientific World Journal*, vol. 2014, 2014.
- [261] CASRAI, *CRedit – Contributor Roles Taxonomy*, CASRAI, 2020. [Online]. Available: <https://casrai.org/credit/> (visited on 29/09/2020).
- [262] SHIELDS, *Detecting known security vulnerabilities from within design and development tools*, 2010. [Online]. Available: <https://cordis.europa.eu/project/id/215995> (visited on 10/10/2020).
- [263] D. W. Hubbard and R. Seiersen, *How to measure anything in cybersecurity risk*. Wiley Online Library, 2016.
- [264] P. Santini, G. Gottardi, M. Baldi and F. Chiaraluce, 'A data-driven approach to cyber risk assessment,' *Security and Communication Networks*, vol. 2019, 2019. doi: <https://doi.org/10.1155/2019/6716918>.
- [265] P. Tubío Figueira, C. López Bravo and J. L. Rivas López, 'Improving information security risk analysis by including threat-occurrence predictive models,' *Computers & Security*, vol. 88, p. 101 609, 2020, ISSN: 0167-4048. doi: <https://doi.org/10.1016/j.cose.2019.101609>. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167404819301592>.
- [266] T. Kissoon, 'Optimum Spending on Cybersecurity Measures: Part II,' *Journal of Information Security*, vol. 12, no. 1, pp. 137–161, 2021. doi: <https://doi.org/10.4236/jis.2021.121007>. [Online]. Available: <https://www.scirp.org/journal/paperinformation.aspx?paperid=106651>.
- [267] N. Al-Hadhrami, M. Collinson and N. Oren, 'Modelling security risk scenarios using subjective attack trees,' *Risks and Security of Internet and Systems*, 2021. doi: [https://doi.org/10.1007/978-3-030-68887-5\\_12](https://doi.org/10.1007/978-3-030-68887-5_12).

- [268] H. Holm, T. Sommestad, M. Ekstedt and N. Honeth, 'Indicators of expert judgement and their significance: An empirical investigation in the area of cyber security,' *Expert Systems*, vol. 31, no. 4, pp. 299–318, 2014.
- [269] G. Falco, M. Eling, D. Jablanski, M. Weber, V. Miller, L. A. Gordon, S. S. Wang, J. Schmit, R. Thomas, M. Elvedi, T. Maillart, E. Donavan, S. Dejung, E. Durand, F. Nutter, U. Scheffer, G. Arazi, G. Ohana and H. Lin, 'Cyber risk research impeded by disciplinary barriers,' *Science*, vol. 366, no. 6469, pp. 1066–1069, 2019, ISSN: 0036-8075. DOI: <https://doi.org/10.1126/science.aaz4795>. [Online]. Available: <https://science.sciencemag.org/content/366/6469/1066>.
- [270] I. Salman, A. T. Misirli and N. Juristo, 'Are students representatives of professionals in software engineering experiments?' In *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, IEEE, vol. 1, 2015, pp. 666–676.
- [271] M. Svahnberg, A. Aurum and C. Wohlin, 'Using students as subjects-an empirical evaluation,' in *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement*, 2008, pp. 288–290.
- [272] M. Höst, C. Wohlin and T. Thelin, 'Experimental context classification: Incentives and experience of subjects,' in *Proceedings of the 27th international conference on Software engineering*, 2005, pp. 470–478.
- [273] J. Hallberg, J. Bengtsson, N. Hallberg, H. Karlzén and T. Sommestad, 'The significance of information security risk assessments: Exploring the consensus of raters' perceptions of probability and severity,' in *Proceedings of the International Conference on Security and Management (SAM)*, 2017, pp. 131–137.
- [274] I. L. Janis, 'Groupthink,' *Psychology today*, vol. 5, no. 6, pp. 43–46, 1971.
- [275] G. Orwell, *Nineteen Eighty-Four*. Secker & Warburg, 1949.
- [276] C. Knez, T. Llansó, D. Pearson, T. Schonfeld and K. Sotzen, 'Lessons learned from applying cyber risk management and survivability concepts to a space mission,' in *2016 IEEE Aerospace Conference*, IEEE, 2016, pp. 1–8.
- [277] T. Llansó, M. McNeil, D. Pearson and G. Moore, 'Blugen: An analytic framework for mission-cyber risk assessment and mitigation recommendation,' in *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017, ISBN: 978-0-9981331-0-2. DOI: <https://doi.org/10.24251/HICSS.2017.724>.
- [278] M. McNeil, T. Llansó and D. Pearson, 'Application of capability-based cyber risk assessment methodology to a space system,' in *Proceedings of the 5th Annual Symposium and Bootcamp on Hot Topics in the Science of Security*, 2018, pp. 1–10.

- [279] R. Anderson, C. Barton, R. Bölme, R. Clayton, C. Ganán, T. Grasso, M. Levi, T. Moore and M. Vasek, 'Measuring the changing cost of cybercrime,' in *18th Workshop on the Economics of Information Security (WEIS)*, 2019.
- [280] M. Brecht and T. Nowey, 'A closer look at information security costs,' in *The Economics of Information Security and Privacy*, Springer, 2013, pp. 3–24.
- [281] GenevaAssociation, *Exploring the opportunity for a Cyber Incident Data Exchange and Repository (CIDER)*, 2020. [Online]. Available: <https://www.genevaassociation.org/news/articles-interest/exploring-opportunity-cyber-incident-data-exchange-and-repository-cider> (visited on 09/02/2021).
- [282] J. R. Nurse, L. Axon, A. Erola, I. Agrafiotis, M. Goldsmith and S. Creese, 'The data that drives cyber insurance: A study into the underwriting and claims processes,' in *2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, IEEE, 2020. DOI: <https://doi.org/10.1109/CyberSA49311.2020.9139703>.
- [283] L. A. Gordon, M. P. Loeb, W. Lucyshyn and R. Richardson, '2006 CSI/FBI computer crime and security survey,' *Computer Security Journal*, vol. 22, no. 3, p. 1, 2006.
- [284] P. A. Wortman and J. A. Chandy, 'Smart: Security model adversarial risk-based tool for systems security design evaluation,' *Journal of Cybersecurity*, vol. 1, p. 8, 2020.
- [285] C. Konradt, A. Schilling and B. Werners, 'Phishing: An economic analysis of cybercrime perpetrators,' *Computers & Security*, vol. 58, pp. 39–46, 2016, ISSN: 0167-4048. DOI: <https://doi.org/10.1016/j.cose.2015.12.001>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167404815001844>.
- [286] M. Saad, A. Khormali and A. Mohaisen, 'Dine and dash: Static, dynamic, and economic analysis of in-browser cryptojacking,' in *2019 APWG Symposium on Electronic Crime Research (eCrime)*, IEEE, 2019, pp. 1–12. DOI: <https://doi.org/10.1109/eCrime47957.2019.9037576>.
- [287] P. Papadopoulos, P. Ilija and E. Markatos, 'Truth in web mining: Measuring the profitability and the imposed overheads of cryptojacking,' in *International Conference on Information Security*, Springer, 2019, pp. 277–296. DOI: [https://doi.org/10.1007/978-3-030-30215-3\\_14](https://doi.org/10.1007/978-3-030-30215-3_14).
- [288] H. Lee and K.-S. Choi, 'Interrelationship between bitcoin, ransomware, and terrorist activities: Criminal opportunity assessment via cyber-routine activities theoretical framework,' *Victims & Offenders*, vol. 16, no. 3, pp. 363–384, 2021. DOI: <https://doi.org/10.1080/15564886.2020.1835764>.

- [289] B. Kordy, S. Mauw and P. Schweitzer, ‘Quantitative questions on attack–defense trees,’ in *International Conference on Information Security and Cryptology*, Springer, 2012, pp. 49–64.
- [290] Z. Aslanyan, F. Nielson and D. Parker, ‘Quantitative verification and synthesis of attack–defence scenarios,’ in *2016 IEEE 29th Computer Security Foundations Symposium (CSF)*, IEEE, 2016, pp. 105–119. DOI: <https://doi.org/10.1109/CSF.2016.15>.
- [291] Z. Aslanyan and F. Nielson, ‘Pareto efficient solutions of attack–defence trees,’ in *International Conference on Principles of Security and Trust*, Springer, 2015, pp. 95–114.
- [292] Z. Aslanyan, F. Nielson and C. W. Probst, ‘Formal analysis of graphical security models,’ DTU Compute PHD-2016; No. 421, Ph.D. dissertation, Technical University of Denmark, 2017. [Online]. Available: [http://orbit.dtu.dk/en/publications/formal-analysis-of-graphical-security-models\(38c22eae-f164-4a45-a336-4722169facbb\).html](http://orbit.dtu.dk/en/publications/formal-analysis-of-graphical-security-models(38c22eae-f164-4a45-a336-4722169facbb).html).
- [293] R. R. Hansen, P. G. Jensen, K. G. Larsen, A. Legay and D. B. Poulsen, ‘Quantitative evaluation of attack defense trees using stochastic timed automata,’ in *International Workshop on Graphical Models for Security*, Springer, 2017, pp. 75–90. DOI: [https://doi.org/10.1007/978-3-319-74860-3\\_5](https://doi.org/10.1007/978-3-319-74860-3_5).
- [294] D. Vitkus, J. Salter, N. Goranin and D. Čeponis, ‘Method for attack tree data transformation and import into it risk analysis expert systems,’ *Applied Sciences*, vol. 10, no. 23, 2020, ISSN: 2076-3417. DOI: [10.3390/app10238423](https://doi.org/10.3390/app10238423). [Online]. Available: <https://www.mdpi.com/2076-3417/10/23/8423>.
- [295] A. Buldas, O. Gadyatskaya, A. Lenin, S. Mauw and R. Trujillo-Rasua, ‘Attribute evaluation on attack trees with incomplete information,’ *Computers & Security*, vol. 88, p. 101 630, 2020, ISSN: 0167-4048. DOI: <https://doi.org/10.1016/j.cose.2019.101630>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167404819301774>.
- [296] M. H. ter Beek, A. Legay, A. L. Lafuente and A. Vandin, ‘Quantitative security risk modeling and analysis with risqflan,’ *arXiv preprint arXiv:2101.08677*, 2021.
- [297] A. Walde and E. G. Hanus, ‘The feasibility of AIS- and GNSS-based attacks within the maritime industry,’ M.S. thesis, Norwegian University of Science and Technology (NTNU), Jun. 2020.

- [298] M. Vasek and T. Moore, 'Analyzing the Bitcoin Ponzi scheme ecosystem,' in *Fifth Workshop on Bitcoin and Blockchain Research*, ser. Lecture Notes in Computer Science, Springer, 2018. DOI: [https://doi.org/10.1007/978-3-662-58820-8\\_8](https://doi.org/10.1007/978-3-662-58820-8_8).
- [299] C. A. Tais, 'General analysis of the economy behind DDoS attacks,' *Hyperion International Journal of Econophysics & New Economy*, vol. 4, no. 2, 2011.
- [300] B. Stone-Gross, T. Holz, G. Stringhini and G. Vigna, 'The underground economy of spam: A botmaster's perspective of coordinating large-scale spam campaigns,' *LEET*, vol. 11, pp. 4–4, 2011.
- [301] A.-M. Jamil, L. b. Othmane and A. Valani, 'Threat modeling of cyber-physical systems in practice,' *arXiv preprint arXiv:2103.04226*, 2021.
- [302] Q. Xiang, A. Neufeld, G. W. Peters, I. Nevat and A. Datta, 'A bonus-malus framework for cyber risk insurance and optimal cybersecurity provisioning,' *arXiv preprint arXiv:2102.05568*, 2021.
- [303] L. D. Bodin, L. A. Gordon, M. P. Loeb and A. Wang, 'Cybersecurity insurance and risk-sharing,' *Journal of Accounting and Public Policy*, vol. 37, no. 6, pp. 527–544, 2018, Special Issue on Cybersecurity and Accounting, ISSN: 0278-4254. DOI: <https://doi.org/10.1016/j.jaccpubpol.2018.10.004>. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0278425418302382>.
- [304] A. Mazzoccoli and M. Naldi, 'Optimal investment in cyber-security under cyber insurance for a multi-branch firm,' *Risks*, vol. 9, no. 1, p. 24, 2021.
- [305] H. Bahşi, U. Franke and E. L. Friberg, 'The cyber-insurance market in Norway,' *Information & Computer Security*, vol. 28, no. 1, pp. 54–67, 2019. DOI: <https://doi.org/10.1108/ICS-01-2019-0012>.
- [306] D. Wrede, T. Stegen and J.-M. G. von der Schulenburg, 'Affirmative and silent cyber coverage in traditional insurance policies: Qualitative content analysis of selected insurance products from the German insurance market,' *The Geneva Papers on Risk and Insurance-Issues and Practice*, vol. 45, no. 4, pp. 657–689, 2020.
- [307] E. Marin, M. Almukaynizi and P. Shakarian, 'Inductive and deductive reasoning to assist in cyber-attack prediction,' in *2020 10th Annual Computing and Communication Workshop and Conference (CCWC)*, 2020, pp. 0262–0268. DOI: [10.1109/CCWC47524.2020.9031154](https://doi.org/10.1109/CCWC47524.2020.9031154).
- [308] A. Deb, K. Lerman and E. Ferrara, 'Predicting cyber-events by leveraging hacker sentiment,' *Information*, vol. 9, no. 11, 2018, ISSN: 2078-2489. DOI: [10.3390/info9110280](https://doi.org/10.3390/info9110280). [Online]. Available: <https://www.mdpi.com/2078-2489/9/11/280>.

- [309] CSO, *CSO's guide to the worst and most notable ransomware*, 2021. [Online]. Available: <https://www.csoonline.com/article/3607649/csos-guide-to-the-worst-and-most-notable-ransomware.html> (visited on 16/02/2021).
- [310] Intel471, *Ransomware-as-a-service: The pandemic within a pandemic*, 2020. [Online]. Available: <https://intel471.com/blog/ransomware-as-a-service-2020-ryuk-maze-revil-egregor-doppelpaymer/> (visited on 16/11/2020).
- [311] H. Lawrence, A. Hughes, R. Tonic and C. Zou, 'D-miner: A framework for mining, searching, visualizing, and alerting on darknet events,' in *2017 IEEE Conference on Communications and Network Security (CNS)*, IEEE, 2017, pp. 1–9. DOI: <https://doi.org/10.1109/CNS.2017.8228628>.
- [312] M. Ball, R. Broadhurst, A. Niven and H. Trivedi, 'Data capture and analysis of darknet markets,' *Available at SSRN 3344936*, 2019. DOI: <https://dx.doi.org/10.2139/ssrn.3344936>.
- [313] D. R. Hayes, F. Cappa and J. Cardon, 'A framework for more effective dark web marketplace investigations,' *Information*, vol. 9, no. 8, 2018, ISSN: 2078-2489. DOI: [10.3390/info9080186](https://doi.org/10.3390/info9080186). [Online]. Available: <https://www.mdpi.com/2078-2489/9/8/186>.
- [314] K. Soska and N. Christin, 'Measuring the longitudinal evolution of the online anonymous marketplace ecosystem,' in *24th {USENIX} security symposium ({USENIX} security 15)*, 2015, pp. 33–48.
- [315] E. Nunes, A. Diab, A. Gunn, E. Marin, V. Mishra, V. Paliath, J. Robertson, J. Shakarian, A. Thart and P. Shakarian, 'Darknet and deepnet mining for proactive cybersecurity threat intelligence,' in *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*, IEEE, 2016, pp. 7–12.
- [316] J. Robertson, A. Diab, E. Marin, E. Nunes, V. Paliath, J. Shakarian and P. Shakarian, 'Darknet mining and game theory for enhanced cyber threat intelligence,' *The Cyber Defense Review*, vol. 1, no. 2, pp. 95–122, 2016.
- [317] F. Dong, S. Yuan, H. Ou and L. Liu, 'New cyber threat discovery from darknet marketplaces,' in *2018 IEEE Conference on Big Data and Analytics (ICBDA)*, 2018, pp. 62–67. DOI: [10.1109/ICBDAA.2018.8629658](https://doi.org/10.1109/ICBDAA.2018.8629658).
- [318] PositiveTechnologies, *Positive technologies: High demand for hackers observed in 90 percent of ads related to hacking sites*, 2021. [Online]. Available: <https://www.ptsecurity.com/ww-en/about/news/positive-technologies-high-demand-for-hackers-observed-in-90-percent-of-ads-related-to-hacking-sites/> (visited on 16/02/2021).



- [319] E. Tekiner, A. Acar, A. S. Uluagac, E. Kirda and A. A. Selcuk, 'SoK: Cryptojacking Malware,' *arXiv preprint arXiv:2103.03851*, 2021. [Online]. Available: <https://arxiv.org/abs/2103.03851>.
- [320] F. P. Spagnoletti, F. Ceci and B. Bygstad, 'An investigation on the generative mechanisms of dark net markets,' in *Proceedings of the 13th Pre-ICIS Workshop on Information Security and Privacy*, 2018. [Online]. Available: <https://aisel.aisnet.org/wisp2018/20>.
- [321] E. Abeer, A. Laura, R. Leonid, D. Goldsmith, T. Alexander and B. Andrea, 'Collective dynamics of dark web marketplaces,' *Scientific Reports (Nature Publisher Group)*, vol. 10, no. 1, 2020. doi: <https://doi.org/10.1038/s41598-020-74416-y>.
- [322] B. Collier, R. Clayton, A. Hutchings and D. R. Thomas, 'Cybercrime is (often) boring: Maintaining the infrastructure of cybercrime economies,' in *The 2020 Workshop on the Economics of Information Security (WEIS)*, 2020. doi: <https://doi.org/10.17863/CAM.53769>.
- [323] A. J. Burstein, 'Conducting cybersecurity research legally and ethically,' *LEET*, vol. 8, pp. 1–8, 2008.
- [324] M. Christen, B. Gordijn and M. Loi, *The Ethics of Cybersecurity*. Springer Nature, 2020. doi: <https://doi.org/10.1007/978-3-030-29053-5>.
- [325] K. Macnish and J. van der Ham, 'Ethics in cybersecurity research and practice,' *Technology in Society*, vol. 63, p. 101 382, 2020. doi: <https://doi.org/10.1016/j.techsoc.2020.101382>.
- [326] D. Dittrich and E. Kenneally, 'The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research,' U.S. Department of Homeland Security, Tech. Rep., Aug. 2012. [Online]. Available: [https://www.dhs.gov/sites/default/files/publications/CSD-MenloPrinciplesCORE-20120803\\_1.pdf](https://www.dhs.gov/sites/default/files/publications/CSD-MenloPrinciplesCORE-20120803_1.pdf).
- [327] M. Zheng, H. Robbins, Z. Chai, P. Thapa and T. Moore, 'Cybersecurity research datasets: Taxonomy and empirical analysis,' in *11th {USENIX} Workshop on Cyber Security Experimentation and Test ({CSET} 18)*, USENIX Association, 2018. [Online]. Available: <https://www.usenix.org/conference/cset18/presentation/zheng>.
- [328] *The technical specification of VDES*, IALA Guideline G1139, Edition 3.0, Jun. 2019. [Online]. Available: <https://www.iala-aism.org/product/g1139-technical-specification-vdes/>.

- [329] CheckPoint, *Attacks targeting healthcare organizations spike globally as covid-19 cases rise again*. [Online]. Available: <https://blog.checkpoint.com/2021/01/05/attacks-targeting-healthcare-organizations-spike-globally-as-covid-19-cases-rise-again/> (visited on 01/03/2021).
- [330] R. Hunter, *Scammer interview*. [Online]. Available: [https://419eater.com/html/user\\_subs/interview/scammer\\_interview.htm](https://419eater.com/html/user_subs/interview/scammer_interview.htm) (visited on 01/03/2021).
- [331] P. H. Meland, *Retrospam: Nigeriabrev i posten*, Norwegian SINTEF infosec blog. [Online]. Available: <https://infosec.sintef.no/informasjonsikkerhet/2018/11/retrospam-nigeriabrev-i-posten/#more-5106> (visited on 01/03/2021).
- [332] EU, *Joint Framework on countering hybrid threats - a European Union response*, 2016. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016JC0018>.
- [333] S. Boes and E. R. Leukfeldt, 'Fighting cybercrime: A joint effort,' in *Cyber-Physical Security: Protecting Critical Infrastructure at the State and Local Level*, R. M. Clark and S. Hakim, Eds. Cham: Springer International Publishing, 2017, pp. 185–203, ISBN: 978-3-319-32824-9. DOI: [https://doi.org/10.1007/978-3-319-32824-9\\_9](https://doi.org/10.1007/978-3-319-32824-9_9).



## INDEX

The first page number is usually, but not always, the primary reference to the indexed topic.

### Symbols

419 fraud, 102

4D trajectory management, 25

### A

action research, 42

actuarial analysis, 33

ADTree, 59, 84, 93

advance fee scam, 102

agent, 33

ALE, 17

annual loss expectancy, 17

approach, 42

APT, 82

artefact creation, 49

asset, 9, 11

ATC, 24

attack (method), 10

attack cost, 42, 59, 60, 67, 86, 92

attacker cost, 3, 64, 69, 72, 86, 88, 92, 94

attacker investment, 86

attacker penalty, 86

attacker profit, 86

attacker supplier profit, 86

aviation, 24

### B

benchmarking, 52

Bonus-Malus, 95

botconomics, 94

bow-tie, 61, 64, 75, 84, 92, 93

BPMN, 15

broker, 34

business model, 22

### C

C-RAT, 91

campaign, 10

capstone, 8

case study, 52

CCA, 14

CCRM, viii

CFO, 34

CIDER, 88

CISO, 34

CLTC, viii

COBIT, 13

communication and consultation, 12

concurrent transformative, 48

confidence, 59

consequence, 10

context establishment, 11

continental model, 4

- control, 10
- CORAS, 14
- cost-benefit analysis, 17
- COVID-19, 97, 102
- CRediT, 55
- critical research, 42
- cryptocurrency, 67, 90, 91
- cryptojacking, 66, 90, 91, 98
- cryptomining, 99
- CTO, 34
- cyber, 9
- cyber insurance, 33, 60, 62, 63, 65, 94, 100
- cyber insurance profile, 63
- cyber kill chain, 93
- cyber physical system, 10
- cyber resilience, 10
- cyber resiliency, 10, 84
- cyber resource, 9
- cyber risk management, 103
- cyber security, 9
- cybercrime economy, 68, 78, 96
- cybercrime ecosystem, 90
- cyberspace, 9
- CySiMS, 30, 61, 75, 77
- CySiMS-SE, 30
- D**
- darknet, 67, 75, 89, 91, 96, 104
- DDos, 94
- defence cost, 59, 60
- defender cost, 3
- defender investment, 86
- defender loss, 86
- defender reactive cost, 86
- defender reimbursement, 86
- design science, 41
- DESMET, 50
- domineering, 105
- Dr Who, 9
- DSR, 41
- dynamic choice model, 19
- E**
- ECDIS, 88
- econometrics of wickedness, 17, 85
- economic motives, 22
- economic threat modeling, 17
- economics of cybercrime, 18
- ecosystem, 22
- engaged scholarship, 3, 41
- ENISA, 17
- ESA, 75
- ETA, 14
- ethical issues, 98
- evaluation, 50
- event, 11
- explanatory science, 41
- externalities, 16
- F**
- formal safety assessment, 28
- FSA, 28
- FTA, 14
- G**
- game theory, 19
- generic risk model, 63, 94, 104
- grounded theory, 42
- groupthink, 84, 102, 105
- GTM, 42
- H**
- hidden-action problem, 16
- historical data, 82

HTMA, 80

hybrid threats, 102

## I

IALA, 100

IDI, v

IDS, 23

IE, v

IEC, 10

IMO, 28

incentive, 85

information asymmetry, 16

information security, 9

InSecurance, 33

insurance ladder, 95

insurance policy, 33

insuree, 33

insurer, 33

interpretivism, 42

interview, 48

interview guide, 48

IRCM, 50

ISO, 10

ISSRM, 11

## K

KRI, 80

## L

likelihood, 10

literature study, 48

## M

Maersk, 29

Maginot Line, 1

maritime, 28

MDS, 77

method, 49

mining, 88

misaligned incentives, 16

MISP, 23

mixed methods, 45

monitoring and review, 12

moral hazard, 95

MP, 42

## N

Nash equilibrium, 20

netnography, 49

network externality, 16

NextGen, 25

Nigeria scam, 102

NIST, 9

NTNU, v

## O

OCTAVE, 13

open research data, 101

opportunity cost, 86

opportunity costs, 19

OSINT, 23

## P

paradigm, 41

perverse incentives, 16

PhD, v

phishing, 90

PKI, 30

positivism, 42

posters, 8

practice research, 44

practice-based research, 3, 41

pragmatism, 42

primary papers, 5, 56

probabilistic analysis, 80

probability, 10  
 problem investigation, 48  
 PSQ, 39  
 psychic benefits, 19  
 psychic costs, 19  
 PUCRS, vii

## Q

qualitative effects analysis, 52  
 qualitative experiment, 51  
 qualitative methods, 45  
 quantitative formal experiment, 51  
 quantitative methods, 45

## R

RaaS, 68, 90, 91, 96  
 ransomware, 68, 90, 97  
 RAT, 91  
 rational attacker, 18  
 rational attacker's paradigm, 18, 85  
 rational behaviour, 20  
 rational choice theory, 18  
 re-insurer, 33  
 regression models, 80  
 remote tower, 25, 76  
 research question, 37, 78, 79  
 resilience, 98  
 resource cost, 70  
 return on security investment, 17  
 risk, 10  
 risk analysis, 11  
 risk assessment, 11  
 risk avoidance, 12  
 risk evaluation, 11  
 risk identification, 11  
 risk management, 11  
 risk manager, 34

risk modification, 12  
 risk retention, 12, 84  
 risk sharing, 12  
 risk treatment, 11, 94  
 ROSI, 17

## S

SCADA, 16  
 scraping, 96  
 SDL, 13  
 secondary papers, 6, 73  
 security econometrics, 16  
 security economics, 12, 16  
 security indicator, 70  
 security requirement, 14  
 semi-structured interview, 48  
 SESAR, 25  
 SESS, v  
 SHIELDS, 59  
 socio-technical system, 10  
 software tool, 50  
 spam, 94  
 Spanish prisoner scam, 102  
 SQ, 37  
 STIP-INST, v  
 storyless, 3, 72, 82, 84, 94, 104  
 subjective logic, 80

## T

test environment, 50  
 threat, 9, 10  
 threat agent, 11  
 threat analysis, 69  
 threat intelligence, 22, 104  
 threat modeling, 13  
 Threat Modeling Manifesto, 13, 104

threat modelling, 12, 13, 22, 82, 91, 102,  
103

TTI, 23

**U**

UFBA, vii

UML, 15

underwriter, 33

utility, 17, 84

utility function, 17

**V**

VDES, 30

VTS, 30

vulnerability, 9, 10





APPENDIX



## PRIMARY PAPERS



**A: ‘Attribute decoration of attack–defense trees’**

A written permission to include this material in its published form [17] has been obtained from IGI Global.

A

# Attribute Decoration of Attack–Defense Trees

*Alessandra Bagnato, TXT e-solutions, Italy*

*Barbara Kordy, University of Luxembourg, Luxembourg*

*Per Håkon Meland, SINTEF ICT, Norway*

*Patrick Schweitzer, University of Luxembourg, Luxembourg*

---

## ABSTRACT

*Attack–defense trees can be used as part of threat and risk analysis for system development and maintenance. They are an extension of attack trees with defense measures. Moreover, tree nodes can be decorated with attributes, such as probability, impact, and penalty, to increase the expressiveness of the model. Attribute values are typically assigned based on cognitive estimations and historically recorded events. This paper presents a practical case study with attack–defense trees. First, the authors create an attack–defense tree for an RFID-based goods management system for a warehouse. Then, they explore how to use a rich set of attributes for attack and defense nodes and assign and aggregate values to obtain condensed information, such as performance indicators or other key security figures. The authors discuss different modeling choices and tradeoffs. The case study led them to define concrete guidelines that can be used by software developers, security analysts, and system owners when performing similar assessments.*

*Keywords:* Attack Trees, Attack-Defense Trees, Attributes, Calculation, Guidelines, Security Analysis, Security Modeling

---

## 1. INTRODUCTION

The security of any sufficiently valuable system is not static. To keep a system secure, it has to be protected against an increasing number of threats of growing complexity. As defenses are added to the system, more sophisticated attacks break these defensive measures anew. To cope with the resulting, intricate systems, a formal modeling and evaluation approach become indispensable.

One of the formal approaches to assess a system's security is the *attack–defense tree* (ADTree) methodology. ADTrees focus on the interaction between two types of players, attackers and defenders, while keeping the complexity of the formalism at a minimum (Kordy et al., 2011b). They are a compromise between attack trees, which are too restrictive in their modeling capabilities, and petri-nets, where modeling is quite intricate and computationally complex. ADTrees retain the easily understandable tree structure and are therefore especially useful in an interdisciplinary work environment, where

DOI: 10.4018/jsse.2012040101

an intuitive understanding of the system is as important as formal foundations. ADTrees even allow a rough first assessment of a system's security purely based on the visual representation of the scenario, making it easy to spot missing or redundant defenses. The theoretical aspects of the ADTree methodology have already been extensively studied by Kordy et al. (2010, 2011a, 2011b).

The purpose of this paper is to present experiences and provide practical recommendations on the use of attributes in ADTrees. Attributes are the part of the ADTree formalism that allows quantitative analysis, something that is of great value for risk analysis either during planning, development or maintenance of a system. There are numerous security attributes to be found in the literature today, and through a case study we show how a selection of them can be applied, how values are assigned to nodes and how they are used for quantitative analysis. Knowing which attributes to choose and how to estimate their values is a non-trivial challenge and is addressed in detail. Attributes are used to answer questions such as: Is it possible to attack the system? How much would it cost to prevent one or all attacks or implement one or all defenses? How long does it take to secure the entire system? We are interested in extending these answerable questions to bivariate questions, i.e., questions where inputs from attackers and defenders are needed. This, for example, includes questions such as: Given a limited defense budget, can the defender at least defend against some attacks? How does the scenario change in case of a power outage?

The case study was based on an operational *Radio-Frequency Identification* (RFID) system for goods management in a warehouse, taking technical, physical and social engineering aspects into account. There were four players from both academia and industry involved, taking roles as defenders and attackers.

The rest of the paper is structured as follows. This section continues with a summary of the theoretical foundations of ADTrees and concludes with a short literature review on related work. In Section 2, we review some of

the attributes that can be found in the literature and elaborate on different calculation methods. In Section 3, we present the case study scenario and the corresponding ADTree. Section 4 shows the attribute decoration and calculation of values for the ADTree. The results of the case study are discussed in Section 5 and we conclude and synthesize our recommendations in Section 6.

### 1.1. ADTrees

ADTrees were introduced by Kordy et al. (2011a) and are an extension of attack trees (Schneier, 1999; Mauw & Oostdijk, 2005) with defense nodes. An ADTree is a node-labeled rooted tree describing the measures an attacker might take in order to attack a system and the defenses a defender can employ to protect the system. ADTrees allow the system modeler to repeatedly interleave attack and defense components. Consequently, an ADTree has nodes of two opposite types, *attack* nodes and *defense* nodes, and can be seen as a game between two players: an attacker and a defender. The *root* node of an ADTree represents the *main goal of the attacker*. Every node of an ADTree may have one or more *children* of the same type representing a *refinement* of the node's goal into sub-goals. The refinement of a node is either *disjunctive* or *conjunctive*. The goal of a disjunctively refined node is achieved when *at least one* of its refining children's goals is achieved. The goal of a conjunctively refined node is achieved when *all* of its refining children's goals are achieved. Nodes which do not have any children of the same type represent *basic actions*. Every node may also have one child of the opposite type, which represents a *countermeasure*. Thus, an attack node may have several children which refine the attack and one child which defends against the attack. A defense node in turn may have several children which refine the defense and one child, being an attack node, which counters the defense. We understand countermeasures in a general sense, i.e., countermeasure ranges from a complete prevention of the parent's goal over a possible prevention to a weak mitigation.

The semantics of an ADTree is the following. The attack tree, obtained from an ADTree by removing all subtrees rooted in defense nodes, represents how an attacker can attack a given system taking all existing defensive measures into account. The remaining parts of an ADTree, i.e., all defense nodes, their refinements, counterattacks, and so on, represent possible measures that can be put in place in order to defend against the original attack on the system, or attack the newly introduced defenses, and so on. This means that the existing defensive mechanisms are *not* explicitly depicted in an ADTree, whereas the original attack tree already represents how they should be overcome.

We depict attack nodes by circles and defense nodes by rectangles. Refinement relations are indicated by solid edges between nodes, and a countermeasure is connected to the countered node using dotted edge. In addition, a conjunctive refinement of a node is depicted by an arc which connects the node's edges to its children of the same type.

We illustrate the attack–defense language by giving an ADTree, depicted Figure 1. The root of the tree depicts the main goal of the attacker. It is disjunctively refined into two subgoals. The Disjunctive Subgoal 1 is countered by a countermeasure, whereas the Disjunctive Subgoal 2 is conjunctively refined into two subgoals. A more extensive ADTree, where the main goal is a *Denial of Service* (DoS) attack on an RFID-based management system, is detailed in Section 3.1 and is depicted in Figures 4 through 7.

## 1.2. Related Work

The literature on attack trees is abundant. Piètre-Cambacédès and Bouissou (2010) have given a historical overview on graphical representations of computer attacks, such as fault trees (Vesely et al., 1981), threat trees (Amoroso, 1994) and privilege graphs (Dacier & Deswarte, 1994), and how these representations led Schneier to coin the term attack tree (Schneier, 1999).

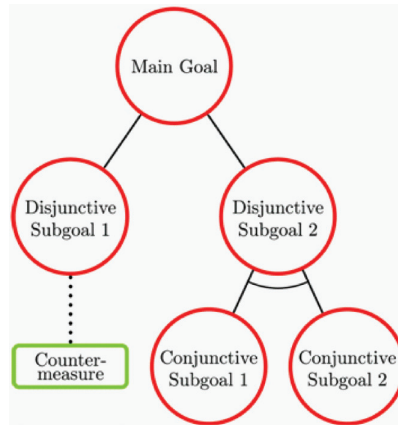
Many authors treat how to add different kinds of values to attack tree nodes. In Schneier's terms these values are called attributes. In 1999, he proposed how to analyze the costs and the success probability of an attack with the help of attack trees. Since then, many authors have followed in his footsteps proposing extensions to attack trees and attributes, as well as describing case studies. For instance, Amoroso (1994), Mauw and Oostdijk (2005), Buldas et al. (2006), Li et al. (2009), and Tanu and Arreympi (2010) demonstrate how an attack tree can be parameterized with different kinds of values and how to deduce aggregated results. Moore et al. (2001) include attacks related to social engineering and physical entering of premises and Saini et al. (2008) show examples from multiple other systems. Baca and Petersen (2010) have extended attack trees with countermeasure graphs with an example from open-source application development, while Edge et al. (2006), Bistarelli et al. (2007), and Roy et al. (2011) have extended attack trees with a notion of defense nodes for the leaves of the trees.

There also exist a number of deeper studies and experience reports with attack tree based methods applied to real-life systems. Some notable examples are Henniger et al. (2009), who have conducted a study using attack trees and a variety of node attributes for vehicle communications systems, Abdulla et al. (2010) with an analysis on the GSM radio network, and Tanu and Arreympi (2010) using vulnerability tree, fault tree and attack tree analysis on a mobile SCADA system for a multiple tank and pump facility. Byres et al. (2004) treat another SCADA case study related to a MODBUS protocol for critical infrastructures. All these work show that attack tree based methodologies constitute a very useful tool for modeling threats and analyzing vulnerabilities of complex systems.

Finally, steps have been made to compare and combine attack tree based models with other modeling techniques in order to obtain a better and more complete way for representation and analysis of threats and vulnerabilities. For instance, Opdahl and Sindre (2009) compare



Figure 1. Generic ADTree



misuse cases with attack trees. They use a class room experiment to gather experimental data and determine which method models the scenario more accurately. According to this study, using attack trees for finding threats is more effective than the use of misuse cases. Moreover, the authors conclude that the perception of a used technique is not correlated with the actual performance of that technique. A

similar approach has been proposed by Diallo et al. (2006), where a comparative evaluation of the common criteria, misuse cases and attack trees is presented. Based on the results obtained in these works, Tøndel et al. (2010) suggest combining misuse cases and attack trees, in order to represent possible threats, attacks and mitigating countermeasures. This suggestion is made with reuse of models in mind, and is

Figure 2. WIMS deployment diagram

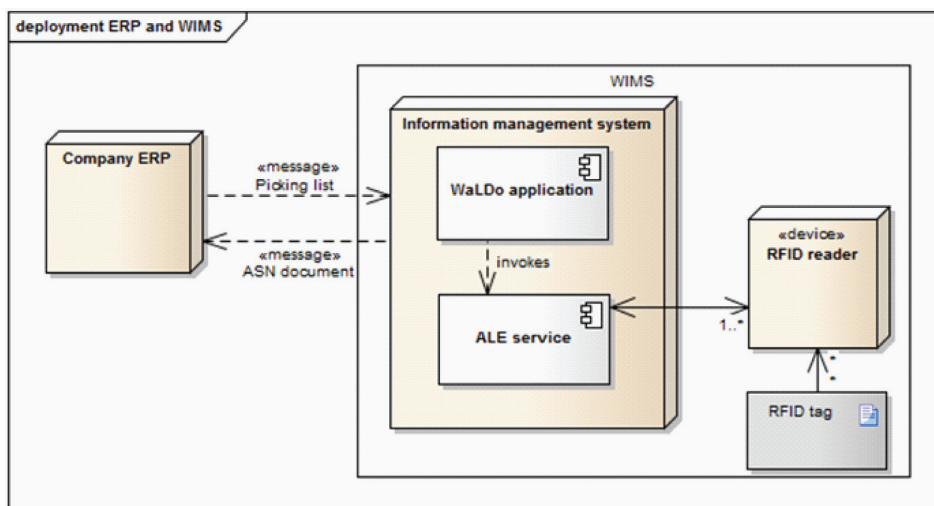
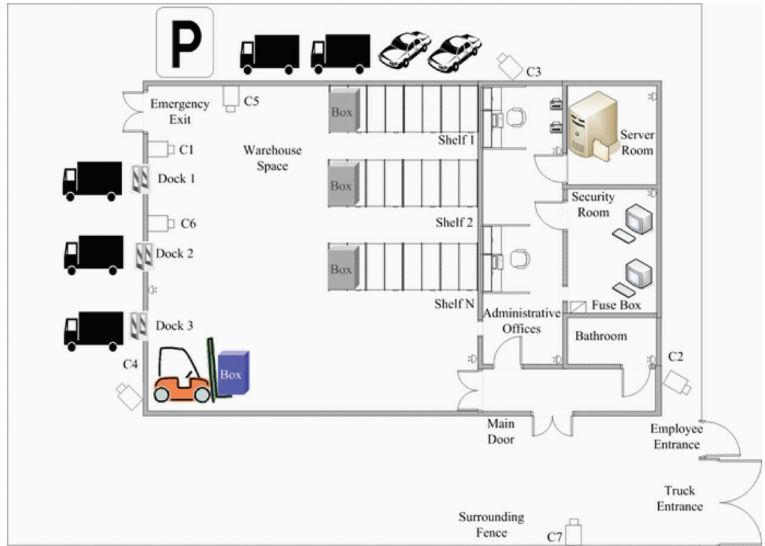


Figure 3. Floor plan



supported by an online repository of security models developed within the SHIELDS project (2008-2010).

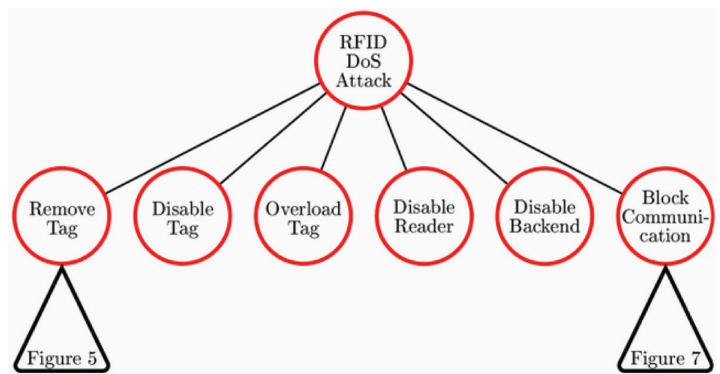
The results and experiences from the related work have been taken into account when designing the formal treatment of attributes for ADTrees. As ADTrees are the only approach that is formal and allows for interleaved attacks and defenses, the ideas needed to either

be formalized or extended to include a notion of interleaving.

## 2. BACKGROUND OF ATTRIBUTES

In order to analyze an attack–defense scenario, we use quantitative measures called metrics or attributes. Some attributes, e.g., **costs**, **probability**, appear frequently in the context of

Figure 4. An ADTree for RFID DoS attack



attack trees, others can rarely be found. In this section, we provide background information about attributes and related calculation methods. For readability, we always write attribute names in bold font.

## 2.1. Attribute Overview

### 2.1.1. Attributes for Attack Trees

Schneier (1999) proposes decorating leaf nodes of attack trees with the values expressing the component's **costs** in order to deduce the cheapest attack. A similar approach applies to metrics such as **probability**, **severity**, **impact**, and **consequence**. According to Schneier, all of these attributes can be combined or conditioned, in order to identify the cheapest low-risk attack, attacks that cost less than a given amount of money, or cheap, highly successful attacks where only medium skill is needed, etc.

Other researchers extend this catalog of attributes with further metrics. Edge et al. (2006) introduce **protection costs**; Byres et al. (2004) use **detectability** which illustrates how easily a defender can discover an attack and **difficulty** which specifies the needed skill of the attacker. As suggested by Henniger et al. (2009), **attack time** - which models an amount of time needed to perform an attack - may be considered independent of an attacker's **skill**, **costs** or **resources**. Attacks that a defender cannot really protect himself against can be annotated as **unmitigatable**. Instead of modeling constraints as additional child nodes of a conjunctive node, we can use attributes, such as **requires an insider** or **needs electricity**. These and similar attributes merely depict a choice between two options and can therefore be modeled with Boolean values, as suggested by Schneier (1999). We subsume Boolean attributes under the keyword **special skill**. A formal treatment of attributes for attack trees, that guarantees compatibility with underlying semantics, was first described by Mauw and Oostdijk in (2005).

### 2.1.2. Attributes for Defense Nodes

Traditionally, attack trees only consider attributes directly related to attacks or attackers. With ADTrees we can also cover attributes that quantify defenses and their behavior. Many attributes related to attacks can straightforwardly be extended to encompass defenses. For instance, **costs** can refer to **costs of performing an attack** or **costs of performing a defense**. Similarly **probability of success**, **probability of occurrence**, **required skill level**, **number of possible countermeasures** and many more can be adapted. Other attributes might only make sense for either attacker or defenders, such as **penalty** which a law-abiding defender would never have to worry about, or **response time** where a defender might be depended on the **response time of a third party**. The usefulness of distinguishing between attributes for attack and defense nodes becomes apparent, once we look at questions related to attributes such as **social costs** (addition of attacker's as well as defender's costs) or attributes where values for one player have a direct consequence on values for the other player, as in the case of **satisfiability** or **probability**. In addition, defense nodes and their associated attribute values allow us to answer bivariate questions, such as how much does a defender have to spend on defenses, if he knows that the attacker has a low **skill level**.

### 2.1.3. Meta-Attributes

Attributes themselves are the main ingredient to perform a quantitative analysis with the help of ADTrees. However, when associating attribute values with nodes, we might still want to distinguish between the associated values, even if the values themselves are the same. For example, an RFID security expert would associate a medium probability of occurrence to a DoS attack and is very confident about that, whereas a person working in computer forensics would also associate a medium probability to the node, but would probably be less confident about it. This additional information, describ-

ing an attribute value we call a *meta-attribute*. An example is **confidence**, which indicates how certain a decorator is, when associating an attribute value with a given node. Another meta-attribute is **coverage**, which expresses the number of people who have associated an attribute value with a given node. Meta-attributes are suitable to be used in combination with any attribute. Using meta-attributes allows us to model the desired properties more accurately, and to improve their quantification. Meta-attributes constitute one of the novelties of the methodology used in the current case study, as they have not yet been mentioned neither in the context of attack trees nor attack–defense trees in the existing literature.

#### 2.1.4. Value Domains

Attribute values can range over diversified types of mathematical domains. One can consider Boolean values, values from a nominal scale, e.g., Low, Medium, and High, real numbers or even discrete or continuous probability distributions. If data is available, the probability distribution could be estimated from histograms. Meta-attributes can also be quantified using the values from any of possible domains. For example, the pair (4.23, 2) could be a possible value for **costs of attack** with the meta-attribute **confidence**, where the domain is modeled as the product space of the real numbers and a nominal scale from 1 to 5. Instead of a single value, it is also possible to associate sets of values to the nodes. For example, if we know that the attack costs are not High, we could associate the set {Low, Medium} to the corresponding node.

### 2.2. General Calculation with Attributes

Attributes provide a powerful analysis tool for vulnerability scenarios. They help us estimate which attacks may happen with a high probability and which countermeasures should be applied. However, to get useful insights from the analysis, it is necessary to have accurate values associated with all the nodes of an ADTree. One possibility is to ask experts to provide the

values. Another strategy is to involve several people, such as the system owner, developers and administrators, to perform the task. In any case, this process can be very time consuming, costly and highly error-prone, depending on the tree complexity and the number of attributes. Thus, numerous approaches have been proposed, allowing us to deduce values for one node, based on values already associated with other nodes, or to combine values for several attributes in order to deduce the value for another attribute. In this section, we give a brief overview of the calculation methods already present in the literature.

As pointed out by Schneier (1999) in the case of attack trees, the most intuitive calculation procedure on attack trees is the bottom-up approach. The idea is to only associate attribute values with the basic actions and then deduce the values corresponding to the refined nodes from the values associated with their children. The functions which are used to calculate the value for a parent node depend on the type of the corresponding refinement. The bottom-up approach presents several advantages. First of all, we only have to decide on values for a small amount of nodes, which reduces the time necessary for the attribution of values. Moreover, estimation of values for basic actions should be feasible in most cases, since such actions can be easily understood and quantified (if this is not the case, a node should be refined further). Finally, this approach is suitable for evaluation of a large number of attributes and it can be automated. A formal framework for the bottom-up approach for attribute evaluation on attack trees has been developed by Mauw and Oostdijk (2005). This approach has been extended to ADTrees by Kordy et al. (2011a), where Example 5 illustrates the calculation procedure.

It is also possible to define a new attribute by combining several existing attributes. Such combinations allow us to estimate values corresponding to more complex properties, for which it would be difficult to provide values directly. As an example, Edge et al. (2006) have defined a variant of a **risk** attribute for attack

trees based on the formula **risk = (probability/costs)\*impact**. A similar approach has been used by Jürgenson and Willemson (2008), where the **costs, success probability, gain and penalty** attributes have been combined in order to define a new attribute called the **exact expected outcome** of the attacker. Henniger et al. (2009) combine the attributes **elapsed time, expertise, knowledge of system, window of opportunity** and **required equipment**, in order to deduce the **required attack potential**. Finally, Fung et al. (2005) show how the **difficulty level** associated to the non-refined nodes can be used to estimate the **survivability** in the root node.

It is also important to observe how multiple attributes relate to each other and how the values for one attribute may influence the values for another one. For instance, the **costs** associated with a given attack component can be used to estimate the corresponding **probability of occurrence** value, i.e., to deduce how probable it is that an action will take place. This is of particular importance, when we have some specific knowledge about the attacker. For instance, if we know that the attacker has a limited budget we can deduce that, with a high probability, he will not perform actions which are more expensive than a certain threshold. Moreover, attribute values can be used to check soundness of a scenario. Rational reasoning lets us deduce that attack components, which require a higher investment than the potential gain of the attacker, do not have to be considered, because in such case the attack would be unprofitable for the attacker. Similarly, as pointed out by Herley (2009), the **protection costs** of a threat should not be greater than the **benefit** gained by following a security advice. Indeed, such a protection would do more harm than good, from an economic point of view. Furthermore, if the **protection costs** are greater than the **impact** of an attack, economically speaking the protection causes more harm than the attack it addresses.

Moreover, in the case of ADTrees, the values associated with the components of one player may influence the values for the other player. For instance, if the **satisfiability** value of an action for the attacker is True, then it

follows that the **satisfiability** of this action for the defender is False, and vice versa. A similar property holds for **probability**: if the attacker is successful at a node with probability  $p$ , then the defender is successful at this node with probability  $1-p$ . However, such a relation does not exist for all attributes. For instance, when considered in isolation, the **attacker's costs** have no impact on the **defender's costs**.

Finally, properties quantified with Boolean values are well suited to model hypothetical scenarios. As an example, let us consider the attribute **electricity needed**, which evaluates to True at every component which requires electricity, e.g., cameras, and to False otherwise. Using this attribute, we can model what happens if we experience a power outage, by projecting the scenario on its part evaluated to False. This would allow us to check whether the existing defensive measures would also be sufficient in emergency situations.

### 3. AN RFID-BASED GOODS MANAGEMENT SYSTEM

In this section we describe the setting of RFID-based goods management system and the creation of a corresponding ADTree model. The system had already been subject of a threat assessment as a part of the EU-funded project SHIELDS (2008-2010). Therefore, we already had attack trees and misuse case diagrams describing potential threats and countermeasures as an initial starting point. In order to extend this work and capture threats related to technical, physical and social engineering, we did a more thorough analysis of the conceptual design of the RFID-based goods management system, the physical layout of the warehouse where it is deployed and how people are involved in the work processes. We also made use of other relevant attack trees, such as Mirowski et al. (2009), to supplement the creation.

In this case study we have chosen to provide detailed examples from the physical and social engineering world and illustrate the ADTree formalism on them. Obviously, this is not a

restriction on the formalism, and the presented methodology can be readily applied in other fields, such as risk management or software engineering. We believe that it is easier to relate to these tangible examples than to fields that require expert knowledge to even understand the depicted scenario. In the following sections we draw up general purpose guidelines that can be easily modified to the field of application by adapting the set of suitable attributes as well as possibly changing their value domain.

### 3.1. Setting

In order to perform a realistic case study, we selected an already deployed and operational system named the *Warehouse Information Management System (WIMS)* with special focus on one of its components, the *Warehouse Loading Docks Management Application (WaLDo)*. This system manages all incoming and outgoing goods to and from a warehouse, keeping track of orders, goods location, picking lists, shipping notifications, etc. The warehouse is a highly automated environment where all goods can be electronically identified using RFID tags. Figure 2 gives a high level overview of the system itself.

The WaLDo application controls all goods that cross the loading docks of the warehouse. The physical warehouse is equipped with RFID enabled loading docks. All RFID readers conform to the EPCGlobal specifications and are managed via an *Application Level Event (ALE)* service that provides a web service interface to upper layer applications like WaLDo. Additionally, the warehouse has an information management system able to interact with the company *Enterprise Resource Planning (ERP)* system, the integrated software application that manages the entire company information flow and resources, to process universal business language documents like *Picking Lists*, used to specify which material is to be shipped to whom, and *Advanced Shipping Notification (ASN)* documents, used to specify which goods are expected to be received.

In order to properly analyze potential threats, we also consider the environment in which the system operates. Figure 3 depicts the physical premises, the equipment and the workspaces inside the warehouse. The WaLDo application operates in a warehouse where eight employees are working. The size of the warehouse building is 500 m. It contains RFID enabled forklifts, shelves for goods and three loading docks with RFID readers, which can only be opened from the inside. All goods pass in and out through the loading docks and are registered by the RFID readers. The building also has one room for computer servers, one administrative office, one security room containing two *Closed-Circuit Television (CCTV)* monitors and a fuse box, one bathroom, one corridor, a main entrance and an emergency exit. The warehouse is surrounded by a fence that encloses the entire area. The fence has two gates, one for trucks and one for employees, which can only be opened remotely from the security room. The area inside the fence has a parking place where trucks can wait before unloading their goods and where the employees can park their cars. The warehouse is equipped with a high-speed Internet connection and a wired LAN Ethernet. The Ethernet network connects the servers with the RFID readers of the loading dock. In total, there are seven surveillance cameras that are linked to external security services, monitoring both the inside and the outside of the warehouse. Cameras 1, 5 and 6 monitor the shelves within the WaLDo building, Camera 2 monitors the main entrance gates, Camera 3 monitors the parking areas, Camera 4 monitors the loading docks and Camera 7 monitors the warehouse's main door. Each day between 10 and 20 trucks deliver goods to or from the warehouse. The drivers load the goods on and of their trucks by accessing the warehouse through the docks. Though we are not specifying exactly what kind of goods is stored in the warehouse, we assume they are worth stealing (otherwise the security assessment would be pointless).

### 3.2. The ADTree Model

The information given in the previous section served as a basis to create an ADTree that we then used for attribute decoration. First, an initial tree was suggested by one player. This tree was then independently examined by the other players who suggested improvements which were merged with the tree through several iterations. We limited the scope by focusing on one high-level attack-defense scenario, namely disrupting the RFID-based part of the system by preventing communication between a specific tag and a specific reader. Each player spent roughly three hours on the tree creation phase, later only minor refinements were necessary. We did not have an automated method of combining trees; however the trees were small enough to do a visual comparison in order to reveal missing or similar nodes.

The top goal node of the high level tree model, shown in Figure 4, is called “RFID DoS Attack.” In order to achieve this goal, an attacker has six options. He can “remove the tag,” “disable the tag,” “overload the tag,” “disable the reader,” “disable the backend” or “block the communication” between all tags and all readers. Even though we initially refined all children, we chose to continue the case study by only refining the nodes “remove tag” and “block communication.” The refinements are depicted in Figure 5 and Figure 7. We deliberately chose to analyze an incomplete tree to reflect that, in most use cases, the modeling time is limited which invariably will lead to incomplete trees.

To physically remove the RFID tag an attacker can either remove the tag himself, or he can convince someone else to remove the tag. In the first case, he either can “infiltrate the building” or he has to “infiltrate the organization” and thereby gain legitimate access. Infiltrating the building can be achieved by “breaking and entering,” as detailed in Figure 6, by “posing as a truck driver,” by executing a “postal Trojan attack” or by staging a “visitor attack.” A postal Trojan attack can be achieved when the attacker “hides in a box” and this box is sent to the warehouse. The owner of the

warehouse could defend against Trojan mail by employing a “sniffer dog” that can detect humans in the incoming goods. The attacker, in turn, could confuse the dog using decoy rats or pepper spray. If the attacker decides to execute a “visitor attack” he can “come as visitor” during daytime and “hide in the bathroom” until everyone else has gone home. The defender could anticipate such an attack and “track visitors” on the warehouse premises. Tracking the visitors can be accomplished by “escorting the visitors,” by requiring visitors to “register in a visitor’s log,” by using a more supervised attended visitor’s log, or by installing “presence detectors on the premises.” A visitor could choose to overcome the defense “register in a visitor log” by “faking a log entry”. In that case, the warehouse owner should switch to “register in an attended visitor’s log”. If the attacker decides that he wants to infiltrate the organization, he can try to “get hired as warehouse staff,” “pose as warehouse employee,” or simply “buy the warehouse” (we deliberately added some extreme nodes to the tree to try and provoke some extreme attribute values). The defender could protect himself against any infiltration by performing “background checks” on everyone he works with.

If the attacker chooses to convince someone else to remove the tag, he can “bribe,” “threaten,” “blackmail,” or “trick” this person. In the first case, he has to “identify a corruptible subject” and then he has to actually “bribe the subject.” The warehouse owner could defend against bribery by “thwarting the employees” from receiving bribes, by providing mandatory “security awareness trainings” or “threatening to fire the employees” in case of infringement. Provided the attacker wants to trick another person into removing the tag, he can either “send false replacement tags” or he can place a “false management order” to replace the tags. Fake orders can be done by “infiltrating the management” and “ordering replacement tags.” A defender can prevent this kind of trickery by mandatory “security awareness training courses.” Last, a defender could prevent any kind of removal of the RFID tag by using a

“stronger adhesive,” i.e., attaching the tag in a way that it cannot be removed.

If an attacker decides to remove the tag himself by breaking and entering he must “get onto the premises” and “get into the warehouse,” undetected by the installed cameras. To get onto the premises, an attacker can “climb over the fence” or he can “enter through the gate” for employees undetected by Camera 2. To prevent attackers from climbing over the fence, the defender could install “barbed wire” on the fence. An attacker, in turn, could circumvent the barbs by guarding against them, which he could achieve by either throwing a “carpet over the barbs” or by “wearing protective clothing.” The attacker also has to get into the warehouse. He can accomplish that by “entering through the door” undetected by Camera 7 or “entering through the loading dock” undetected by Camera 4. The defender could prevent an attacker from entering through the main door by monitoring the door with biometric sensors. Another defensive measure would be to install and monitor the premises with additional security cameras. These new cameras would monitor the parts of the property not yet covered, but could be rendered useless if an attacker disables them. Disabling could be done by shooting a strong laser at the cameras or by “video looping the camera” feed. Alternatively, guards patrolling the premises could protect against this kind of threat.

Blocking communication can be done by blocking the communication between the tag and the reader or by blocking communication between the reader and the backend, as depicted in Figure 7. To do the former, there exist several options: It is possible to “shield the tag,” to use a malicious reader that constantly requests information from the tag and this way “blocks the tag,” to use a different tag that “blocks the reader,” or to “jam all signals.” Shielding a tag can be achieved by “being in the vicinity of the tag” and by using a “Faraday cage.” An obvious defense against attackers being in the vicinity of the tags would be to increase the “security of

the warehouse”. A Faraday cage can be installed around the reader or around the tag. To prevent attackers from jamming the signal, the defender could “isolate the entire warehouse network,” which could be achieved by “securing the warehouse” or “encasing the entire warehouse inside a Faraday cage.” If the attacker decides to block the communication between the reader and the backend, he can achieve it by evoking “DoS in the wired network.”

## 4. ATTRIBUTE CASE STUDY

Having established our ADTree, we are ready to focus on the novelty of the case study: the decoration of the ADTree with attributes and the corresponding values. Figure 8 gives an overview on how this case study was performed. As explained in Section 3, the *ADTree creation* activity was done through several iterations. This was followed by an *attribute decoration* activity, which consisted in selecting a relevant set of attributes, choosing players and estimating attribute values. Section 4.1 explains this part in more detail, while Section 4.2 describes the objective observations we recorded. The next step was to prepare the attribute values for calculation, as explained in Section 4.3. Finally, the bottom-up algorithm was applied in order to derive the minimal **costs** of the attacker in our scenario. This step, as well as the corresponding findings, is presented in Section 4.4.

### 4.1. ADTree Decoration with Attributes

In this case study, we have experimented with a game-based approach for the decoration of the ADTree. We chose a set of attributes that we felt were useful and possible to provide accurate values for (in other studies a different set might be more suitable). The selected attributes with their detailed descriptions, example references and the corresponding value domains can be found in Table 1. In many cases there are variants of the attribute names in the



Figure 5. The remove tag subtree

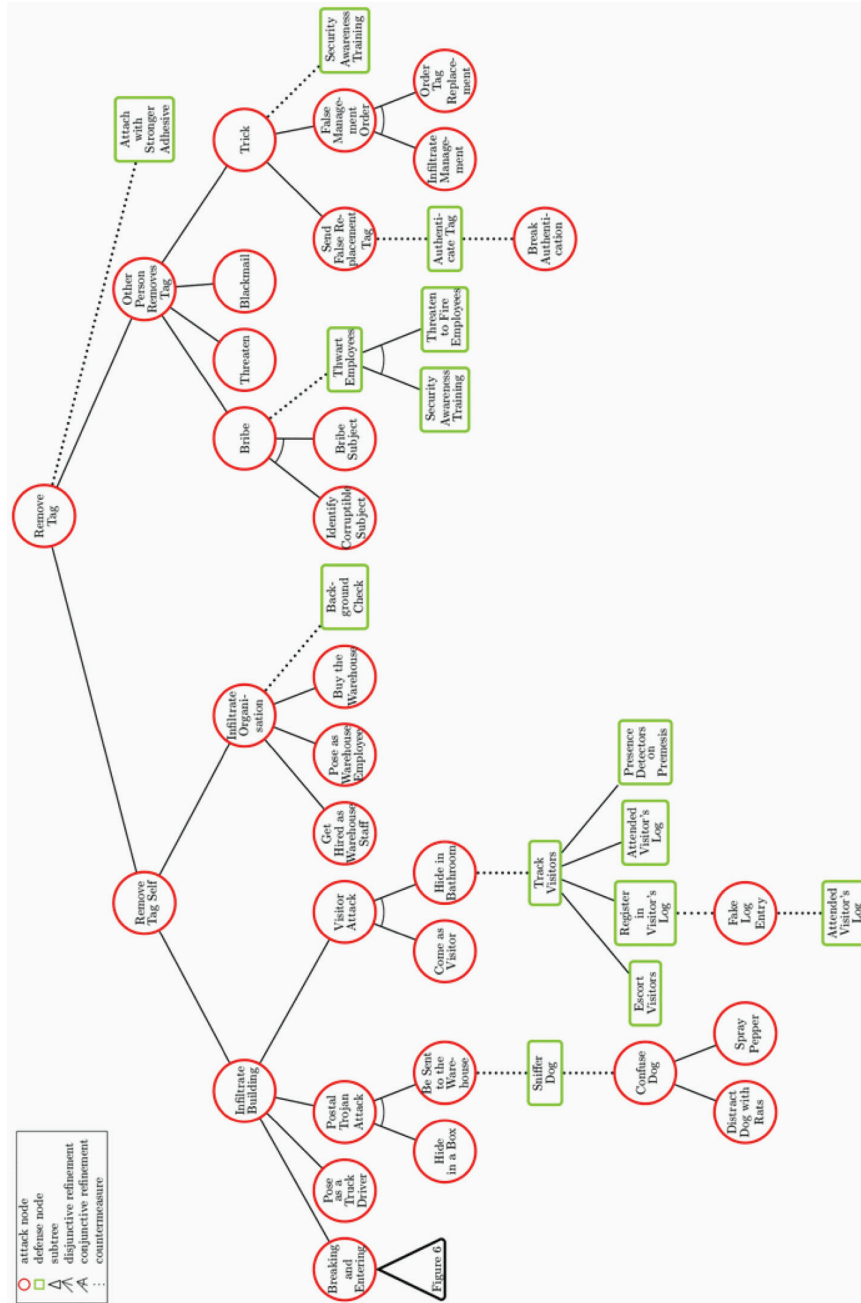
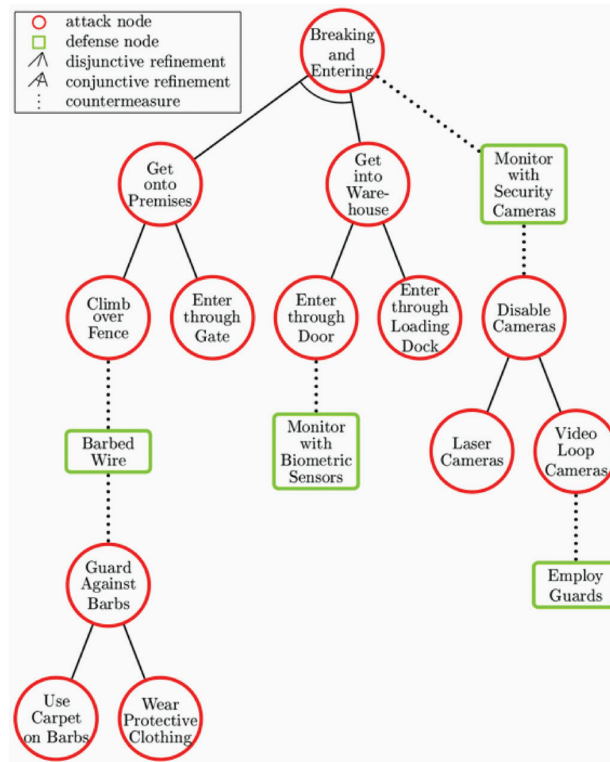


Figure 6. The breaking and entering subtree



literature, e.g., we have used the term **impact** to cover **impact, severity, consequence, damage, criticality**, as well as **seriousness**. Most of the chosen attributes can be applied to both attack and defense nodes. We also decided to make use of the meta-attribute **confidence** with the domain  $\{1, \dots, 5\}$ , where 1 represents total lack of confidence and 5 very high confidence.

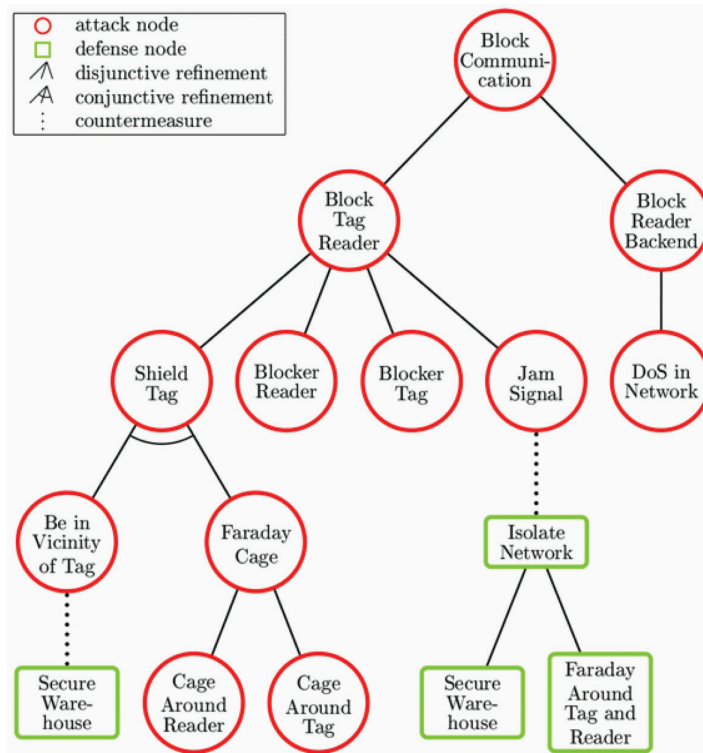
There were two players taking roles as attackers, and two as defender. The players estimated the attribute values for the nodes of the ADTree. For this purpose we created an empty table with 9 columns for the attributes and 79 rows for all nodes. Our intention for the table was to prevent illegible values that would have occurred if the players had estimated all values directly on the ADTree printed on a sheet of paper. All players were allowed to use the labeled ADTree, the warehouse and the system

description given in Section 3.1, the attribute description from Table 1 and the empty table. The values were estimated independently over a period of one week, but we did not set a specific time limit. Neither did we require estimated values for every node and attribute, we rather suggested to apply a best effort strategy.

#### 4.2 Observations from the Attribute Decoration

The entire ADTree consisted of 79 nodes, where 59 were attack nodes and 20 were defense nodes. An extract of the resulting estimated values is shown in Table 2, where each line for an attack node represents a player who had been given the role of an attacker; similarly for lines representing defense nodes. A dash indicates a conscious decision not to estimate a value; an

Figure 7. The block communication subtree



empty field indicates that the player was not considering estimating a value. The letters [D] and [B], inserted in front of the node labels, indicate that the node was a defensive node and a basic action, respectively.

None of the players used a real number values for any of the attributes, instead all four players used the ranged values given in Table 1. In Table 2, the first letter of these ranges has been used to indicate the value, and the following number gives the confidence value.

The two players who were initially given the role of attackers spent approximately 90 min and 120 min, the players with the role of the defender spent 40 min and 90 min.

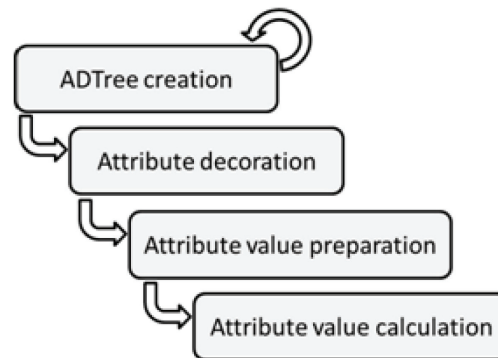
We would like to mention that two of the players consulted the ADTree to reexamine the context of the nodes, whereas the two other players estimated the values without taking

existing values of similar nodes or values of the parents and children into account.

The percentage of the nodes, where both players have given values, ranges between 0% and 93%, for different attributes. For five attributes (**detectability**, **impact**, **penalty**, **profit** and **probability**) either an attacker or a defender did not estimate a value. For the **special skill** attribute, we discovered that the attackers had not considered the same special skill. One defender estimated values only for basic actions and not for refined nodes.

In the remainder of this section, we elaborate on the difficulties we encountered while estimating the values. We have grouped the difficulties into *attribute related*, *node related* and *methodological* problems. They are based on our analysis of the data and the players' own evaluation.

Figure 8. The ADTree case study process



#### 4.2.1. Attribute Related Problems

Several problems, encountered while estimating values, are worth mentioning. First of all, we learned that the players tended to estimate values by using their own interpretation of the attributes and not the one provided in Table 1. This happened when the description given in Table 1 was not immediately understandable or when it was contradictory to the player's belief of what the attribute represented. Apparently, we forgot to mention that only the attribute description given in the table should be used. For example, the attribute **costs of attack** or **costs of defense** to "disable cameras" could be understood from two perspectives. From an attacker's point of view, it could mean how much the attacker has to pay to disable the cameras. From a defender's point of view, however, the **costs** of this attack could be the money the defender has to pay to get the cameras running again after the attack.

Table 1 describes **penalty** and **profit** only from the perspective of an attacker that performs an attack. Slightly different, the **impact** attribute describes consequences from a system's owner or defender's point of view, the scale given in Table 1 however indicates attacks. When attackers only estimate values for attack nodes and defenders only for defense nodes, both descriptions lead to nodes which did not have any values.

Even when Table 1 did not make a reference to only the defender or the attacker, some players believed that a certain attribute only concerns the other player's actions. As a result, the **detectability** attribute was interpreted by an attacker as detectability of the defense, and **probability** was interpreted by a defender as the probability that the attack succeeds. This again led to nodes without estimated values.

A third source of missing values was the description of the attribute itself. This occurred for the attributes **profit** and **special skill**. More concretely, the **profit** attribute is missing information on what can be gained with the attack. For example, we should have mentioned whether we plan to steal valuables from the warehouse or to annoy people that work there. Furthermore, the special skill description allows for a use of a Boolean domain, but does not specify an attribute such as whether or not **insider knowledge**, **electricity**, or certain **technical skills** are required. Table 1 offers an explanation for the different attribute value categories. However in some cases, the given category explanation was not precise enough, or only the keyword classifying the category was taken into account. For example, depending on the node, the value Medium for the attribute **time** can be understood as anything ranging from minutes to weeks. Indeed, consider the nodes "hide in bathroom" and "get hired": *being patient for a while* could mean *a couple of days* when we were hiring

Table 1. Decoration attributes for the RFID case study

Attribute	Description	Values
<b>Costs</b> (Schneier, 1999; Buldas et al., 2006; Tanu & Arreymbi, 2010; Baca & Petersen, 2010; Saini et al., 2008; Mauw & Oostdijk, 2005; Yager, 2006; Abdulla et al., 2010; Roy et al., 2011; Byres et al., 2004; Amenaza, 2011; Wang et al., 2011; Edge et al., 2006)	The amount of real money needed to finance the attack or defend against it (depending whether it is the attacker's or defender's point of view), referring to, e.g., equipment or software costs, educational expenses, development costs or size of a bribe.	Cheap (C): Any attacker or defender can afford this without thinking twice. Average (A): The costs of the attack will fend off most attackers without a steady income. Defenders will typically do a cost-benefit analysis before the expenses can be justified. Expensive (E): The attacker will need substantial funding in order to perform the attack. It is unlikely the defender will invest in this. Optional: Real number value.
<b>Detectability</b> (Tanu & Arreymbi, 2010; Byres et al., 2004; Amenaza, 2011)	The chance that the defender will notice the attack during its execution or the attacker will notice the defense mechanism.	Easy (E): Any attacker or defender with a clear state of mind will detect that something was out of the ordinary right away. Possible (P): Some attackers or defenders are able to detect this defense or attack. Difficult (D): Very few attackers or defenders are qualified to notice that something is wrong before it is too late.
<b>Difficulty</b> (Byres et al., 2004; Fung et al., 2005; Tanu & Arreymbi, 2010; Henniger et al., 2009; Amenaza, 2011; Mauw & Oostdijk, 2005; Abdulla et al., 2010; Amoroso, 1994; Wang et al., 2011)	The technical or social skill level needed for the attacker or defender to succeed.	Trivial (T): Little technical skill required. Moderate (M): Average cyber hacking or defense skills required. Difficult (D): Demands a high degree of technical expertise, the attacker is a professional con artist. Unlikely (U): Beyond the known capability of today's best attackers or defenders.
<b>Impact</b> (Schneier, 1999; Tanu & Arreymbi, 2010; Henniger et al., 2009; Li et al., 2009; Saini et al., 2008; Mauw & Oostdijk, 2005; Amoroso, 1994; Abdulla et al., 2010; Roy et al., 2011; Edge et al., 2006; Wang et al., 2011)	The severity or consequence from the system owner's point of view. Can refer to loss of money, but also other less tangible resources such as loss of reputation.	Low (L): The system owner will not care or notice. Moderate (M): Acceptable but unwanted loss. High (H): Unacceptable loss, must be avoided. Extreme (E): Will terminate business. Optional: Real number value.
<b>Penalty</b> (Buldas et al., 2006; Jürgenson & Willemsen, 2008; Wang et al., 2011)	The consequences for the attacker given that the attack fails, for instance a fine, jail sentence or being blacklisted. Here, we do not consider any penalty of a successful attack.	Low (L): The attacker will not care. Medium (M): The attacker will think twice before performing the attack. High (H): Very few attackers will take the risk. Optional: Real number value of a fine or years in prison.
<b>Profit</b> (Amoroso, 1994; Jürgenson & Willemsen, 2008; Bistarelli et al., 2007; Roy et al., 2011)	The economic profit or gain the attacker will receive should the attack succeed. This value does not include costs of attack.	None (N): A successful attack does not lead to any direct income. Marginal (M): Economic gain is not enough by itself to justify the attack. Lucrative (L): The attacker can obtain a substantial profit. Optional: Real number value.

*continued on the following page*

Table 1. Continued

<b>Probability</b> (Schneier, 1999; Buldas et al., 2006; Henniger et al., 2009; Li et al., 2009; Manikas et al., 2011; Yager, 2006; Abdulla et al., 2010; Roy et al., 2011; Byres et al., 2004; Edge et al., 2006; Wang et al., 2011)	The assumed chance that the attack or defense will succeed. Could be based on heuristics of similar attacks or cognitive estimations.	Unlikely (U): Below 5%. Low (L): Between 5% and 25%. Medium (M): Between 25% and 75%. High (H): More than 75%. Certain (C): Close to 100%. Optional: Specific percentage value.
<b>Special skill</b> (Mauw & Oostdijk, 2005; Abdulla et al., 2010; Schneier, 1999)	A specified skill or property the attacker or defender will need in order to succeed. This is orthogonal to the <b>difficulty</b> attribute. Examples are access to insiders or need of electricity.	Binary value: True (T): A special skill is required. False (F): No special skill is required.
<b>Time</b> (Henniger et al., 2009; Schneier, 1999; Wang et al., 2011)	For the attacker this is the time needed to perform the attack, independent of <b>difficulty</b> and <b>costs of attack</b> . For the defender this is the time needed until the defense is effective.	Quick (Q): The attack or defense can be performed in an instance. Medium (M): The attacker or defender will need to be patient and wait for a while. Slow (S): The attack or defense takes really long time to complete. Optional: Real number in terms of minutes.

someone, but it would mean *a couple of minutes* if we were hiding in the toilet.

For some attributes, e.g., **costs**, **probability**, **time**, it felt easier to estimate the more refined nodes, while for other attributes, e.g., **penalty**, **profit**, it was easier to estimate the less refined nodes. As a result, some players chose to only estimate attribute values they were moderately sure about. When uncertain about an attribute value, they favored not estimating over estimating a value with low confidence.

The description and the use of the meta-attribute confidence should have been described more clearly. One of the players chose to only use one confidence level per node, with the intention to save time at the expense of less accurate values. Other players selected a different confidence level for every estimated attribute value. All players concluded that the confidence level scale should be reduced to fewer (e.g., three) options.

#### 4.2.2. Node Related Problems

Issues directly concerning attributes and their specifications are not the only reasons why associating values with the nodes of an ADTree was a difficult task. Several problems were in fact related to the nodes themselves. Associating attribute values with the nodes helped us to realize that the user's understanding of the presented scenario is in many cases incomplete or even incorrect. One of the main problems was that node labels were often not self-explanatory and may lead to confusion. We used simple labels in order to be able to graphically represent the nodes. Thus, the labels were often short, e.g., "false management order," imprecise, e.g., "blocker reader," or did not contain verbs, e.g., "barbed wire." This implied that, without looking at the context in which the nodes had been used, i.e., parent, sibling and child nodes as well as the corresponding main goal, it was impossible to estimate the related attribute values. As an example, we can consider the "enter through door" node. Without taking its parent

node “get into warehouse” into account, it is impossible to estimate the corresponding values for the **time** and **difficulty** attributes, because the values may differ depending on which door we are interested in: the warehouse door, the bathroom door or the administrative office door.

Another issue is whether attribute values should be assigned to the non-refined nodes only or also to the refined ones? This problem is related to the meaning of refined nodes. In fact, some of them represent understandable attacks or defenses, e.g., “get onto premises” or “block tag reader,” others play the role of dummy placeholders, e.g., “trick.” In the first case, it was possible to associate values with such a refined node. In the latter case, they were only used to connect several options that could be attacked or defended in the same way. In such situations, the attribution was more problematic and could not be performed without taking the corresponding child nodes into account.

Finally, in order to quantify some of the considered properties, additional information may be required. Such information does not have to be related to the considered ADTree structure. For instance, it is hard to decide how long it takes to employ guards, without knowing what kind of goods are stored in the warehouse. If the goods are not very valuable, anybody could be hired as a guard. Thus, the execution **time** of “employ guards” would be Medium. However, if the goods are expensive, sensitive or dangerous for the environment, the guards have to be chosen more carefully. In this case, the selection process would take longer, because it would have to be accompanied by additional measures such as background checks. Thus, the corresponding attribute value would then be Slow.

#### 4.2.3. Methodological Problems

The case study was performed by four people: two of them played a role of the attacker and two of the defender. However, it was not explicitly specified to which nodes the players should assign the values. This led to inconsistencies in the gathered data. Each of the players provided

values for nodes related to his or her role, but one of the attacker players also estimated some values for the defender nodes. So a question arises: Which player should provide values for which nodes?

This issue is closely related to the problem of what the knowledge of the player is. Should we assume that the attacker and the defender are only able to estimate values for their own actions, or should the game give them the freedom to assign values to the adversary’s nodes as well? Furthermore, should the players only take the part of the scenario corresponding to their role into account, or can they base their decisions on knowledge about the other player too?

Another issue is how to assign values to the nodes that have the same labels. Nodes such as “secure warehouse” or “attended visitor’s log” were mentioned twice in the initial table, because they appear twice in Figure 5. On the one hand, the values assigned to two different occurrences of the same action might be different in the case when the context is taken into account. On the other hand, if the nodes are handled independently of the tree, it would be more reasonable to associate the same values with similarly labeled nodes.

We also identified an issue concerning the role of the players in the creation phase of the considered ADTree. As we were the creators of the tree, we had a good understanding of the tree. However, in general players might be asked to estimate values for trees which they have not seen before.

### 4.3 Preparation of Attribute Values

The attribute values have been estimated by four players. As a result we have obtained several values for a given attribute and a given node, as described in Table 2. In practice, we would like to have a single value that can give us some indication about the aspects we are interested in. In this section, we show how a single value can be derived from the raw data gathered in the previous step of the case study.

We start from the data provided by the players, as given in Table 2. If every player has

Table 2. Extract of raw-data of the case study. Each row represents the estimates of one player. The first letter of every field represents an attribute value, as abbreviated in Table 1 and the second represents the confidence level on a scale from 1 to 5.

Name of the Node	Costs	Det	Diff	Imp	Pen	Prof	Prob	Skill	Time
Breaking and Entering	A,3	P,3	M,3	M,3	M,3	-	-	F,3	Q,3
	-		M,4		-		M,4	F,4	Q,4
Get onto Premises	C,4	P,4	T,4	M,4	M,4	-	-	F,4	Q,4
	C,4		T,4		L,4		H,4	F,4	Q,4
Get into Warehouse	C,4	P,4	T,4	M,4	M,4	-	-	F,4	Q,4
	C,3		M,3		M,3		H,3	F,3	Q,3
[D,B] Monitor with Security Cameras	E,4	P,4	T,4		H,4			F,4	S,4
	E,4	E,4	M,4				M,3	T,4	Q,3
[B] Climb over Fence	C,5	D,5	T,5	L,5	L,5	-	M,5	F,5	Q,5
	C,4		T,4		L,4		H,4	F,4	Q,4
[B] Enter through Gate	C,5	E,5	T,5	L,5	L,5	-	H,5	F,5	Q,5
	C,4		M,4		M,4		M,4	F,4	Q,4
[B] Enter through Door	C,5	E,5	T,5	M,5	M,5	-	M,5	F,5	Q,5
	C,4		M,4		M,4		M,4	F,4	Q,4
[B] Enter through Loading Dock	C,5	E,5	T,5	M,5	M,5	-	L,5	F,5	Q,5
	C,4		M,4		M,4		M,4	F,4	Q,4
Disable Cameras	A,3	E,3	T,3	L,3	M,3	-	M,3	F,3	Q,3
	C,3		M,3		L,3		-	F,3	M,3
[D,B] Barbed Wire	E,4	D,4	D,4		H,4			T,4	S,4
	C,4	E,4	T,4				M,3	T,3	Q,2
[D,B] Monitor with Biometric Sensors	E,4	D,4	D,4		H,4			T,4	S,4
	E,4	E,3	M,3				M,3	T,3	Q,2
[B] Laser Cameras	A,3	E,3	M,3	L,3	M,3	-	L,3	F,3	Q,3
	A,2		M,2		L,2		M,2	F,2	Q,2
[B] Video Loop Cameras	A,2	D,2	D,3	L,2	M,2	-	U,2	T,2	M,2
	A,2		M,2		L,2		L,2	F,2	M,2
Guard Against Barbs	C,4	-	T,4	L,4	L,4	-	H,4	F,4	Q,4
	A,4		T,4		L,4		H,4	F,4	Q,4
[D,B] Employ Guards	A,4	P,4	T,4		H,4			F,4	M,4
	E,4	E,4	T,4				M,3	F,4	M,2
[B] Use Carpet on Barbs	C,5	E,5	T,5	L,5	L,5	-	C,5	F,5	Q,5
	C,4		T,4		L,4		H,4	F,4	Q,4
[B] Wear Protective	A,5	E,5	T,5	L,5	L,5	-	H,5	F,5	Q,5

continued on the following page



Table 2. Continued

Clothing	A,4		T,4		L,4		H,4	F,4	Q,4
:	:	:	:	:	:	:	:	:	:
Nodes with less than two values	11	65	6	79	24	79	47	9	12
Attack nodes with two values	54	0	59	0	55	0	32	56	53
Defense nodes with two values	14	14	14	0	0	0	0	14	14

estimated the same value, we can immediately use it as input to produce an indicator for the scenario. In practice, this perfect world scenario seldom occurs. There are several reasons why the values were not identical. First, some players may have opted not to estimate a value at all. Second, the understanding of the node may have been different for each player, as we pointed out in Section 4.2. Finally, it might also happen that no player has estimated a value for a node. Therefore, in a non-perfect world scenario, we need a method to choose one representative value for each node.

Provided all players had estimated a value for a given node, we automated the process of combining the values by using the following procedure. First, we transformed the attribute values into natural numbers, e.g., the **costs** attribute values Cheap, Average and Expensive, were transformed into 1, 2 and 3, respectively. We let  $n$  be the number of independently gathered pairs (attribute, confidence), i.e., in our case  $n$  was equal to 2. Then, we used the following formula.

$$\left( \text{rnd} \left( \frac{\sum \text{attribute} * \text{confidence}}{\sum \text{confidence}} \right), \left\lfloor \frac{\sum \text{confidence}}{n} \right\rfloor \right) \quad (1)$$

Here  $\lfloor \cdot \rfloor$  symbolizes rounding down to the nearest integer and rnd regular rounding. We chose to round the first component to use the best estimate. This rounding also allows us to transform the first component back into the original categories. We chose to round the second component down, to reflect risk averseness. The results of the application of the formula are

the non-bold values in Table 3. To illustrate the application of Formula 1, we combine the difficulty attribute values estimated for the “Enter through Gate” node. From Table 2 we can see that one of the players estimated the **difficulty** of entering through the gate as being Trivial (T) and his confidence level in this value was 5. The other player found the considered action as Moderate (M) in **difficulty** and his confidence level in this value was 4. The value domain for the **difficulty** attribute, as defined in Table 1, contains 4 values: Trivial, Moderate, Difficult and Unlikely, which we transform into 1, 2, 3 and 4, respectively. Thus, we use (1,5) and (2,4) as inputs for Formula 1. We obtain:

$$\left( \text{rnd} \left( \frac{1 * 5 + 2 * 4}{5 + 4} \right), \frac{5 + 4}{2} \right) = (1, 4)$$

This means that, after combining the values estimated by the players, we obtain that the **difficulty** of “entering through the gate” is Trivial (1), of which we are certain with a confidence level of 4. The **cost**, **difficulty**, and **time** values, as well as the corresponding confidence levels, obtained for the remaining nodes by applying Formula 1, are depicted as non-bold pairs in Table 3. We defer a discussion of other possible methods to combine values to Section 5.2.

In our case study, we encountered several exceptional cases, when using the formula was either impossible, because, for example, not a single value was given, or doubtful, because the values differed substantially. These critical cases we classified and discussed at a *consensus meeting*.

Table 3. The table after the consensus meeting. The first letter of every field represents an attribute value, as abbreviated in Table 1 and the second represents the confidence level on a scale from 1 to 5.

Name of the Node	Costs	Diff	Time
Breaking and Entering	A,3	M,3	Q,3
Get onto Premises	C,4	T,4	Q,4
Get into Warehouse	C,3	T,3	Q,3
[D,B] Monitor with Security Cameras	E,4	M,4	S,4
[B] Climb over Fence	C,4	T,4	Q,4
[B] Enter through Gate	C,4	T,4	Q,4
[B] Enter through Door	C,4	T,4	Q,4
[B] Enter through Loading Dock	C,4	T,4	Q,4
Disable Cameras	A,3	M,3	Q,3
[D,B] Barbed Wire	A,4	T,4	S,4
[D,B] Monitor with Biometric Sensors	E,4	D,3	S,5
[B] Laser Cameras	A,2	M,2	Q,2
[B] Video Loop Cameras	A,2	D,2	M,2
Guard Against Barbs	A,4	T,4	Q,4
[D,B] Employ Guards	E,4	T,4	M,3
[B] Use Carpet on Barbs	C,4	T,4	Q,4
[B] Wear Protective Clothing	A,4	T,4	Q,4

The main goal of the consensus meeting was to obtain values that can then be used as unbiased input for further calculations on the ADTree. For example, we reduced the unbalance that would be introduced if we considered nodes with four times the same value equal to those where only one of the players had estimated this value. Since for several attributes we did not have enough data, we focused only on the attributes **costs**, **difficulty**, and **time**, at the consensus meeting.

We first ensured that the nodes were correctly classified, i.e., that the players had not mistakenly estimated wrong value. Whenever mistakes were discovered, they were corrected, and the nodes were reassigned to the correct category, before the actual conflict resolution started. We also uncovered one inconsistency where the scenario and the tree were not matching. To repair this mistake, we corrected the

tree. To obtain agreed values, the nodes were analyzed in context, i.e., we looked at the parents and the children of the nodes, but without considering any values assigned to these nodes. More concretely, we identified the following categories and resolved the problems in the following way

- *Nodes where no one had estimated a value:* We opted to discuss the value and eventually assign a single value. The players who had taken the opposite role commented on plausible values, then we selected one with which we all agreed.
- *Nodes where not every player had estimated a value:* We also decided on a single value at the consensus meeting. Concretely, the player who had not given a value, first explained why he had not done so, then

the player who had given a value explained his choice. Then, a consensus was formed.

- *Nodes that had non-neighboring values:* The player with the lower value explained his choice, and then the player with the higher value explained his. After that, the involved players agreed on one of the given values, or a compromise was chosen. Whenever a compromise was chosen, we lowered the confidence value.
- *Nodes where all given values have low confidence levels:* We also planned to discuss these uncertain values, but we ran out of time, skipped this step and applied Formula 1.

The final results of the consensus meeting are given as pairs in **bold font** in Table 3. Instead of the allocated hour, we spend two hours discussing the values. Out of the  $3 \times 79$  possible values, there were 188 cases for which we applied the Formula 1, 8 cases without any assigned value, 24 cases to which we had assigned only one value and 17 cases where the values diverged significantly.

#### 4.4 Bottom-Up Calculation of Attribute Values

As already mentioned in Section 4.2, assigning relevant values to all nodes of an ADTree may be difficult or even impossible. Fortunately, the ADTree methodology allows us to automate the calculation of values on ADTrees with the help of the bottom-up procedure. To use the bottom-up evaluation, we first have to initialize values at all non-refined nodes of the tree. Then, the values for all subtrees, and in particular for the entire tree, are calculated, using functions which depend on the type of the root of the subtree and the considered attribute. In this section, we first show how to calculate, in the bottom-up way, the values for the minimal **costs** of the attacker. Then, we use the obtained values to analyze the warehouse scenario.

In the case study, we were interested in calculating the minimal **costs** of a successful attack in the RFID warehouse scenario. We

considered the situation where the attacker did not have any precise information on how the defender will decide to protect the warehouse. Thus, for this calculation, we assumed that all possible defenses illustrated on the ADTree were in place and that they were fully functional, i.e., a defense attached to an attack node defeats the corresponding attack component, unless the defense itself is rendered useless by a counterattack.

We started by initializing the values for the non-refined nodes of the tree. In the case of the attacker's non-refined nodes, we used the pairs (costs, confidence) from Table 3 as initial values. As the defender's **costs** do not have influence on the attacker's **costs**, we did not use the values from Table 3 in the case of the defender's non-refined nodes. Instead, we introduced an additional cost value Infinite, denoted by  $X$ , which represents infinite **costs** and we assumed the attacker is not able to afford it. The non-refined nodes of the defender were initialized with the pair  $(X, 5)$ . This indicates that we are fully confident that it is infinitely expensive (and thus impossible) for the attacker to be successful at a defender's action. Such initial values allow us to express the **costs** of the considered scenario from the point of view of the attacker.

Since we were interested in the minimal **costs**, we have to know how to compare different values. Thus, we considered the following linear order Cheap < Average < Expensive < Infinite. Now, we are ready to describe how we calculated the minimal attacker's **costs** for all subtrees. The bottom-up procedure is recursive, i.e., we start from the leaves and we calculate the value for every subtree rooted in a parent node based on the values previously calculated for the subtrees rooted in its child nodes.

In this framework, we chose the minimal value for an attacker disjunctive subtree, because we were interested in the minimal **costs**. Thus, we supposed that the attacker always performs the least expensive option. Moreover, we assumed that performing several actions belonging to the same **costs** category is not more expensive than performing one of such

actions. Therefore, we chose the maximal **costs** in the case of a conjunctive subtree with the attacker's root. Conversely, in order to successfully disable a disjunctively refined defensive countermeasure, the attacker has to disable all corresponding refining options. Therefore, we used the maximum operator in this case. On the other hand, to successfully disable a conjunctively refined defensive countermeasure, it is sufficient for the attacker to disable only one of the corresponding refining actions. Here again, according to our assumption, the attacker chooses the least expensive solution. Thus, the operator used in this case is minimum. Finally, we always propagated the maximal confidence level, corresponding to the chosen **costs** value. This allowed us to express how certain we can be about this value.

The three paragraphs below summarize which operators are used for calculation of **costs** values for all possible subtrees.

#### 4.4.1 Subtrees Rooted in a Node which is Refined but not Countered

- The **costs** calculated for a subtree rooted in a disjunctively refined attack (resp. defense) node is defined as the minimum (resp. maximum) of the **costs** calculated for its refining subtrees.
- The **costs** calculated for a subtree rooted in a conjunctively refined attack (resp. defense) node is defined as the maximum (resp. minimum) of the **costs** calculated for its refining subtrees.

The maximal confidence level corresponding to the chosen **costs** is propagated.

#### 4.4.2. Subtrees Rooted in a Node which is not Refined but Countered

The **costs** calculated for a subtree rooted in an attack (resp. defense) node is defined as the maximum (resp. minimum) of the initial value for the non-refined root node and the value calculated for the countering subtree.

The maximal confidence level corresponding to the chosen **costs** is propagated.

#### 4.4.3. Subtrees Rooted in a Node which is Refined and Countered

In this case, first a pair corresponding to a refining part of the tree is calculated, as in the case of a subtree rooted in a refined but not countered node. Then, the functions for a subtree rooted in a non-refined but countered node are used, where the initial value for the root is replaced with the calculated value for the refining part.

We would like to remark that the functions used to calculate values depend on the considered attribute and additional assumptions. Thus, if we would be interested in calculation of the defender maximal **costs**, for instance, the used functions would have to be redefined accordingly. It is easy to see that the functions presented in this section are also suitable for calculation of the minimal **difficulty level** of the attacker and the minimal **time of an attack**, under the assumption that all defensive components are present and fully functional.

With the assumption that all the possible defenses are present and fully functional, the real minimal **costs** of a successful attack can be lower than the one obtained using our calculation. Indeed, in reality, the defender may decide not to implement some of the defenses and thus the **costs** of the corresponding counterattacks will not be taken into account for the final **costs** of the attacker. However, by taking the described approach we use a safe solution, in the sense that

- The calculated minimal **costs** will not be lower than the actual minimal **costs**, i.e., the minimal **costs** will not be underestimated,
- And the resulting set of attack components that have to be executed in order to achieve the cheapest attack forms a successful attack which, according to our scenario, the defender cannot defend against.

In the rest of this section, we use the calculated values for minimal **costs**, minimal

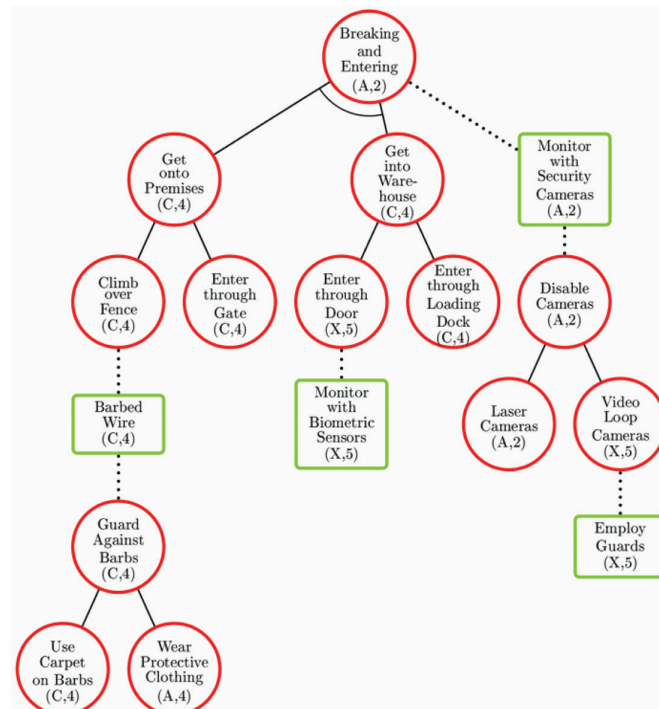
**difficulty** and minimal **time** to analyze the RFID-based warehouse scenario. These attributes serve as representative examples. Our calculation shows that, to achieve an attack with the minimal **costs**, an attacker needs to spend Average amount of money. We have a confidence level 3 in this value. The corresponding attack consists of “disabling the backend”. For the attack of minimal **difficulty**, the attacker should also “disable the backend”. Its **difficulty** is Medium of which we are confident with level 3. The **time** it takes to perform the fastest attack is Quick of which we are confident with level 2. To achieve the fastest attack, an attacker should “disable the RFID tag”. We observe that for all three attributes, the optimal attack option is something which we chose not to refine, see Figure 4. To be able to give a more insightful example, we look at the subtree rooted in the node “breaking and entering”. The values

resulting from the bottom-up approach for the **costs** attribute are depicted in Figure 9.

From the bottom-up calculation presented in Figure 9, we deduce that an attacker can “break and enter” when he spends an Average amount of money, and we are confident with level 2 about that. To perform the attack, an attacker has two options: either he has to “use a carpet on the barbs,” “climb over the fence,” “enter through the loading dock” and “laser the cameras,” or he has to “enter through the gate,” “enter through the loading dock” and “laser the cameras.”

Using the bottom-up approach, we computed the minimal **costs** of an attacker for every subtree. This is the information that is depicted in every node. Comparing these values with the ones gathered in Table 3 seems like a natural consequence. In Section 5.3 we elaborate on the

Figure 9. The breaking and entering subtree with costs calculated in the bottom-up way



this and other questions which illustrate the differences between attack trees and ADTrees.

## 5. DISCUSSION

While performing this case study we encountered numerous design choices concerning the ADTree methodology. Some options were outright inadmissible, some easy to pinpoint, whereas for others the multiplicity of possible solutions proved the versatility of the ADTree methodology. In Section 5.1 we discuss issues related to estimation of values by players. Then, in Section 5.2, we present the alternative we were faced with when fusing several values into one final value that we then use as initial assignment in the bottom-up algorithm. Section 5.3 elaborates on the choices and problems we had during the actual bottom-up calculation. Section 5.4 contains a general discussion about the usefulness and the benefits of the entire ADTree methodology and elaborates on the four conflicting modeling goals. Finally, Section 5.5 shows the hindsight guidelines we have learned based on our specific case study.

### 5.1 Lessons Learned from ADTree Decoration

Some attributes like **penalty**, **profit** or **impact** were only estimated by either attackers or defenders. Assigning a specific role to a player initially seemed like a good idea, as it caused the players to minimize their work, such that (with a small exception) attackers only estimated values for attack nodes and defenders for defense nodes. This reduces the number of estimated values by a half, but it is doubtful that the quality was twice as good. So, the first lesson here would be to have a clear understanding of what kind of players you have available and how to assign them. If you have specialists within certain domains available, make sure to exploit that. For example, a janitor could estimate nodes related to physical building security, i.e., nodes depicted in Figure 6, whereas a psychologist might be better suited to estimate values for nodes related to social engineering, i.e., nodes

depicted in Figure 5. A system administrator will in most cases know more about historical attacks than a software developer, and so on.

Another alternative would be to let specialists estimate values for all nodes but only for the attributes related to their field of expertise. For instance, accountants would be better suited to provide **costs** values and technical personnel could take care of deciding whether **electricity is needed** to perform the considered actions.

If the players are not specialized, e.g., you do not have that kind of people available or can afford to hire a seemingly trustworthy black hat hacker; we believe that a random assignment of nodes to be estimated is hardly justifiable. In such a case, it might be helpful to assign roles according to the node types to different players to transform them into attackers and defenders. We do not have sufficient evidence to recommend whether the players should then estimate only one type of nodes (i.e., attack node or defense nodes) or both, i.e., whether they should predict the strategy of their opponents and reflect that in the node values belonging to their role, or whether they should influence the results by estimating values for all non-refined nodes. The best data from non-specialists would probably be obtained from having the players play as both defenders and attackers; however we do suspect that having a friendly competition between two opposing teams would serve as a good motivation.

Even though all players were involved in the creation of the tree, there were cases where some of the players did not completely agree with its structure when estimating the values. Here, we feel it is best to run with a dual strategy. First, the player should nonetheless provide a value, but assign a very low confidence. Second, we should introduce a new attribute called **disagree with node** with a binary value range. Any node that has been flagged with this attribute should then be discussed at the consensus meeting. It is important to remember that other players may feel differently about the model and too many structural changes will make other values in the tree insignificant. To avoid such complications and especially repeti-

tive work, it is important that the model of the tree is sufficiently accurate before any values are assigned.

To make full use of meta-attributes, they should always be estimated on a per attribute value basis. In particular, not all attribute values for a given node should be assigned with the same confidence level. This increases the time it takes to assign values, but it also increases the accuracy of any calculation. Furthermore, a 3-valued confidence scale should be enough. As for the attribute values, each confidence level should be clearly explained. Otherwise distinguishing between different confidence levels is somewhat arbitrary.

If only the bottom-up approach is used during the analysis phase, estimating values for refined nodes is questionable. If an action is already sufficiently comprehensible that a reliable value could be suggested, it would not need to be refined anymore. Hence, from the fundamental modeling idea behind ADTrees, assigning values to intermediate nodes is an inherent contradiction.

On the one hand, node labels are very important because they help the players to understand the scenario without reading the scenario description in detail. Node labels that are too short may lead to confusion. On the other hand, the labels should be concise because if they are too long and detailed they are difficult to display and reduce the possibility of reusability. Therefore, to satisfy both criteria, we propose to always use a noun and a verb as node labels.

During the game, players raised the question whether nodes should be considered without their context, e.g., neighboring nodes and previously assigned values to similar nodes. We believe that if the values had been assigned inside the actual ADTree, and not in a separate table, one would have to consider the context. If the context is taken into account, the node labels might be easier understandable (due to the additional information the parent node gives). We are aware that assigning values without context can be used to detect inconsistencies and random assignments. This however, we would

pay for by less accurate values, because the node by itself is then less descriptive. Estimating a value without context might not even be possible, because a bias that might have been introduced during the creation of the tree. The player may simply remember the context. Even though refined nodes should not hold less information than the children, repeating the label of the parent clutters the label of the child nodes.

## 5.2 Lessons Learned from Attribute Preparation

Due to the different background or knowledge of the players, the estimated values will rarely be the same. Furthermore, some players may chose not to insert a value at all. For these reasons, we end up with heterogeneous data that needs to be homogenized. In Section 4.3 we describe one possible option to proceed by using Formula 1 and to discuss the remaining values at a consensus meeting. The formula consists of a weighted average for the attribute value and an estimation of the confidence value, which reflects risk averseness. Instead of applying this formula, it is possible to for example, choose the average, the median, the most often used or the lowest value for the attribute value and an average, a reduced average, the lowest or possibly a new level of confidence as the confidence value. The desired method may vary, depending on the scenario and the attribute to be calculated.

When using any formula, it is, in general, preferable to have as many input values as possible, since this increases the significance of the result. However, the meaningfulness of the values may depend on the actual values that were estimated. To differentiate the inputs, we classify the input values into six categories. Then, for each of the following six categories we can choose a different approach of how to combine the estimated values in to one, e.g., using an average or a minimum value, using ranges as values or deciding on the final value at the consensus meeting. The proposed categories are:

- Category 1: Nodes with as many attribute values as players.
- Category 2: Nodes where at least one player has not estimated a value.
- Category 3: Nodes where all estimated values have a low confidence.
- Category 4: Nodes where the values diverge significantly.
- Category 5: Nodes where the **disagree** flag is set.
- Category 6: Nodes where no player has estimated a value.

The categories are ordered according to a descending scale of automation. Whereas for Category 1 it is entirely reasonable to combine the input values automatically into a single value, this is not even possible for Category 6. A decreasing automatic treatment is equivalent to an increased necessity for a consensus meeting. We again observe the conflict between modeling time and modeling accuracy. Holding a consensus meeting, will result in an improvement of model accuracy for most of the nodes from Category 5 or 6, because either the model was actually wrong or the model was not described clearly enough. Categories 3 and 4 only vaguely depict a design option, since the terms *low confidence* and *diverge significantly* would need to be defined, in more detail. Nonetheless, even defining low confidence as only values with confidence level of 1 or 2 and defining *diverge significantly* as are not neighboring in the natural order, already improves the model. For example, for the attribute attack **time**, the node “postal Trojan attack” was put in Category 4. While reviewing this node at the consensus meeting, we discovered that one of estimated values was mistakenly given. Another example of model improvement was the node “hide in bathroom”, where the divergent values started a discussion which led to the insight that an attack component was missing.

A different classification we have to consider is the domain of the attributes. When actually estimating values for non-refined nodes, naturally the question arises of why we only have three or four different possible values for each

attribute. From a theoretic point of view, it is entirely possible to use real numbers, intervals or even discrete probability functions as value domains. However, the more detailed the information a person has to estimate, the less likely he is inclined to provide a value. Using a more fine-grained scale to achieve more exact results is counterproductive, if the number of people, who estimates a value, decreases. Furthermore, increasing the graining of the scale may make it more difficult to distinguish between values.

### 5.3 Lessons Learned from Calculation

Comparing the values from Table 3 with the values calculated using the bottom up approach shows that the countermeasures are usually disregarded when we try to assign values to nodes on an intuitive basis. Therefore, we should not perform such a comparison. Figure 9 shows, for instance, that video looping cameras is infinitely expensive and thus impossible when guards are employed. This is contradicted by the estimated **costs** value mentioned in Table 3, which is Average with confidence level of 2. A similar negligence of countermeasures and subsequent counterattacks occurs if we consider attack trees instead of ADTrees. If we remove all subtrees rooted in defense nodes from Figure 9, we do not model that an attacker should worry about possible defenses, such as “barbed wire” or “monitor with security cameras.” Then, the **costs** value of the cheapest scenario would be Cheap, with the confidence level of 4. The corresponding attack would be to enter through the gate and the loading dock, undetected.

Since ADTrees allow us to combine information about the attacker and the defender, the ADTree formalism allows us to answer questions that depend on both players. We can, for instance, compute the minimal **difficulty** of an attack, assuming that the budget of the defender is limited to Average. Using Table 3 and Figure 9 we see that, in this case, monitoring with biometric sensors as well as with security cameras would be too expensive



for the defender. Hence, there would be four possible attack scenarios:

- Using the carpet on the barbs, climbing over the fence and entering through the door of the main building,
- Using the carpet on the barbs, climbing over the fence and entering through the loading dock,
- Entering through the gate and the door of the main building,
- Entering through the gate and the loading dock.

Similarly, it is possible to compute all combined attributes mentioned in Section 2.1, provided we know the values given in Table 3.

In the minimal **costs** calculation performed in Section 4.4, we have chosen to use the functions minimum and maximum. It is possible to redefine the used functions, in order to more accurately express how costly a combination of actions from different categories is, for instance, that performing a sequence of actions which are Cheap and Average is actually Expensive. However, in this case study we are more interested in the proof of concept rather than in precise computations, thus we use simple functions. It is clear that when our assumptions change, we have to redefine the used functions accordingly. For instance, the function minimum cannot be used any more for a disjunctively refined node, if we assume that the attacker is able to implement several among existing options and not only the cheapest one.

The minimal costs calculation, performed in Section 4.4, can be made more precise with the help of attributes expressed using Boolean values. Such attributes are well suited to reason about hypothetical scenarios. If we are sure that some of the hypothetically possible attacks or defenses do not occur, we can model this by pruning the tree. Pruning nicely fits in the framework of the bottom up propagation, when we use Boolean values. Let us, for instance, consider the attribute **is electricity needed**. We can prune the tree to simulate what happens if there is a power outage. Since power outage affects the

attacker as well as the defender, ADTrees are the formalism we want to employ. In this case, pruning the tree is not done by simply cutting off defenses, rather we cut off the subtrees rooted in the nodes which need electricity and all parents until we either hit a node of the other player or a disjunctively refined node. Pruning can also be used to reason about parts of the scenario that satisfy a certain property, like for instance their **costs** is lower than a certain threshold. In such cases, one player can prune the tree according to his knowledge and assumptions about the other player, to get a better overview of a realistic scenario.

#### 5.4 Lessons Learned from the Methodology in General

The attack tree obtained from our ADTree by removing all subtrees rooted in defense nodes depicts the actual attack scenario, all other nodes describe hypothetical defenses and attacks that may or may not be in place. Therefore, the security cameras from the floor plan, see Figure 3, are not explicitly modeled as defense nodes in our ADTree. The cameras mentioned in Figure 6 are additional cameras that could be put into place. Modeling this way clearly distinguishes between the considered scenario and hypothetical attacks and defenses. When modeling an actual scenario this approach might be appropriate, whereas when we want to store possible attacks and defenses in a library it is preferable that all defenses, including the already existing ones, are depicted in the ADTree. Consequently, when using libraries as the starting point for an ADTree, we have to adapt the tree to the scenario. Moreover, we can include information we have about the attacker/defender, by adjusting the tree to the considered situation before actually starting a bottom-up calculation. For example, if we know that the defender only has a limited budget, we could remove any attack that is too expensive and then start the bottom-up calculation on the pruned tree.

Since an incorrect or missing value anywhere in the tree can affect the resulting value

for an attribute, this indicates that the level of node refinement is crucial. To avoid biased results, the level of refinement should roughly be the same for all branches. For us this means that for most of the non-refined nodes we have the same intuitive level of understanding.

The level of refinement may be influenced by who created the tree and how it was created. First, the players can be given the tree and act as independent security experts or they can have created the tree (even described the corresponding scenario). Second, tree creation can start from templates available in security repositories, as suggested by Meland et al. (2010) or from an empty sheet of paper. In either combination, there is a trade-off between time and creativity.

An observation that we found interesting is that some of the attributes, such as **skill**, **attack costs** and **insider required**, tell us a lot about the requirements for the attacker. This would actually allow us to generate attacker profiles based on specific projections of the tree. Having such profiles would be of benefit for the defender in order to identify potentially harmful candidates.

There are four conflicting modeling goals we would like to emphasize: time, reusability, accuracy and simplicity, which all have implications on the complexity of the analysis. In modeling, *time* is always a concern. According to our experience, companies spend between one hour and one week on threat modeling before the implementation starts. From a theoretic point of view this might not be enough, but unless we see a paradigm shift in security modeling, time is always a scarce resource. The amount of time (and therefore money) spent always has to be justified by either allowing the analysis to be highly reusable or require a high degree of accuracy.

If we spend a lot of time modeling, we prefer our analysis to be *reusable*. For graphic security modeling, libraries immediately come to mind. For this reason, the SHIELDS project (SHIELDS, 2008-2010) developed an online library for (among others) attack trees. This library could be extended to also include attack-defense trees. Whereas a repository for the

structure of attack scenarios already exists, there has not yet been an attempt to also store node values together with the structural information. The degree of reusability might not be as high for actual values. Therefore, instead of storing concrete values, it might be preferable to store ranges of admissible values which serve as possible and not actual values. The more values for different attributes are retrievable, the more likely some information will be reusable. Using stored values may again conflict with other modeling goals, such as a fast scenario analysis (the stored node values most likely still have to be adapted) and, unless a computer tool is used, the visual appeal of the ADTrees is diminished, because the tree feels cluttered.

A third conflicting modeling goal is the *accuracy* of the model and the values. It is necessary to find an acceptable compromise between the required time and the necessary accuracy. Also, more accurate ADTrees and values reduce the reusability of the ADTree. In general, we can say that the coarser the value, the more raw data we get, because more people feel comfortable with actually providing the value. The finer the value, the more precise the result will be, but if the values are too fine, only experts might be able to estimate values. A coarse value range for a **costs** attribute would, for example, be Low, Medium, and High, a fine grade would be if the value was given as a real number expressing a monetary value, e.g., in €.

As a last modeling goal, we want the ADTree methodology to be easily *understandable*. We use a simple tree structure which is a main advantage over the generalized petri-net approach. But we do not only want the relation between the basic actions depicted in an easy way, we also want non-experts to be able to make full use of the ADTree methodology. Therefore, we also want a common user, developer, administrator or system owner to be able to estimate values for basic actions. By doing this, we benefit from a larger resource pool of potential attribute assigners, which might reduce the costs, because we do not need to hire an expert for tree creation and providing values.

This however, might have implications on the accuracy of the values.

### 5.5 Hindsight Guidelines for the Warehouse Case Study

Earlier in this section, we have elaborated on possible methodology design choices that typically occur in case studies such as ours. The “right” choices depend on the actual scenario, the security relevant questions to be answered, the modeling goals, the client, and last but not least, the people performing the case study. None of them should be treated in isolation. In Table 4, we take this discussion into account and list numerous design choices for the presented RFID case study. The bold options indicate which of the choices we would select with hindsight, but are of course not necessarily the right choices for other system settings.

## 6. CONCLUSION AND FUTURE WORK

In this paper, we have looked at the use of attributes for attack–defense trees (ADTrees). After explaining the ADTree formalism and giving an overview of typical attributes for attack trees found in the literature, we have described a case study for an RFID-based system managing goods in a warehouse. An ADTree depicting possible attacks on the considered system and the subsequent countermeasures was created. Then, a set of suitable attributes was selected and their values were determined. The results of the attribute evaluation can be used as a part of a redesign process or risk analysis in order to improve the security of the system.

The main contributions from this case study are practical experiences and user feedback. Taking the lessons learned during the case study into account, we have extended the original case study process graph depicted in Figure 8. Below we present the resulting six steps guideline which suggests a work flow and lists possible design choices that we recommend for applying the ADTree methodology when performing case studies.

1. *Create an ADTree for the scenario:* An ADTree is created using all available information and support tools. The attack tree, obtained from an ADTree by ignoring all defense nodes and the corresponding subtrees, depicts the main attack scenario. All other nodes describe hypothetical defenses and counterattacks.
  - People with different knowledge and relationship to the system, e.g., developers, security experts, system owner and end users, should be involved in the tree creation.
  - Different material, such as system specifications, floor plans, blueprints, work descriptions, attack tree libraries and attack patterns, should be used to create the tree.
  - The creation of the tree should be an iterative process which should end when there is mutual agreement between the involved parties. Modifying the tree after Step 3 should be avoided.
  - Node labels should contain a verb and a noun and concisely represent an attack or defense action.
2. *Choose and describe attributes:* Relevant attributes and meta-attributes are chosen, based on the security questions to be answered.
  - A clear, written description of chosen attributes and meta-attributes should be provided.
  - The description of each attribute and meta-attribute should include a domain which is used to quantify the values.
  - In the case of discrete domains, a written definition for each introduced category, such as small, medium, big, should be provided.
  - A user should be allowed to express whether he disagrees with a part of the tree, e.g., by including the **disagree with node** attribute in the list of attributes.
3. *Choose who estimates attribute values:* Decide which and how many people esti-

Table 4. Work flow–exemplary guidelines for the use of ADTrees for our case study

Step	Task	Design choices
1	<b>Create ADTree for scenario</b>	<ul style="list-style-type: none"> <li>- Create tree from root node on/<b>adapt tree from existing templates</b>.</li> <li>- <b>Use</b>/do not use incomplete trees.</li> <li>- Continuously improve trees/<b>freeze tree structure at some time</b>.</li> <li>- Use <b>concise noun and verb</b>/detailed textual description as node labels.</li> <li>- <b>Security expert</b>/system owner/random person creates tree.</li> <li>- Use <b>same</b> level of detail for refinements/limit number of nodes.</li> <li>- Allow/<b>disallow</b> pruning.</li> <li>- <b>Assume</b>/do not assume players are the creator of the tree.</li> </ul>
2	<b>Choose and describe attributes</b>	<ul style="list-style-type: none"> <li>- Use attribute description given in <b>Table 1</b>/provide new descriptions.</li> <li>- Select attribute domains: <b>discrete</b>/real numbers/fuzzy sets/ intervals/probability measures.</li> <li>- <b>Allow</b>/disallow (<b>disagree with node</b> attribute).</li> <li>- <b>Always</b>/sometimes use meta-attribute confidence.</li> </ul>
3	<b>Choose who estimates what</b>	<ul style="list-style-type: none"> <li>- Who estimates: attackers/defenders/<b>specialists</b>/ random people.</li> <li>- Which nodes: according to role of player/to <b>background</b>/ depending on attribute/all nodes.</li> </ul>
4	<b>Estimate values</b>	<ul style="list-style-type: none"> <li>- Evaluate meta-attributes for all attributes <b>separately</b>/ together.</li> <li>- Consider nodes <b>in</b>/without context.</li> <li>- Allow/<b>disallow</b> different values for repetitive nodes.</li> <li>- <b>Do not estimate</b>/estimate values for intermediate nodes.</li> </ul>
5	<b>Combine values</b>	<ul style="list-style-type: none"> <li>- Apply standard combining procedure for <b>Categories 1–4</b>/for other categories.</li> <li>- Use <b>Formula 1</b>/something else as standard procedure.</li> <li>- Use averaging/minimization/majority/<b>consensus meeting</b> for alternative categories.</li> <li>- <b>Restrict</b>/do not restrict time in case of consensus meetings.</li> </ul>
6	<b>Calculate values</b>	<ul style="list-style-type: none"> <li>- Use <b>predefined</b>/other functions from software tool or literature.</li> <li>- Compare/<b>do not compare</b> with intermediate values.</li> </ul>

mate which values. Optimize the number of people with respect to the available resources.

- In order to avoid errors and take into account different perspectives, more than one player should estimate attribute values.
  - Each player should obtain clear, written instructions detailing which values to estimate. It is not necessary that each player estimates the values for all nodes and/or all attributes, however it should be mandatory that he provides the values he is assigned to estimate.
4. *Value estimation*: The players selected in Step 3 estimate the values of the attributes chosen in Step 2 with the help of the support material identified in Step 1.
- The values should be estimated *only* based on the attribute and meta-attribute descriptions provided in Step 2.
  - When the bottom-up approach is used, the values should only be estimated for non-refined nodes.
  - The **confidence** meta-attribute should express a user's confidence in the provided attribute value. It should therefore be given for each estimated attribute value separately.
  - The attribute values should be estimated based on the node's context in the tree.
5. *Value combination*: When the attribute estimates from different people diverge, a combined value needs to be obtained. This value should be the best representative for all input values.

- Nodes should be partitioned into categories, depending on clear objective criteria, such as percentage of **coverage**.
  - The best way of deriving the representatives should be selected independently for each category, e.g., use a suitable formula, average or decide at a consensus meeting.
  - In case a consensus meeting is called for, its duration should be limited.
6. *Value calculation*: If the bottom-up approach is to be applied, suitable functions need to be chosen in order to calculate values for all the subtrees of a considered tree.
- The used functions should be in accordance with the attribute descriptions provided in Step 2.
  - Scientific papers discussing attribute evaluation and existing attack tree tools can be consulted in order to define the appropriate functions.
  - Estimated values of refined subtrees should not be compared with values resulting from the bottom-up algorithm, unless the tree does not contain any defense nodes.

When performing the case study, it became apparent that a software tool supporting the security analysis using attributes on ADTrees would be of great value. Such a tool is currently under development at the University of Luxembourg. Its main objective is to facilitate the work with the ADTree formalism by answering questions pertaining to security aspects based on realistic models. In particular, such a tool should lend support during input of values, show different tree views that focus on specific parts of a scenario, prevent repetitive tasks, lend support while computing values and be able to generate attack scenarios.

In the future, we intend to carry out another case study using the ADTree methodology. This will help us to further substantiate the ADTree formalism. We hope that it will allow us to expand our recommendations on modeling

choices, fine-tune the attribute descriptions, lead to more insights about which attribute domains to choose in which case and test our software tool.

Another line of research that we foresee is to consider the use of the ADTree methodology in diagnostics and forensics. More explicitly, we would like to look at the question whether ADTrees can be usefully employed once an attack has occurred in order to reconstruct what happened.

## ACKNOWLEDGMENTS

We would like to thank Ton van Deursen and Domenico Rotondi for their invaluable insights in RFID technology and their help to create the ADTree used in this case study. B. Kordy and P. Schweitzer were supported by grants from the National Research Fund, Luxembourg, with No. C08/IS/26 and PHD-09-167, respectively. P. H. Meland and A. Bagnato have received funding leading to these research results from the European Union Seventh Framework Programme (FP7/2007-2013) under grant No. 215995 and 257930, and 215995, respectively.

## REFERENCES

- Abdulla, P. A., Cederberg, J., & Kaati, L. (2010). Analyzing the security in the GSM radio network using attack jungles. In T. Margaria & B. Steffen (Eds.), *Proceedings of the 4th International Conference on Leveraging Applications of Formal Methods, Verification, and Validation - Volume 1* (LNCS 6415, pp. 60-74).
- Amenaza. (2011). *SecurITree*. Retrieved from <http://www.amenaza.com/>
- Amoroso, E. G. (1994). *Fundamentals of computer security technology*. Upper Saddle River, NJ: Prentice Hall.
- Baca, D., & Petersen, K. (2010). Prioritizing countermeasures through the countermeasure method for software security (CM-Sec). In M. A. Babar, M. Vierimaa, & M. Oivo (Eds.), *Proceedings of the 11th International Conference on Product-Focused Software Process Improvement* (LNCS 6156, pp. 176-190).

- Bistarelli, S., Dall'Aglio, M., & Peretti, P. (2007). Strategic games on defense trees. In T. Dimitrakos, F. Martinelli, P. Y. A. Ryan, & S. Schneider (Eds.), *Proceedings of the 4th International Workshop on Formal Aspects in Security and Trust* (LNCS 4691, pp. 1-15).
- Buldas, A., Laud, P., Priisalu, J., Saarepera, M., & Willemson, J. (2006). Rational choice of security measures via multi-parameter attack trees. In J. Lopez (Eds.), *Proceedings of the First International Workshop on Critical Information Infrastructures Security* (LNCS 4347, pp. 235-248).
- Byres, E. J., Franz, M., & Miller, D. (2004, December). *The use of attack trees in assessing vulnerabilities in SCADA systems*. Paper presented at the International Infrastructure Survivability Workshop, Lisbon, Portugal.
- Dacier, M., & Deswarte, Y. (1994). Privilege graph: an extension to the typed access matrix model. In D. Gollmann (Ed.), *Proceedings of the Third European Symposium on Research in Computer Security* (LNCS 875, pp. 319-334).
- Diallo, M. H., Romero-Mariona, J., Sim, S. E., Alspaugh, T. A., & Richardson, D. J. (2006, June). A comparative evaluation of three approaches to specifying security requirements. In *Proceeding of the 12th International Working Conference on Requirements Engineering: Foundation for Software Quality*.
- Edge, K. S., Dalton, G. C., II, Raines, R. A., & Mills, R. F. (2006, October). Using attack and protection trees to analyze threats and defenses to homeland security. In *Proceedings of the IEEE Military Communications Conference* (pp. 953-959). Washington, DC: IEEE Computer Society.
- Fung, C., Chen, Y., Wang, X., Lee, J., Tarquini, R., Anderson, M., & Linger, R. (2005, October). Survivability analysis of distributed systems using attack tree methodology. In *Proceedings of the IEEE Military Communications Conference* (pp. 583-589). Washington, DC: IEEE Computer Society.
- Henniger, O., Aprville, L., Fuchs, A., Roudier, Y., Ruddle, A., & Weyl, B. (2009). Security requirements for automotive on-board networks. In *Proceedings of the 9th International Conference on Intelligent Transport Systems Telecommunications* (pp. 641-646). Washington, DC: IEEE Computer Society.
- Herley, C. (2009). So long, and no thanks for the externalities: the rational rejection of security advice by users. In *Proceedings of the Workshop on New Security Paradigms* (pp. 133-144). New York, NY: ACM.
- Jürgenson, A., & Willemson, J. (2008). Computing exact outcomes of multi-parameter attack trees. In R. Meersman & Z. Tari (Eds.), *Proceedings of the On The Move to Meaningful Internet Systems* (LNCS 5332, pp. 1036-1051).
- Kordy, B., Mauw, S., Melissen, M., & Schweitzer, P. (2010). Attack-defense trees and two-player binary zero-sum extensive form games are equivalent. In T. Alpcan, L. Buttyán, & J. S. Baras (Eds.), *Proceedings of the First International Conference on Decision and Game Theory for Security* (LNCS 6442, pp. 245-256).
- Kordy, B., Mauw, S., Radomirović, S., & Schweitzer, P. (2011a). Foundations of attack-defense trees. In P. Degano, S. Etalle, & J. Guttman (Eds.), *Proceedings of the 7th International Workshop on Formal Aspects of Security and Trust* (LNCS 6561, pp. 80-95).
- Kordy, B., Pouly, M., & Schweitzer, P. (2011b). Computational aspects of attack-defense trees. In P. Bouvry, M. A. Klopotek, F. Leprévost, M. Marciniak, A. Mykowiecka & H. Rybinski (Eds.), *Proceedings of the International Joint Conferences on Security & Intelligent Information Systems* (LNCS 7053, pp. 103-116).
- Li, X., Liu, R., Feng, Z., & He, K. (2009). Threat modeling-oriented attack path evaluating algorithm. *Transactions of Tianjin University*, 15(3), 162-167. doi:10.1007/s12209-009-0029-y
- Manikas, T. W., Thornton, M. A., & Feinstein, D. Y. (2011). Using multiple-valued logic decision diagrams to model system threat probabilities. In *Proceedings of the 41st IEEE International Symposium Multiple-Valued Logic* (pp. 263-267). Washington, DC: IEEE Computer Society.
- Mauw, S., & Oostdijk, M. (2005). Foundations of attack trees. In D. H. Won & S. Kim (Eds.), *Proceedings of the International Conference on Information Security and Cryptology* (LNCS 3935, pp. 186-198).
- Meland, P. H., Tøndel, I. A., & Jensen, J. (2010). Idea: Reusability of threat models - two approaches with an experimental evaluation. In F. Massacci, D. Wallach, & N. Zannone (Eds.), *Proceedings of the International Symposium on Engineering Secure Software and Systems* (LNCS 5965, pp. 114-122).
- Mirowski, L., Hartnett, J., & Williams, R. (2009). An RFID attacker behavior taxonomy. *IEEE Pervasive Computing / IEEE Computer Society [and] IEEE Communications Society*, 8, 79-84. doi:10.1109/MPRV.2009.68

- Moore, A. P., Ellison, R. J., & Linger, R. C. (2001). *Attack modeling for information security and survivability*. Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University.
- Opdahl, A. L., & Sindre, G. (2009). Experimental comparison of attack trees and misuse cases for security threat identification. *Information and Software Technology*, 51(5), 916–932. doi:10.1016/j.infsof.2008.05.013
- Piètre-Cambacédès, L., & Bouissou, M. (2010). Beyond attack trees: Dynamic security modeling with Boolean Logic Driven Markov Processes (BDMP). In *Proceedings of the European Dependable Computing Conference* (pp. 199-208). Washington, DC: IEEE Computer Society.
- Roy, R., Kim, D. S., & Trivedi, K. S. (2011). Attack countermeasure trees (ACT): towards unifying the constructs of attack and defense trees. *Security and Communication Networks*.
- Saini, V., Duan, Q., & Paruchuri, V. (2008). Threat modeling using attack trees. *Journal of Computing Sciences in Colleges*, 23(4), 124–131.
- Schneider, B. (1999). Attack trees. *Dr. Dobb's Journal of Software Tools*, 24(12), 21–29.
- SHIELDS. (2008-2010). *FP7 project, grant agreement 215995*. Retrieved from <http://www.shields-project.eu/>
- Tanu, E., & Arreymbi, J. (2010). An examination of the security implications of the supervisory control and data acquisition (SCADA) system in a mobile networked environment: An augmented vulnerability tree approach. In *Proceedings of the 5th Annual Conference on Advances in Computing and Technology* (pp. 228-242).
- Tøndel, I. A., Jensen, J., & Røstad, J. (2010). Combining misuse cases with attack trees and security activity models. In *Proceedings of the International Conference on Availability, Reliability and Security*, Krakow, Poland (pp. 438-445). Washington, DC: IEEE Computer Society.
- Vesely, W. E., Goldberg, F. F., Roberts, N. H., & Haasl, D. F. (1981). *Fault tree handbook* (Tech. Rep. No. NUREG-0492). Washington, DC: U.S. Regulatory Commission.
- Wang, J., Whitley, J. N., Phan, R. C.-W., & Parish, D. J. (2011). Unified parametrizable attack tree. *International Journal for Information Security Research*, 1(1), 20–26.
- Yager, R. R. (2006). OWA trees and their role in security modeling using attack trees. *Information Sciences*, 176(20), 2933–2959. doi:10.1016/j.ins.2005.08.004

*Alessandra Bagnato is a research scientist and project manager within the Corporate Research Division of TXT e-solutions. She holds an MSc in computer science from the University of Genoa, Italy. She has worked in numerous EU projects related to software/service development and security, and she was also the dissemination and exploitation manager of the EU project SHIELDS. Her research interests include secure software development as well as security and privacy issues, model driven engineering, model driven modernization of complex systems, model-based methods and tools for embedded systems development.*

*Barbara Kordy is a postdoctoral researcher at the University of Luxembourg. She is currently leading the ATREES project which focuses on modeling and evaluating vulnerability scenarios using attack-defense tree methodology. Her research interests include formal methods for security assessment, privacy and anonymity issues and analysis of security protocols. She received her PhD and master degree in computer science from the University of Orléans in France. She also holds a master degree in mathematics obtained at the University of Silesia in Poland.*

*Per Håkon Meland graduated from NTNU in 2002, and has since been employed as a research scientist at SINTEF ICT in Trondheim. His research interests include software security and service engineering within domains such as health care and telecommunication, and with a special focus on early security awareness and improvements during the software development life cycle.*

*Patrick Schweitzer is a PhD student at the University of Luxembourg. Since 2009 he has been working in the ATREES project led by Dr. Barbara Kordy. His main research interests lie in formal security modeling, especially in models involving graphs, such as attack trees, defense trees and attack-defense trees. He is also researching various aspects of combinatorial graph theory, including the development of new graph class characterizations used to resolve standing conjectures. Patrick Schweitzer received a master's degree in mathematics from the Technical University of Berlin in 2008 and a master's degree in industrial engineering, also from the Technical University of Berlin in 2009.*



**B: ‘Mitigating risk with cyberinsurance’**

©2015 IEEE. Reprinted, with permission, from Per Håkon Meland, Inger Anne Tøndel and Bjørnar Solhaug, Mitigating Risk with Cyberinsurance, IEEE Security & Privacy, November/December 2015.

Included here is the accepted article version, which has been revised by the authors to incorporate review suggestions, and which has been accepted by IEEE for publication. The final, published version is the reviewed and accepted article, with copy-editing, proofreading and formatting added by IEEE. This is the version that appears in IEEE Xplore and can be found following the reference [18].

## Mitigating risk with cyber insurance

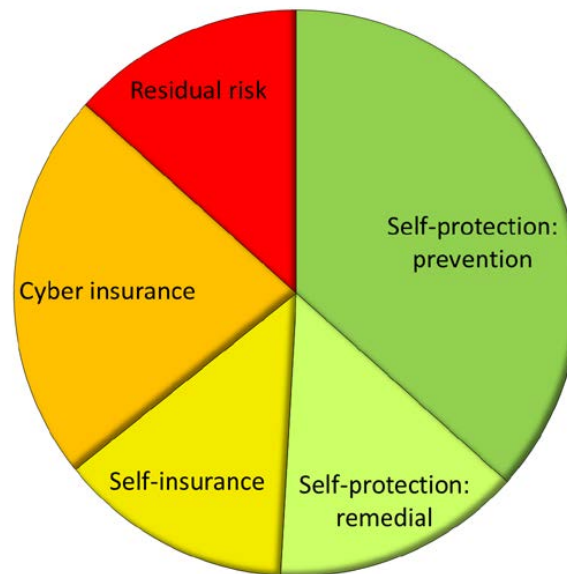
Per Håkon Meland, Inger Anne Tøndel and Bjørnar Solhaug  
SINTEF ICT

Information security is not merely a technical discipline, but something that must be considered based on economic incentives as well; how much and what kind of security pays off? A business can decide to drop proven security controls if the financial penalty of an attack is cheaper, but knowing this in advance is usually very hard. In other cases, vast amounts of money are spent on security that is strictly not necessary. Even if a business manages to get a proper balance between security investments and acceptable loss, there will always be some residual risk that is not well understood due to uncertainty and unpredictability. Such uncertainty may, for instance, be related to security knowledge in general, future proofing of defenses, attacker motivation, and value of intangible assets. To address such risks, businesses are increasingly buying cyber insurances, which is a cost saving, but still immature strategy of managing risk.

### 1 What's the deal here?

Information and Communication Technology (ICT) has brought a lot of benefits to businesses and the society. Nevertheless, as dependence on ICT has grown, the cyber threats towards these systems have become more prominent. Today's businesses should all expect to suffer from a cyber incident sooner than later.

Businesses can take preventive measures in order to reduce the risk of cyber attacks, but it is not economically feasible to fully protect all systems; the goal is rather to make them secure *enough*. Every year we are spending an increased amount of money on the global ICT security budget, studies from 2013 estimated around 50 billion USD, but the number of breaches is increasing by 20% yearly, and the costs of those breaches increase by 30% [1]. Most businesses do not have the economic backbone to compete in this race. Thus, they need to take on a mixed approach to risk management, including preventive and remedial actions, and self-insurance in the form of setting aside funds. Cyber insurance offers an additional type of risk management strategy, and has been defined in the literature as “*the transfer of financial risk associated with network and computer incidents to a third party*” [2]. A cyber insurance goes beyond traditional business interruption and crime insurances (though the borderline is vague), and can for instance cover liability issues, property loss and theft, data damage, loss of income from network outage and computer failures or web-site defacement [3]. Figure 1 shows an example of how risks can be distributed, with a remaining red section for risks that cannot be properly managed, and that businesses therefore just have to live with. Achieving an economically optimal mix of risk management strategies is the real challenge here.



**Figure 1. Risk management strategies.**

The task of dealing with cyber risks is commonly transferred to the ICT department. As a result, cyber security is often considered to be primarily a technical issue. To appeal to business leaders, cyber security needs to be translated into numbers and objective information that can underlie strategic management decisions on how much to invest and on which types of measures. Risk analysis can provide such input, as well as metrics such as *Annualized Loss Expectancy*, *Return on Investment* and *Total Cost of Ownership*. However, the data underlying risk analysis and the calculations of such metrics are in many cases sparse or with a high degree of uncertainty. Larger businesses usually have internal resources to work seriously on establishing a proper decision basis, and work strategically with cyber security. Decisions to buy cyber insurance policies can then be made based on cost-benefit trade-offs, and as a way of mitigating catastrophic events. Smaller businesses will usually have more difficulties making proper trade-offs, but with much lower premiums, cyber insurances are still attractive. In contrast to many beliefs, about 72% of all cyber breaches occur at small-to-medium sized businesses [4].

Cyber insurance is a relatively new product, but it is actually the fastest growing niche insurance product in the U.S., and gaining a lot of attention in Europe as well. For the insurance industry, this represents a large market, and insurance providers are jumping into it because their customers are actually demanding it. Ty Sagalow, president of an insurance consulting group, recently stated that "*insurers can't afford not to be in this thing*" [5].

## 2 What's so tricky?

Due to the still immature nature of cyber insurance, a series of challenges arise for both the insurer and the insured. Cyber security is a new domain for the insurance business, and without technical know-how or actuarial data, insurers do not really know what to require from their customers. Few breaches are publicly reported, and attacks may even go unnoticed. Even when available, historical information on attacks quickly becomes outdated due to the rapid technological development of both attack and defence techniques [6].

On the other side of the table, businesses struggle to implement and document security best practices that would eventually give them a better premium. Policies are supposed to close this gap, but to quote Selena Linde, who practices insurance law, “*cyber policies are still the Wild West of insurance policies*” [7], missing a standardized form, content and vocabulary. For the insured, it is no cakewalk to compare offers, and to grasp what is really covered. To give an example, *ransomcrypt* attacks, which encrypt files on the victims drive and demand a ransom to decrypt them, do not really fall under the category "theft", since the data has not come into possession of the attacker. Still, these kinds of attacks are clearly on the rise due to their simple nature and high profitability. Furthermore, data that accidentally goes into the wrong hands may not be considered a breach by the insurance policy, even though 29% of all data breaches stems from accidents (e.g. lost laptops, data sent to the wrong person, incorrect access control). Intentional hacking, on the other hand, contributes to 34% of the data breaches [8].

With traditional house insurance, you will have a hard time getting compensation if your house was already on fire when you signed the policy. Figure 2 illustrates how a software breach compares to a fire, taking the timeline into account. It typically takes a bit more than a year before a breach is discovered [7], but that does not mean that the incident is over. If we assume that the breach discovery is concurrent with the vulnerability discovery, we basically need to invent and produce fire extinguishers before the fire can be put out. In software terms this means waiting for a patch or creating a custom fix for homegrown code, followed by some additional time for deployment and testing. Studies by Symantec on the top-five zero day vulnerabilities show that average time between vulnerability publication and patch was 59 days [9]. A report [10] from WhiteHat Security states that the average time for fixing serious vulnerabilities in web applications is around 193 days. After this, a variable amount of time will go by before we see the real effect and damage of the breach, for instance misuse of stolen information. We are indeed looking at a very long timeline even for a single instance of a breach, which can easily cause disagreement on the validity of the policy. Ideally, the policy should of course have been signed before the breach, or even before an exploitable vulnerability was present. However, there has to be a general assumption that there exist such vulnerabilities in any ICT system at any time, and that there will be new ones introduced from time to time as the system evolves. The known or unknown presence of vulnerabilities will affect the policy coverage and premium, while attacks and their effects are related to insurance claims. With long-tail effects for years to come after a breach, the insured will be in a difficult position for re-negotiation before the incident lifecycle has ended.

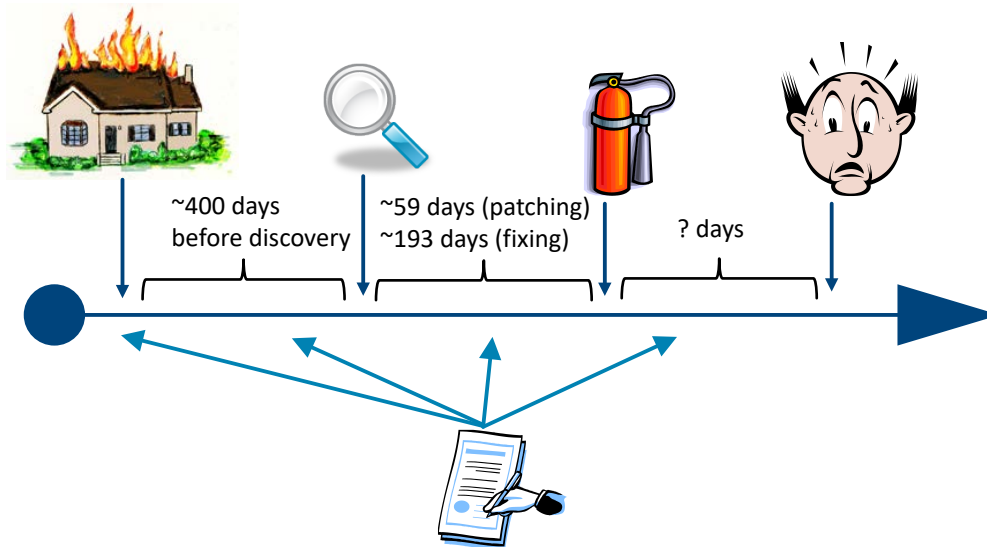


Figure 2. The signing date of the policy is critical when considering the timeline of a security incident.

If we consider last year's mega vulnerability affecting OpenSSL, the Heartbleed bug, we know that it was introduced already in 2012. It was first reported on April 3<sup>rd</sup> 2014 by the security teams of Codenomicon and Google Security. The news of it was made public on the 7<sup>th</sup> of April along with a fixed version of OpenSSL, but a few central service providers had been given an early warning a couple of days in advance. However, rolling out patches and renewing certificates is a process that is still on-going, and there are still a lot of unpatched web servers out there that can be exploited. To what extent this vulnerability has been abused is uncertain. Nothing abnormal is recorded in the event of an exploit, but there is no doubt that economic consequences have been severe considering all the remedial actions that are needed.

It is important to remember that most attacks occur after vulnerabilities have been discovered and made public. This brings us into another set of challenges related to cyber insurance; *when do you tell your insurance company about the flaw and what do you tell.* Leaking information about vulnerabilities or a breach could easily expose the system to further attacks or lead to exploits of stolen data, so it is important to preserve confidentiality when notifying the insurance company and the affected parties. Studies have shown that the average time taken by hackers to exploit a vulnerability is smaller than that taken by the vendor to issue a fix [11].

In contrast to traditional property or health insurance, computer systems tend to have a much more distributed and complex nature. For instance, there can be service chains where a software application that is used in Asia, may be provided by a European company, installed on a virtual platform from a company in the US, which again is deployed on physical servers in Australia. Composite software services are horizontally puzzled together from service components provided by numerous service providers, who may not even be aware of each other. Cyber attacks on a service component on any level of the service chain can therefore affect many other parties, and insurance lawyers can have a hard time figuring out liability, coverage and jurisdictions.

### 3 There's work to be done!

Which cyber risks a business is exposed to depends on a number of factors, external as well as internal; the threat level, the security of the code, security technology, security culture in the organization, policies and procedures for information security, internal resources on information security, the types of assets relevant, vulnerabilities, what risk is considered acceptable, compliance with legislative requirements, etc. These factors may influence the ability to protect the organization's assets when it comes to cyber threats, as well as the ability to react properly in case there is a breach, and then take actions that minimize the consequences of the breach. This is important with respect to well-known types of incidents, as well as the new and unexpected.

For businesses considering buying insurance, the basic questions to answer are:

- What is our risk?
- Which security measures in place should we tell the insurance company about?
- Does the insurance really cover what we need to survive?
- Is the price and terms right for us?

For insurers wanting to offer cyber insurance products, relevant questions are:

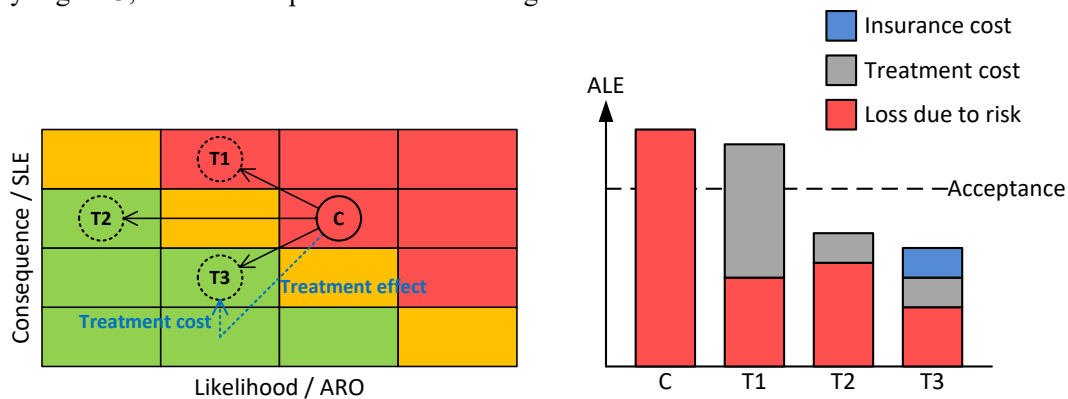
- Which factors impact the type of risk we are covering?
- How can we estimate the influence of these factors on the risk?
- How can we get hold of measurements for these factors, measurements that are of acceptable quality and that we can trust?
- What should we require from our customers?

In order to better address these questions, there is a lot of research ground to cover. Many open topics have already been mentioned in the literature, for instance in the studies from George Washington University [12] and ENISA [6]. Suggested topics range from studies on regional variations on attitude and uptake, collecting and analyzing loss data in a global database, quantification of risk related to technology, the role of legal frameworks and governments, and more. We would like to highlight some additional areas we are currently working on, and that we think will help making cyber insurance more mature. These are tied to risk assessment, collaboration between the insurer and insured, cost estimations and measurements of information security.

Security risk assessment can, and perhaps should, serve as a tool for increasing the understanding of the economic aspects of security and risk. Risk can be defined as the combination of the consequence of an incident and the likelihood of its occurrence. The consequence is usually estimated in terms of harm to assets, such as the number of disclosed database entries or the impact on company reputation. In principle, any kind of harm can be mapped to cost in terms of monetary loss for the stakeholder in question. For risk assessments in which we are able to express all consequences in terms of monetary loss, the overall risk picture can be expressed in terms of annualized loss expectancy (ALE). The ALE is simply the product of the single loss expectancy (SLE) of an incident and the annual rate of occurrence (ARO), where the SLE is the risk consequence and the ARO is its likelihood in terms of frequency per year. For deciding whether or not to buy a cyber insurance, and to what cost, security risk assessment methods and techniques that facilitate monetary cost assessment is a necessity. Such methods and techniques could also be an important tool for insurance providers in determining the adequate premium. While simple in theory, mapping security risks to ALE is not a trivial task as it can be very hard to understand or estimate the

economic implications of security incidents. Several security standards stress the need for understanding risk in terms of cost, e.g. ISO/IEC 27005:2008 and NIST SP 800-30, but there is still a lack of established and efficient techniques for how to do this in practice. This is especially true for costs affecting intangible assets such as reputation, which is vital for most businesses.

From a cyber insurance and economic perspective, we can leverage security risk assessment also for increasing our understanding of cost in relation to security. Basically, a security risk assessment that identifies unacceptable risks indicates that the system under assessment is not sufficiently secure. The risk treatment activity that follows a security risk assessment aims to identify options for mitigating the unacceptable risks and ensure a sufficient level of security. In principle, the selection of the optimal option for risk treatment can be done by means of a cost-benefit analysis. This involves the estimation of the cost of each treatment, as well as its effect in terms of risk reduction. In previous work [13], we have developed support for modeling and reasoning about treatment cost and benefit where risk, treatment cost and benefit are all represented in terms of annualized cost and benefit. Our approach is illustrated by Figure 3, which we explain in the following.



**Figure 3: Left: Risk matrix with treatment cost and effect. Right: Cost-benefit of treatment options.**

To the left in Figure 3 we have used the risk matrix to document the current risk (C) and the effect of three different options for risk treatment (T1, T2 and T3). Importantly, we also include here the treatment cost to explicitly show the total cost and not only the risk consequence. The importance for a cost-benefit assessment is perhaps clearer when comparing options T1 and T2 to the right in the figure. T1 is preferred over T2 when considering only the treatment effect, but the cost is obviously not acceptable. By considering risk and treatment cost together and in terms of ALE, as illustrated by the risk matrix where T1 is in the categorized as unacceptable from a cost-benefit perspective, we get a more appropriate basis for decision making based on security cost. As illustrated to the right in Figure 3, we can moreover use our approach to include cyber insurance in the cost-benefit analysis. In the example, we see that option T3 is preferred over T2, even with the additional cost of buying a cyber insurance. Our approach comes with support for risk and treatment modeling, estimation of risk level, estimation of treatment cost and treatment effect, as well as a calculus for accumulating the estimates and conducting the cost-benefit assessment. We have moreover developed means for graphical representation of the estimates and the analysis results similar to the illustration in Figure 3. We believe that for organizations to make well-founded decisions regarding security risk mitigation, including cyber insurance, both decision makers and security staff need to increase the awareness of cost in relation to security and

risk. We have developed support for reasoning about and for visualizing cost, but an important challenge that still remains is how to measure and estimate the cost of security and risk in an efficient way.

The benefits of cyber insurance are more than getting a claim pay-out in case there is breach. It is in the interest of the insurance companies that the pay-outs are as few and small as possible. Buying a policy thus also implies a new ally, and to some extent the access to the resources of this ally: *"Insurance companies will have on-staff and outsourced resources such as lawyers to help fight class-action lawsuits, security people to help advise about protections before breaches and incident response after breaches, and credit monitoring services to help consumers after a breach"* [7]. For smaller businesses with less cyber security resources inhouse, buying cyber insurance is a great way to increase the access to security competence. In addition, the very process of buying insurance, being reviewed by the insurance company and aiming to meet requirements in order to get premium discounts, can increase security awareness and lead to new or improved measures in order to reduce the company risk. This, however, depends on the quality of the underwriting process, and the indicators used by the insurer [12].

We also argue that when considering cost and likelihood, we need more knowledge on both the defender and attacker side of the trenches. In Figure 4 we have considered various types of costs and how they relate to each other. The defender side needs to set aside a budget for implementing preventive actions (a), such as building-security-in and hardening the infrastructure. Remedial actions (b) include activities such as fixing security problems, and are in general more costly for incidents caused by vulnerabilities that should have been eliminated at an earlier stage (a). The expected loss (c) should be much larger than the sum of the preventive and remedial actions, otherwise a business would not care about implementing security at all. For the same reason, implementing some *basic* security (a+b) must be cheaper than simply buying insurance (d) instead. The insurance pay-out (e) should not exceed the actual loss, since businesses should not profit by attacks to their own system. Self-insurance (not included in the figure) is a type of cost a business will have to reserve up-front to cover potential settlements, fees and remedial actions (b), but not lost as long there are no incidents.

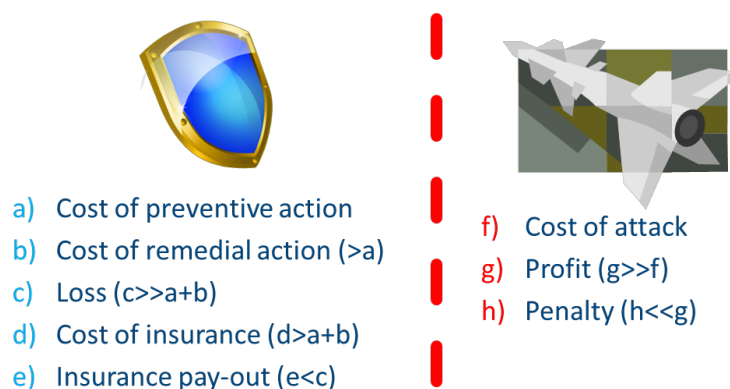


Figure 4. Costs on both the defender and attacker side.

On the attacker side, we should maintain information about the cost perspective in order to estimate likelihood of attacks. Firstly, we need to have an up-to-date overview of the cost of actually performing attacks (f), since launching a cheap attack is more probable than an



expensive one. Secondly, we have to assume that it is not very likely that an attacker will spend money on attacks without any substantial form of profit (g) compared to the cost of attack (f). Finally, the attacker is not likely to perform attacks where the penalty is very tough compared to the profit. In our previous work [14], we have developed attack-defence models that includes cost and likelihood, and can be used as a tool for cyber insurance. We suggest to model breach incidents using this approach to record historical data of the type shown in Figure 4. Note that defenders and attackers in many cases will have different views on what cost and profit are. Attacks may stem from motivations that are not financially motivated at all, such as political or religious actions, personal revenge or just plain fun. On the other hand, we believe that the lucrative underground economy is one of the main reasons for increased cybercrime. This advocates for a belief that attackers are financially aware and are making calculated risks before their actions.

Research and practice on measuring information security has progressed, and there are many indicators and measurement frameworks that can be used to gain insight into the state of information security. Still, we need to overcome the following challenges in order to provide useful measurements for cyber insurance:

- **Lack of standardized sets of indicators:** There is no agreed upon set of metrics that are considered most important to predict information security risk in the general case. Thus, businesses and insurance companies may need to make individual considerations regarding risk.
- **Risk is constantly changing:** Threats and technology, as well as the organization itself, its people, processes and relations, are not static. As a consequence, evaluations of information security risks need to be performed regularly. But following up all clients this way can be costly.
- **Impact of the organization's maturity level:** Information security can be measured at several levels, and NIST SP 800-55 explains how the types of measurement that can be made depend on the organization's information security program maturity. Measurements relating to business impact are at the highest maturity level. But cyber insurance should also be available for organizations that have not reached that high a maturity level on information security, and this must be taken into account in the underwriting process applied.
- **Measurement effect:** Any measurement alters the state of what is measured to some extent. The insured will likely change their self-protection strategies to be more in line with the indicators used by the insurer. This is also desired from the insurer's point of view. However, such a shift in focus will not increase the security if the insurers rely on the wrong indicators [12].
- **Trust in the underlying data:** Sharing of details related to the information security in organizations is often considered difficult due to the sensitivity of this information. In addition, the information that is shared may be of unknown quality – especially in cases where the organisation is not that mature in their information security work. Using third parties to perform evaluations of the cyber risk of an organization is a possible way to increase trust in the cyber security evaluations.
- **Ability to consider the unexpected:** Cyber insurance is likely to be relevant for risks that are low in probability but high in consequences. The organisation's resilience, i.e. the capability of recognizing, adapting to and coping with the unexpected, is thus important when considering this type of risk. Research from the area of safety identifies risk awareness, response capacity and support as key resilience attributes [15]. In particular, risk awareness can be difficult to measure quantitatively.

Considering all the open research topics we have ahead of us, it is unfeasible that we will overcome them without a closer collaboration between the cyber security and insurance community. That should be of mutual benefit and give us the opportunities to learn from each other's domains. More knowledge and further collaboration with authorities all over the world can help stop attacks even before they are launched. We have already seen promising trends for this, where governments are pushing insurance companies to participate in standardization work for cyber security, and hope this will continue in the future.

## References

- [1] S. Corner, "Billions spent on cyber security and much of it 'wasted'," *The Age*, 2014, <http://www.theage.com.au/it-pro/security-it/billions-spent-on-cyber-security-and-much-of-it-wasted-20140403-zqprb.html>.
- [2] R. Böhme, and G. Schwartz, "Modeling Cyber-Insurance: Towards a Unifying Framework," in *The Ninth Workshop on the Economics of Information Security (WEIS 2010)*, Harvard University, USA, 2010.
- [3] T. Bandyopadhyay, V. S. Mookerjee, and R. C. Rao, "Why IT managers don't go for cyber-insurance products," *Commun. ACM*, vol. 52, no. 11, pp. 68-73, 2009.
- [4] S. E. Needleman, "Cybercriminals Sniff Out Vulnerable Firms," *The Wall Street Journal*, 2012, <http://www.wsj.com/articles/SB10001424052702303933404577504790964060610>.
- [5] N. Perlroth, and E. A. Harris, "Cyberattack Insurance a Challenge for Business," *The New York Times*, 2014, <http://www.nytimes.com/2014/06/09/business/cyberattack-insurance-a-challenge-for-business.html>
- [6] N. Robinson, "Incentives and barriers of the cyber insurance market in Europe," 2012, [http://www.enisa.europa.eu/activities/Resilience-and-CIIP/national-cyber-security-strategies-ncsss/incentives-and-barriers-of-the-cyber-insurance-market-in-europe/at\\_download/fullReport](http://www.enisa.europa.eu/activities/Resilience-and-CIIP/national-cyber-security-strategies-ncsss/incentives-and-barriers-of-the-cyber-insurance-market-in-europe/at_download/fullReport).
- [7] E. Chickowski, "10 Things IT Probably Doesn't Know About Cyber Insurance," *InformationWeek*, 2014, <http://www.darkreading.com/operations/10-things-it-probably-doesnt-know-about-cyber-insurance/d/d-id/1316862>.
- [8] Symantec, "Internet Security Threat Report," 19, 2014, [http://www.symantec.com/content/en/us/enterprise/other\\_resources/b-istr\\_main\\_report\\_v19\\_21291018.en-us.pdf](http://www.symantec.com/content/en/us/enterprise/other_resources/b-istr_main_report_v19_21291018.en-us.pdf).
- [9] Symantec, "Internet Security Threat Report," 20, 2015, <http://know.symantec.com/LP=1123>.
- [10] WhiteHat, "The WhiteHat Website Security Statistics Report," 2013, [https://www.whitehatsec.com/assets/WPstatsReport\\_052013.pdf](https://www.whitehatsec.com/assets/WPstatsReport_052013.pdf).
- [11] M. Shahzad, M. Z. Shafiq, and A. X. Liu, "A large scale exploratory analysis of software vulnerability life cycles," in *Proceedings of the 34th International Conference on Software Engineering*, Zurich, Switzerland, 2012, pp. 771-781.
- [12] C. Toregas, and N. Zahn, *Insurance for Cyber Attacks: The Issue of Setting Premiums in Context*, The George Washington University 2014.
- [13] L. Tran, B. Solhaug, and K. Stølen, "An Approach to Select Cost-Effective Risk Countermeasures," *Data and Applications Security and Privacy XXVII*, Lecture Notes in Computer Science L. Wang and B. Shafiq, eds., pp. 266-273: Springer Berlin Heidelberg, 2013.
- [14] A. Bagnato *et al.*, "Attribute Decoration of Attack–Defense Trees," *International Journal of Secure Software Engineering (IJSSSE)*, vol. 3, no. 2, pp. 1-35, 2012.

- B**
- [15] K. Øien *et al.*, “Development of early warning indicators based on Resilience Engineering,” in PSAM10, International Probabilistic Safety Assessment and Management Conference, 2010, pp. 7-11.

## **C: ‘Visualizing cyber security risks with bow-tie diagrams’**

A license to reproduce the published material [19] for inclusion in this thesis has been obtained from Springer Nature.

C

# Visualizing Cyber Security Risks with Bow-Tie Diagrams

Karin Bernsmed<sup>1</sup>, Christian Frøystad<sup>1</sup>, Per Håkon Meland<sup>1,3</sup>, Dag Atle Nesheim<sup>2</sup>, and Ørnulf Jan Rødseth<sup>2</sup>

<sup>1</sup> SINTEF Digital

{karin.bernsmed,christian.froystad,per.h.meland}@sintef.no

<sup>2</sup> SINTEF Ocean

{dag.atle.nesheim,ornulfjan.rodseth}@sintef.no

<sup>3</sup> Norwegian University of Science and Technology

per.hakon.meland@ntnu.no

C

**Abstract.** Safety and security risks are usually analyzed independently, by different people using different tools. Consequently, the system analyst may fail to realize cyber attacks as a contributing factor to safety impacts or, on the contrary, design overly secure systems that will compromise the performance of critical operations. This paper presents a methodology for visualizing and assessing security risks by means of bow-tie diagrams, which are commonly used within safety assessments. We outline how malicious activities, random failures, security countermeasures and safety barriers can be visualized using a common graphical notation and propose a method for quantifying risks based on threat likelihood and consequence severity. The methodology is demonstrated using a case study from maritime communication. Our main conclusion is that adding security concepts to the bow-ties is a promising approach, since this is a notation that high-risk industries are already familiar with. However, their advantage as easy-to-grasp visual models should be maintained, hence complexity needs to be kept low.

**Keywords:** security, safety, risk assessment, bow-tie diagrams, maritime communication

## 1 Introduction

One of the least understood challenges for cyber physical systems (CFS) is uncertainty in the environment, cyber attacks and errors in connected physical devices [46]. The tight coupling between the cyber and physical world leads to new forms of risks that have not been considered adequately, such that the cyber element adversely affects the physical environment [4]. Safety risks, where the system can harm the environment in which it operates, and security risks, where the environment (e.g. malicious actors and other systems) can harm the system, tend to be analyzed independently [42], by different people using different standards, tools and notations. As pointed out by Sun et al. [50], safety and security

goals interact synergistically or conflictingly, and should therefore be evaluated together. If not, conflicts can result in either (a) overly secure systems that compromise the reliability of critical operations or (b) create insecure systems where back-doors are easily found.

An inherent challenge when combining safety and security in an analysis is the increased complexity. Graphical visualizations are helpful when you want to make complex problems easier to understand and navigate [20]. The purpose of this paper is to bridge the gap between safety and security during risk assessment by utilizing the graphical bow-tie diagram methodology [14, 11, 15, 25]. Bow-tie diagrams are very suitable for communicating the results of a risk assessment to different stakeholders within an organization due to the clear diversification of causes and effects for a given unwanted event, and to clarify which barriers have (or have not) been implemented. Bow-tie analysis, which includes the generation of one or more bow-tie diagrams, is a common approach to map the risks associated with unwanted events in, for example, the oil and gas industry. Our approach is to take advantage of the familiarity of this graphical notation among industry experts, analyze use cases within the safety-critical maritime sector, and try to answer the following research questions:

1. How can bow-tie diagrams be extended to include security considerations in addition to safety considerations?
2. How can the likelihood of cause and severity of cyber attacks be visualized in bow-tie diagrams?

In order to answer these questions, we apply a *design science* research methodology [48], with focus on the extended bow-tie diagram methodology as an artefact with a high priority on relevance for the cyber physical domain. Evaluation is done through analysis of descriptive, constructed use cases for maritime service scenarios to demonstrate its utility [21].

Our goal has not been to create yet another theoretical model for risk assessment, but to propose a solution to a real, existing problem we experience in the maritime domain when introducing new technology that may have effect both safety and security. This follows the research paradigm of *pragmatism* [19], which is associated with action, intervention and constructive knowledge. Furthermore, it should be based on real problems and have practical usefulness beyond the specific case studies.

This paper is organized as follows. Section 2 presents related work. In Section 3, we introduce the marine communication case study in which we have developed the proposed methodology. Section 4 explains the concepts and terminology that we use and Section 5 presents the proposed bow-tie risk assessment methodology, which is exemplified in Section 6. Finally, in Section 7 we discuss the results and Section 8 concludes the paper.

## 2 Related work

The most common way of documenting and visualizing risks is in a risk matrix, where the seriousness of the evaluated risks can be easily compared based on

the combination of likelihood and consequence. The US Air Force developed the Risk Matrix Approach (RMA) [18] in 1995, and after that it has spread out to a multitude of domains, such as weapons manufacturing, finance, transport and project management [38]. Still, RMA is a very simplistic notation that does not properly visualize the causes of the risks, and how to address them.

Within the field of security, there are many more specialized modelling notations that are in general concerned about “*identifying system behaviour, including any security defenses; the system adversary’s power; and the properties that constitute system security*” [5]. Security modelling comes in many different forms and flavors, but they all share the common aim of understanding security issues so they can be dealt with effectively. Which one to choose usually depends on what the analyst wants to focus on, level of abstraction/details and personal preference (e.g. familiarity). To quote Shostack [47]: “*different diagrams will help in different circumstances*”. For instance, an attack tree [45, 31] is a tree-based notation showing how an adversary can choose among different paths or branches to obtain an overall attack goal. The attack-defense trees [26] extend this notation by also adding preventive nodes, which again can be attacked by attack nodes. Attack graphs [40] and vulnerability cause graphs [8] are examples of a graph-based notation used for analyzing vulnerabilities, and CORAS [30] contains several graphical notations for a risk analysis process. There also exist different types of security extensions to more general purpose graphical modelling notations, such as Data flow diagrams [47], UML [24, 49] and BPMN [32].

For safety, there are many notations that go even further back in history. The fault-tree analysis (FTA) method was developed in the 1960s for safety and reliability [29], and a recent survey of usage is provided by Ruijters and Stoelinga [43]. Event tree analysis (ETA) is an established technique originating from the nuclear industry [3], and is used to analyze how a series of events can lead to a potential accident scenario. Similarly to ETA, cause-consequence diagrams (CCA) [39] are also used to analyze safety causes.

When considering safety and security in combination, there have been quite a few related studies. For instance, Winther et al. [52] show how to handle security issues as part of HAZOP studies, which is a systematic analysis on how deviations from the design specifications in a system can arise, and whether these deviations can result in hazards. Raspotnig et al. [42] have use UML-based models within a combined safety and security assessment process to elicitate requirements. Bieber and Brunel [7] show how common system models for security and safety can be used for airworthiness certification within aviation. Kumar and Stoelinga [28] have married fault and attack trees so that both safety and security can be considered in combination. Further examples of methods, models, tools and techniques in the intersection of safety and security can be found in the surveys by Zalewski et al. [53], Piètre-Cambacédès and Bouissou [41], Chockalingam et al. [12], as well as Kriaa et al. [27].

There have been several efforts by practitioners related to the use of bow-tie diagrams for cyber security, but they differ from what we are presenting in this paper in several ways. For instance, a report from SANS Institute [35] outlines



how a bow-tie risk assessment methodology can be applied to conduct a cyber security risk assessment in an engineering environment. There is no change to the diagram notation as such, but they argue that *“the first step towards obtaining Engineering community buy-in”* is to compare concepts from security to bow-tie, and basically evaluate cyber threats in the same manner as hazards. They also include considerations related to actors and motivation, but this is done in order to reduce the number of possible scenarios before modelling, and not part of the notation itself. A report from DNV-GL [16] also proposes the use of bow-tie diagrams as a key component in a cyber security assessment program for the maritime sector. Here, standard safety notation is used, and the focus is on visualization of barriers. Quantitative indicators are explicitly left out, and even though vulnerability consideration is central in the overall assessment process, this is not included as diagram concepts. Similarly, the *Bow Tie for Cyber Security* series [22] at PI Square gives numerous examples where the standard notation is used for security. The US Coastguard has also published a report [34] on how to use bow-ties to identify preventive and responsive responses to cyber attacks for marine transportation systems. Their examples are on a very high abstraction level, where causes are for instance *hactivists*, *technical errors* and *insider threats*. Two additional examples of bow-tie diagrams that visualize IT security risks are provided in [10]. The focus here is more on chains of barriers, although it seems like vulnerabilities are represented as escalation factors.

### 3 Case study: Maritime communication

In order to give a better understanding of the methodology and examples used in the later sections, we would like to explain our maritime case study and why security is a growing concern intertwined with safety in this domain.

Shipping has become increasingly dependent on digital data exchanges. As dependence grows and the functions supported becomes more entangled in the ship operations and critical interactions with on-shore authorities, the need to consider consequences of digital attacks on the data exchanges also increases. This calls for a more systematic approach to maritime cyber security.

In 2011, ENISA pointed out [13] that the *“awareness on cyber security needs and challenges in the maritime sector is currently low to non-existent”*. Come 2015, the Lysne commission of Norway [2] reaffirmed this message. The lack of general awareness regarding cyber security, makes the industry more vulnerable to attacks.

Maritime navigational systems of today rely heavily on Global Navigation Satellite Systems (GNSS), such as GPS and GLONASS, to navigate safely, avoid collisions or groundings and for voyage optimization. The GNSS signals available for civilians are unencrypted and unauthenticated and are easily jammed or even spoofed [6]. Automatic Identification System (AIS) is used to identify other ships and their intentions, but can also be used to transmit short safety messages, e.g. to act as virtual aids to navigations. AIS is becoming part of the more extensive VHF Data Exchange System (VDES), which will extend the use

of AIS to include even more digital information exchanges. The AIS messages are unencrypted and unauthenticated, and relatively easy to jam or spoof. Furthermore, IOActive [44] conducted tests on SATCOM firmware from multiple vendors and found vulnerabilities such as hardcoded credentials, undocumented protocols, insecure protocols, backdoors, and weak password reset. Our attention is on digital data exchanges between ships and between ship and shore and the possible consequences of cyber-attacks on these exchanges.

Ships spend most of their time at sea with a minimal crew, and remote monitoring and maintenance is becoming more and more common. If not organized in an appropriate way, this could allow an attacker extensive and easy access to the systems on the ship. Additionally, there are multiple actors connected to the network on-board a ship, including passengers, crew, and operational systems. These actors have different requirements regarding safety, security and separation. For instance, some vessels have physically separated networks, while others only provide logically separated networks. The mechanisms for logical separation of networks vary, but are often just a simple firewall.

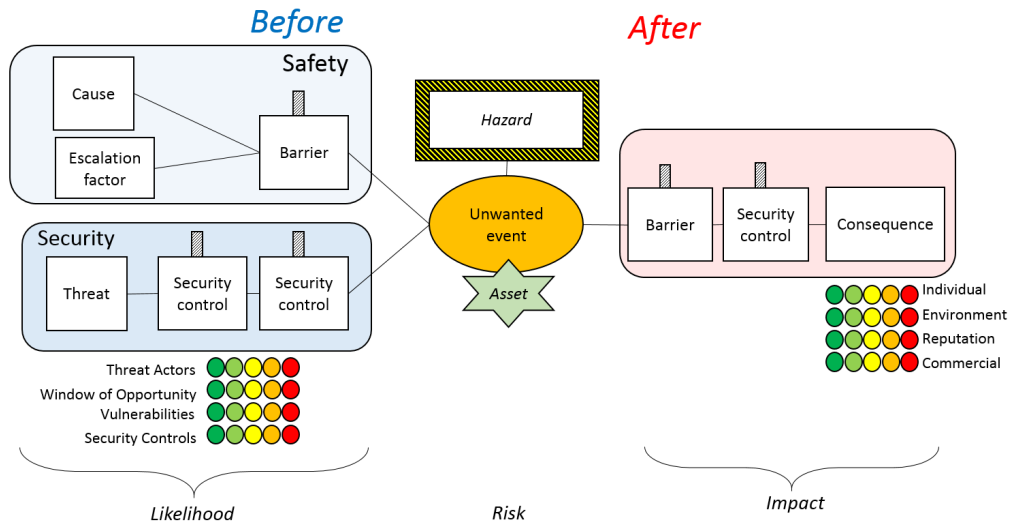
## 4 Concepts and terminology of bow-ties

A bow-tie diagram is shaped like a bow-tie, where the central *knot* typically represent an accident scenario, or as we will later refer to, an unwanted event. The diagram can be seen as a combination of a fault tree and an event tree [17], where the left side shows which causes can lead up to the accident, and the right side the potential effects once the accident has occurred. As pointed out by the tool provider CGE Risk Management<sup>4</sup>, the power of this diagram is that it gives a clear distinction between proactive and reactive risk management, in combination with an overview of multiple plausible scenarios.

To combine security with bow-tie safety assessment, we need to synchronize the terminology and concepts from the safety and security domains. The bow-tie diagram in Fig. 1 shows the traditional layout, notation and concepts from safety assessments in the upper left horizontal part (*cause, barrier, escalation factor*), with concepts we introduce from security in the lower left horizontal part (*threat, security control*). *Hazard* and *unwanted event* are mainly from safety, while *asset* comes from security. On the right side of the figure, the *consequence* concept is shared between safety and security, and can be remedied with safety barriers and security controls, often in combination. We describe these concepts further below.

As defined by International Maritime Organization (IMO) [23], the first step in a Formal Safety Assessment (FSA) [23] is to identify all potential hazards that can contribute to accidents. A *hazard* is a potential to threaten human life, health, property or the environment. Examples of maritime hazards are off-shore operations, hazardous substances and sources of ignition onboard and external hazards, such as storms, lightening and other ships. Hazards may give rise to

<sup>4</sup> <https://www.cgerisk.com/knowledge-base/risk-assessment/thebowtiemethod>



**Fig. 1.** Our combined approach for modelling safety and security in a bow-tie diagram.

scenarios in which people, the environment or property will be damaged. The list of identified hazards and their associated scenarios will be used as input to the safety risk assessment. Basically, a hazard can be anything with the potential to cause harm, but which is also necessary to perform business. From a risk analysis perspective, the hazard needs to be controlled so that unwanted events will not occur.

An *unwanted event* in safety assessment, also known as top event, loss event, or loss of control, represents what will happen if one loses control over a hazard, which again can have severe *consequences*. An unwanted event is typically caused by an accident, or a random failure. In security assessments, the equivalent is often called *incident*, something that typically affects the confidentiality, integrity or availability of a critical system, data, or processes necessary for the operation of the business. Such incident may have malicious or accidental causes. In our model, we are using the term unwanted event for anything that can cause harm to the asset(s) associated to the hazard, regardless if they stem from safety or security causes. In real life, it is often a combination of different causes that lead to unwanted events, therefore we want to evaluate them together.

Related to security, an *asset* is anything that has value to an organization. The ISO/IEC 27005 standard [1] distinguishes between primary assets, which are core business processes and their corresponding information, whilst supporting assets are those required to be in place to support the activities of the primary assets. Typical examples of (primary) assets in a maritime context are Maritime Safety Information (MSI), ship certificates, and electronic nautical charts. Asset is not a concept that is used in traditional safety assessment, but is usually the first thing to identify when it comes to security assessments. Therefore, we

include a mapping between hazard and which assets will be damaged in case the unwanted event occurs.

A *threat* is anything that can potentially cause an unwanted event [1]. Within safety assessments, the term *cause* is very often used directly for the same meaning. A *barrier* is a mechanism that aims to interrupt causes of unwanted events, or that it is possible to recover from the unwanted event without severe consequences. In a security context, the term barrier corresponds to the term *control*, which is a means of managing risk, including policies, procedures, guidelines, practices or organizational structures, which can be of administrative, technical, management, or legal nature [1]. These can be preventive controls used to avoid, detect or mitigate threats, or reactive controls, which are intended to limit the damage caused by an incident. Note that in a security context, the word safeguard, mitigations, or countermeasure, are sometimes used as a synonym for control. An *escalation factor* is anything that may cause a safety barrier to fail. There is no one-to-one mapping between this concept and security terminology, however, to succeed with a threat, a threat actor will need to exploit one or more vulnerabilities, which often is only feasible at a certain point of time (window of opportunity).

In our model, we use threats to explicitly represent malicious activities, while causes are related to traditional safety accidents. We continue to use both barrier and security control for both sides of the bow-tie, though they may have the same implementation (e.g. through redundancy). Note that there can be chains of both barriers or security controls (the latter is illustrated in Fig. 1). Such chains follow the principle of *defence in depth* - if the first barrier fails or control is circumvented, there is another one still operating.

We also introduce a set of color coded indicators for each threat branch on the left side, and for each consequence branch on the right side of the diagram. These indicators are meant to help visualize the likelihood of an unwanted event, and the severity of a consequence in similar manner that is used for risk matrices. This allows us to adopt the RMA framework as described in Section 2 as apart of the notation, and make use of the color indicators that the industry community is already familiar with. For a threat branch, we associate indicators related to *threat actors*, *window of opportunity*, *vulnerabilities* and *security controls*. For instance, the threat actors indicator informs whether or not it is likely that there exists groups or individuals who have the competence, resources and motivation necessary to perform an attack and instantiate the threat. Similarly, we indicate the likely existence of the other indicators. For a consequence branch, the indicators represent the severity of the impact related to *individuals*, the *environment*, the *reputation* of a company and *commercial* (monetary) loss.

In the next section, we focus on how to identify what color should be used for each indicator, and how to quantify the overall risk of a bow-tie diagram for an unwanted event.

## 5 Risk assessment

As illustrated in Fig. 1, the risk of an unwanted event will be a combination of the likelihood and the impact of the unwanted event. Our contribution in this paper focuses on a subset of all potential unwanted events, which are those caused by hostile cyber attacks. In our model, an unwanted event  $U$  will be a function of one or more threats. Each unwanted event will lead to one or more consequences  $C$ , where each identified consequence is associated with a corresponding impact (i.e. severity, or loss,) value  $L$ . The risk  $R$  associated with a certain unwanted event  $U$ , which we denote  $R(U)$ , will then be approximated as the probability that the unwanted event occurs, i.e.  $p(U)$ , multiplied with the worst-case consequence impact value that has been identified, which we denote  $L_C$ , and the likelihood that this consequence occurs, i.e.  $p(C)$ . The formal expression for this is

$$R(U) \approx p(U) \times L_C \times p(C) \quad (1)$$

To quantify the risk of an unwanted event, we hence need to assess 1) the probability of the unwanted event (as a function of one or more identified threats) and 2) the impact value and probability of the worst-case consequence of the unwanted event.

### 5.1 Assessing the left side of the bow-tie (cause)

Assessing the probability of a cyber attack is a notoriously difficult problem. In our model, we assume that all the threats are *mutually independent*. This means that all the identified cyber attacks will be executed independently of each other and that any of them can manifest itself and cause the unwanted event during the time for which the system, or service, is being assessed. Under this assumption, the probability of the unwanted event  $U$  can be computed as

$$p(U) = p(\text{at least one } T_i \text{ occurs}) = 1 - \prod_{i=1}^n (1 - p(T_i)) \quad (2)$$

where  $p(T_i)$ ,  $i = 1 \dots n$ , is the probability of threat  $T_i$ . The problem will hence be reduced to assessing the probabilities, or likelihoods, of the individual threats that have been identified.

Compared to more simplistic probability models, in which the threats are modelled as mutually exclusive (i.e.  $p(U)$  will be computed as a sum of the individual threats), the proposed Equation 2 is much more realistic, since it allows more threats to manifest within the same time interval, which corresponds more closely to the real world. By using Equation 2, we can also model cases in which multiple attackers work simultaneously to exploit different vulnerabilities, and cases where one attacker exploits all the vulnerabilities he can find. However, the assumption that all the threats are independent may not always be true. In particular, it is questionable whether one can model scenarios in which an attacker is aware of all the potential threats that can be carried out, since this

may affect the probabilities of the individual threats, hence violating the independence assumption. Another issue may be that, for some unwanted events, once the unwanted event has happened, it will be less likely to happen again due to increased awareness. This is a common situation in a security context, where threats are manifesting themselves through the actions of human beings rather than through random failures, and the malicious actors will lose their *element of surprise*.

Another characteristic of Equation 2 is that the more threats one identifies, the higher the probability of the unwanted event. A side effect of using this model could therefore be that a more thorough risk assessor, who manages to identify more threats, will also end up with a higher probability of the unwanted event. However, the influence of the number of identified threats will be negligible, as long as both the threat probabilities and the number of identified threats are sufficiently small (which is the case in most real-life scenarios).

In our opinion, in spite of the aforementioned issues, this is the simplest and most straightforward alternative we have for computing the probability of an unwanted event  $p(U)$  as a function of the identified threats. This same model is frequently used in system reliability analysis, in which a system analyst models the system as a set of components, assesses the individual failure rates of the components and evaluates the effect of the total system reliability. In our case, we model malicious threats rather than random failures, however, the underlying line of thought is similar; we are considering multiple sources of error that can cause the system, or service, to fail, regardless of cause. Note that, when using this approach, care must be taken to ensure that all the identified threats are independent and, as explained above, the risk assessor must understand the characteristics of the underlying mathematical model.






**Assessing the threat actors, window of opportunity, vulnerabilities and security countermeasures.** We move on to describe how factors, such as the actors who pose the threat, the needed window of opportunity for the threat to be successful and any vulnerabilities and security countermeasures present in the system can be assessed and visualized. As explained in Section 4, we use color coded indicators to represent these factors in the graphical model.

*Threat actors* Threat actors are the attackers who will represent a security risk against the system that is being assessed. Threat actors can be classified in terms of characteristics, such as skill, capabilities, resources, intent and access [9]. The risk assessor can estimate the threat actors by using the values of Table 1.






*Window of Opportunity* The "window of opportunity" depends on how often/long the threat actor theoretically could gain access to the target (system or data) and how often/long the target of interest is within reach of the attacker. The risk assessor can estimate the window of opportunity by using Table 2.

*Vulnerabilities* No system is perfect, nor are the security measures that are put in place to prevent the threat from manifesting itself. Vulnerabilities can range

**Table 1.** Color coding for representing the threat actors






Threat actors		
Dangerousness	Description	Color coding
<i>Severe</i>	There are threat actors highly capable of pursuing this threat	
<i>High</i>	There are threat actors capable of pursuing this threat	
<i>Moderate</i>	There are threat actors somewhat capable of pursuing this threat	
<i>Low</i>	There are threat actors interested in pursuing this threat, but their capability is limited	
<i>None</i>	There are threat actors interested in pursuing this threat, but they are not capable of acting on this interest	

**Table 2.** Color coding for representing the window of opportunity

Window of opportunity		
Window	Description	Color coding
<i>Always</i>	This threat is always possible.	
<i>Frequent</i>	This threat is frequently possible (there will be an opportunity about once every week).	
<i>Rare</i>	This threat is rarely possible (there will be an opportunity about once every year).	
<i>Extremely rare</i>	This threat is extremely rarely possible (there will be an opportunity about once every 10th year).	
<i>Never</i>	This threat is never possible.	

from simple programming errors to large design flaws of software, hardware and processes. The presence of vulnerabilities increases the likelihood of a threat manifesting. The risk assessor can estimate the existence of vulnerabilities by using Table 3.

**Table 3.** Color coding for representing the presence of vulnerabilities

Vulnerabilities		
Vulnerability	Description	Color coding
<i>Known easy</i>	One or more known vulnerabilities exist, which are easy to exploit.	
<i>Known-difficult</i>	One or more known vulnerabilities exist, but they are either not publicly known, or they are difficult to exploit.	
<i>Unknown</i>	No known vulnerabilities exist, however, vulnerabilities are expected to appear in the near future.	
<i>Very unlikely</i>	It is very unlikely that the system has, or will have, any vulnerabilities in the near future.	
<i>Formally proven absence</i>	Formal methods, or the like, have been applied to demonstrate that no vulnerabilities exist. It is extremely unlikely that vulnerabilities will appear in the near future.	

*Security controls* Finally, the risk assessor will need to input information about the existence of security control and assess their effectiveness (Table 4).

**Assessing the threats** For each threat  $T_i$  and preventive security controls  $Ctrl_1 \dots Ctrl_m$ , the risk assessor choose values for *Threat Actors*, *Window of Opportunity*, *Vulnerabilities* and *Security Controls* according to Table 1, 2, 3 and 4. This is visualized as extended traffic lights as shown in Figure 2. In addition to the traffic lights, the relevant controls for each threat are shown as separate boxes to give an overview of which threats are mitigated by which controls.

The visualization in Fig. 2 serves as domain specific assistance to the risk assessor when assessing  $p(T_i)$ ,  $i = 1 \dots n$ , i.e. the probability of each of the identified threats. We do not dictate exactly how this estimation should be done in practice, as there are different ways of doing threat prediction, and any model depends a lot on the available information used as input. When working with maritime threat scenarios, we have been using averages from generic threat intelligence data, and then adjusted these based on the case specific domain data using expert opinions.

### 5.2 Assessing the right side of the bow-tie (consequence)

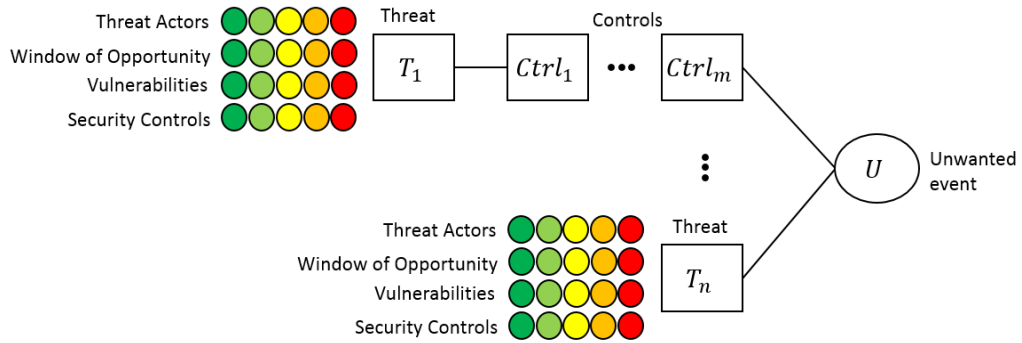
The consequence of an evaluated risk can manifest itself in many ways. FSA normally only consider individual risk and societal risk which represents the main





**Table 4.** Color coding for representing the effectiveness of security controls

Control	Description	Color coding
<i>Known to be ineffective</i>	No security countermeasure exists, or, one or more security countermeasures exists but they are known to be ineffective.	
<i>Probably not effective</i>	One or more security countermeasures exists but they can be circumvented.	
<i>Effective</i>	One or more security countermeasures exists, which are believed to be effective.	
<i>Very effective</i>	One or more security countermeasures exists, which are very effective.	
<i>Formally proven effective</i>	Formal methods, or the like, have been applied to demonstrate that existing security mechanisms are sufficient and work as intended.	



**Fig. 2.** The relation between an unwanted event, threats, threat actors, window of opportunity, vulnerabilities and (preventive) security controls

scope of the Maritime Safety Committee in IMO where the FSA was developed. We have found it useful to also include other aspects, such as the environmental (pollution), commercial (monetary losses) and reputational (loss of confidence by e.g. customers, business partners, bank, insurance, regulatory bodies) damage caused by each identified unwanted event in our model. As an example of reputational damage, the Paris MoU<sup>5</sup> publishes a black list for all ships depending on results from Port State Controls. Once your ship is on this list, you are much more eligible for inspections and your operation may suffer.



**Table 5.** Consequence type and severity level

Consequences					
Level	Individual	Environment	Reputation	Commercial	Color coding
<i>Catastrophic</i>	Multiple deaths	Uncontained release with potential for very large environmental impact	International coverage, unrecoverable damage	\$ 50 000 k	Red
<i>Critical</i>	One death	Uncontained release with potential for major environmental impact	National and some international coverage, impact lasting more than a year	\$ 5 000 k	Orange
<i>Moderate</i>	Multiple severe injuries	Uncontained release with potential for moderate environmental impact	National media coverage, impact lasting more than 3 months	\$ 500 k	Yellow
<i>Negligible</i>	One minor injury	On site release contained without external assistance	Local complaint/ recognition, impact less than one month	\$ 5 k	Light Green
<i>None</i>	No injuries	No effect	No damage	\$ 1 k or less	Green

The risk assessor can estimate the consequence of each identified unwanted event using Table 5. One obvious problem with comparing these different outcomes is to compare consequences for life and health with purely economic or environmental damages. However, it is possible to compare the economic consequences of a lost life or health damage to other more direct economic consequences of a cyber attack. Our approach is to follow this (semi-) quantitative assessment, and leave a more qualitative societal risk acceptance analysis to later stages.

Individual consequence represents the direct danger to life or health of persons on board the ship, on other ships or on shore. It does not include secondary effects due to, e.g. pollution or other factors. As noted above, it is not trivial

<sup>5</sup> <https://www.parismou.org/>

to assess the value of life and health in purely economic terms. The problem is, for instance, complicated by the different economic values assigned to lives in different parts of the world [51]. For example, this value was estimated to be at USD 0.8 million in South Korea in the year 2000, and at USD 9.7 million in Japan the same year. In our model, we will use the mean value of USD 5 million for one life as baseline. This represents the mean value from [51], but not weighted according to population in the different areas.

We follow the defined severity levels for economical loss as shown in Table 5. This maps critical to the above value corresponding to loss of one life and adjusts other levels accordingly.

The inclusion of reputational and economical loss in the risk assessment has been a matter of some discussion. Our rationale for doing this and not only focusing on individual and environmental risks, is that in many cases the motivation for and the consequences of a successful cyber-attack is likely to be much higher in the commercial domain than in the general safety domains. This assumption is strengthened by today's ship bridge operational regime where all received information must be checked against other sources of information, including making visual assessment of the ships situation. Thus, including commercial consequences will likely lead to more risks being assessed as not acceptable and by that lead to a higher overall safety level.

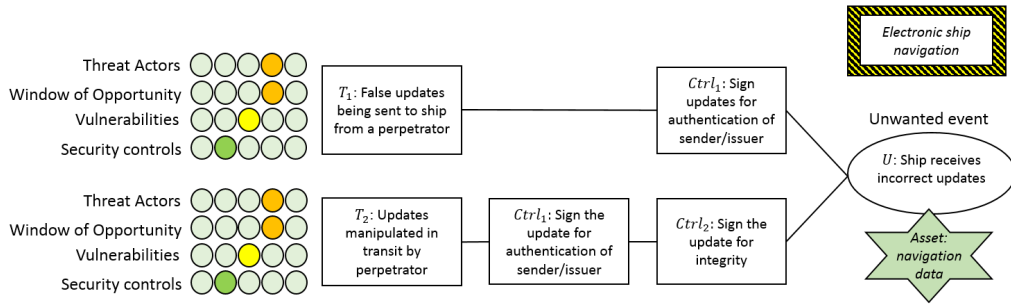
## 6 Use case example: Navigational Information Update

In this section, we demonstrate the use of our proposed methodology to represent unwanted events in a bow-tie diagram and to assess the corresponding risk. The context is cyber security threats in the maritime communication case study introduced in Section 3. The use case we investigate is called *Navigational Information Update*. The objective here is to illustrate the visualization, and not to present the complete description.

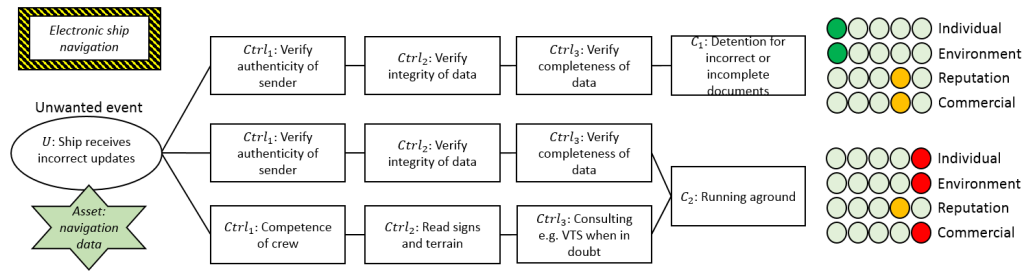
Ships are required to keep critical electronic databases up to date. Such databases include electronic charts and lists of navigation signals. Updating can be done by requesting updates as the voyage progresses and getting data from the chart provider. In the near future, this will be implemented over an Internet based service via satellite or other high capacity carriers. Failing to get the right data can cause safety hazards as well as a danger of detention by the Port State Control in the next port. In addition, some of this information is provided by commercial companies that need to protect the supplied information from copying to non-paying ships.

In this example we address electronic ship navigation as a potential hazard and we want to assess the risk of the unwanted event “Ship receives incorrect updates”. The affected asset is the navigation data that is being transferred. Fig. 3 and 4 illustrate the identified threats, security controls and potential consequences that we have identified in our analysis.

To compute the risk, we need to assess the probabilities of all the identified threats, as well as the impact value and probability of the worst-case consequence



**Fig. 3.** The left hand threat side with preventive controls for the unwanted event “Ship receives incorrect updates”



**Fig. 4.** The right hand consequence side with reactive controls for the unwanted event “Ship receives incorrect updates”

identified for this unwanted event. The assessment of a risk assessor, who has considered the threat actors, window of opportunity, vulnerabilities and security controls, is used as a source for this threat prediction. If we for instance set probability of threat  $T_1 = 0.45$  and probability of threat  $T_2 = 0.23$ , and then apply Equation 2, we can compute the probability of the unwanted event:

$$p(U) = 1 - (1 - p(T_1)) \times (1 - p(T_2)) = 1 - (1 - 0.45) \times (1 - 0.23) \approx 0.57 \quad (3)$$

Furthermore, let’s assume the consequence  $C_1 = 0.3$ ,  $p(C_1) = 0.5$ ,  $C_2 = 0.7$  and that  $p(C_2) = 0.2$ . By applying Equation 1, we find that the risk of the unwanted event to be:

$$R(U) \approx 0.57 \times 0.7 \times 0.2 \approx 0.08 \quad (4)$$

This number does not mean much by itself, but can be used as a relative number when comparing with other unwanted events, and to justify the addition of barriers/controls.

As illustrated by this simple example, the bow-tie diagram provides an illustrative overview over the identified threats, security controls and potential consequences of the unwanted event.

## 7 Discussion

To make useful cyber security visualizations with bow-tie diagrams, we needed to identify which security concepts to include and what kind of quantified input data would be meaningful as input to the diagrams. In our case, we have done this in separate processes, one for each side of the diagrams. For the left side (potential causes and threats, including likelihood), security and domain experts participated in a workshop setting (n=10), while the right side (consequences and their severity) was evaluated by representatives from maritime industry and coastal authorities through an online survey (n=18). Both groups were working with the same set of seventeen service scenarios for maritime communication, and twenty use cases that overlapped between the services. Note that none of these groups worked directly with bow-ties as a graphical notation, but were focused on types of threats, consequences and estimating values based on their experience and expert opinion. Based on these results, which are documented in [36], we have developed the methodology for visualizing concepts and quantified values for cyber security with the bow-tie notation, addressing research question 1 from Section 1. This has then been applied to a sample of the use cases from the service scenarios, as shown in Section 6, to demonstrate the utility of our approach. We consider this to be a first step of evaluation, where we have shown that the main security concepts can be contained and visualized. We have also tried to address research question 2 by adding color coded indicators to the diagrams, which are there to justify the likelihood and impact of an unwanted event. However, further work is needed to do in-depth evaluation on how this is perceived and found useful by other analysts, stakeholders from the maritime domain, as well as stakeholders from other safety domains.

Some general observations we have made when working with bow-tie modelling is that they are very suitable to show the broadness and distribution of different causes and consequences for unwanted events, along with protective and reactive barriers. However, this approach also has its limitations. For instance, a bow-tie diagram will struggle to represent the depth and details of how attacks can be performed. Furthermore, a single cause or threat can lead to different unwanted events, therefore, there can easily be repetition/redundancy between a collection of bow-ties addressing different hazards. We therefore recommend that the diagrams are complemented with more established methods for threat modelling, and that these are reused and referred to from nodes within the bow-ties. This can for instance be fault-trees for safety, or generic attack trees or misuse cases for security, that Meland et al. [33] have already showed can be shared and reused between different projects, organizations or domains with benefit. A prerequisite to realize this would be modelling tool support beyond simple drawing tools, as well as collaboration and willingness to share knowledge between risk analyst addressing both safety and security.

To capture more security related information within a bow-tie, it is also possible to add specific nodes in the model for concepts such as threat actors and vulnerabilities. We believe that this would lead to an unnecessary complexity of the diagram, and it would lose some of its advantage as an easy to grasp

graphical representation. The number and types of nodes would increase, and there would in many cases be many-to-many relationships between threat actors, threats, vulnerabilities, and security controls. Therefore, we rather use the more simplified notation of indicators related threat and consequence branches, that sums up for instance whether it is likely there are many relevant threat actors.

## 8 Conclusion

Safety assessments with bow-tie diagrams give a good pictorial understanding of major risks and how they are controlled. This is a technique that many of the high-risk industries are already familiar with, such as oil and gas, mining, aviation, maritime and public health services [37]. Due to the increasing connectivity of cyber physical systems, these are the same industries that are now becoming more and more exposed to cyber attacks. To avoid conflicting goals and requirements between safety and security, we believe that adding security to the bow-tie notation is more accommodating than inducing yet another specialized, separate modelling technique that tries to capture all aspects of safety and security. Bow-tie diagrams are meant to be easy to understand, and by combining a minimal set of security concepts along with associated indicators, we can show both safety and security considerations without overflowing the diagrams.

**Acknowledgments** The research leading to these results has been performed as a part of the Cyber Security in Merchant Shipping (CySiMS) project, which received funding from the Research Council of Norway under Grant No. 256508, and the SafeCOP-project, which received funding from the ECSEL Joint Undertaking under Grant No. 692529. We appreciate all the feedback and comments from anonymous reviewers that helped us improve this paper.

## References

1. ISO/IEC 27005 Information technology – Security techniques – Information security risk management. Tech. rep. (2008), [http://www.iso.org/iso/catalogue/\\_detail?csnumber=56742](http://www.iso.org/iso/catalogue/_detail?csnumber=56742)
2. Digitale Sarbarheter Maritim Sektor. Tech. rep. (2015), <https://www.regjeringen.no/contentassets/fe88e9ea8a354bd1b63bc0022469f644/no/sved/7.pdf>
3. Andrews, J.D., Moss, T.R.: Reliability and risk assessment. Wiley-Blackwell (2002)
4. Banerjee, A., Venkatasubramanian, K.K., Mukherjee, T., Gupta, S.K.S.: Ensuring safety, security, and sustainability of mission-critical cyber–physical systems. *Proceedings of the IEEE* 100(1), 283–299 (2012)
5. Bau, J., Mitchell, J.C.: Security modeling and analysis. *IEEE Security & Privacy* 9(3), 18–25 (2011)
6. Bhatti, J., Humphreys, T.: Hostile control of ships via false gps signals: Demonstration and detection. *Navigation* (2016)
7. Bieber, P., Brunel, J.: From safety models to security models: preliminary lessons learnt. In: *International Conference on Computer Safety, Reliability, and Security*. pp. 269–281. Springer (2014)

8. Byers, D., Ardi, S., Shahmehri, N., Duma, C.: Modeling software vulnerabilities with vulnerability cause graphs. In: Proceedings of the International Conference on Software Maintenance (ICSM06). pp. 411–422 (2006)
9. Casey, T.: Threat agent library helps identify information security risks (2007), <https://communities.intel.com/docs/DOC-1151>
10. CGE Risk Management Solutions: Using bowties for it security (2017), <https://www.cgerisk.com/knowledge-base/risk-assessment/using-bowties-for-it-security>
11. Chevreau, F.R., Wybo, J.L., Cauchois, D.: Organizing learning processes on risks by using the bow-tie representation. *Journal of hazardous materials* 130(3), 276–283 (2006)
12. Chockalingam, S., Hadziosmanovic, D., Pieters, W., Teixeira, A., van Gelder, P.: Integrated safety and security risk assessment methods: A survey of key characteristics and applications. arXiv preprint arXiv:1707.02140 (2017)
13. Cimpean, D., Meire, J., Bouckaert, V., Vande Casteele, S., Pelle, A., Hellebooge, L.: Analysis of cyber security aspects in the maritime sector (2011)
14. Cockshott, J.: Probability bow-ties: a transparent risk management tool. *Process Safety and Environmental Protection* 83(4), 307–316 (2005)
15. De Dianous, V., Fiévez, C.: Aramis project: A more explicit demonstration of risk control through the use of bow-tie diagrams and the evaluation of safety barrier performance. *Journal of Hazardous Materials* 130(3), 220–233 (2006)
16. DNV-GL AS: Recommended practice. cyber security resilience management for ships and mobile offshore units in operation (2016)
17. Ferdous, R., Khan, F., Sadiq, R., Amyotte, P., Veitch, B.: Analyzing system safety and risks under uncertainty using a bow-tie diagram: An innovative approach. *Process Safety and Environmental Protection* 91(1), 1–18 (2013)
18. Garvey, P.R., Lansdowne, Z.F.: Risk matrix: an approach for identifying, assessing, and ranking program risks. *Air Force Journal of Logistics* 22(1), 18–21 (1998)
19. Goldkuhl, G.: Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems* 21(2), 135–146 (2012)
20. Hall, P., Heath, C., Coles-Kemp, L.: Critical visualization: a case for rethinking how we visualize risk and security. *Journal of cybersecurity* 1(1), 93–108 (2015)
21. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. *MIS Q.* 28(1), 75–105 (Mar 2004), <http://dl.acm.org/citation.cfm?id=2017212.2017217>
22. Hpaul: Security: Bow Tie for Cyber Security (0x01): Ho... — PI Square (2016), <https://pisquare.osisoft.com/groups/security/blog/2016/08/02/bow-tie-for-cyber-security-0x01-how-to-tie-a-cyber-bow-tie>
23. IMO: Revised guidelines for formal safety assessment (fsa) for use in the imo rule-making process (2013)
24. Jürjens, J.: Umlsec: Extending uml for secure systems development. In: International Conference on The Unified Modeling Language. pp. 412–425. Springer (2002)
25. Khakzad, N., Khan, F., Amyotte, P.: Dynamic risk analysis using bow-tie approach. *Reliability Engineering & System Safety* 104, 36–44 (2012)
26. Kordy, B., Mauw, S., Radomirović, S., Schweitzer, P.: Foundations of attack-defense trees. In: International Workshop on Formal Aspects in Security and Trust. pp. 80–95. Springer (2010)
27. Kriaa, S., Pietre-Cambacedes, L., Bouissou, M., Halgand, Y.: A survey of approaches combining safety and security for industrial control systems. *Reliability Engineering & System Safety* 139, 156–178 (2015)

28. Kumar, R., Stoelinga, M.: Quantitative security and safety analysis with attack-fault trees. In: High Assurance Systems Engineering (HASE), 2017 IEEE 18th International Symposium on. pp. 25–32. IEEE (2017)
29. Lee, W.S., Grosh, D.L., Tillman, F.A., Lie, C.H.: Fault tree analysis, methods, and applications; a review. *IEEE Transactions on Reliability* R-34(3), 194–203 (Aug 1985)
30. Lund, M.S., Solhaug, B., Stølen, K.: Model-driven risk analysis: the CORAS approach. Springer Science & Business Media (2010)
31. Mauw, S., Oostdijk, M.: Foundations of attack trees. In: International Conference on Information Security and Cryptology. pp. 186–198. Springer (2005)
32. Meland, P.H., Gjære, E.A.: Representing threats in BPMN 2.0. In: Availability, Reliability and Security (ARES), 2012 Seventh International Conference on. pp. 542–550. IEEE (2012)
33. Meland, P.H., Tøndel, I.A., Jensen, J.: Idea: reusability of threat models—two approaches with an experimental evaluation. In: International Symposium on Engineering Secure Software and Systems. pp. 114–122. Springer (2010)
34. Michel, C.D., Thomas, P.F., Tucci, A.E.: Cyber Risks in the Marine Transportation System. The U.S. Coast Guard Approach
35. Mohr, R.: Evaluating cyber risk in engineering environments: A proposed framework and methodology (2016)
36. Nesheim, D., Rødseth, Ø., Bernsmed, K., Frøystad, C., Meland, P.: Risk model and analysis. Tech. rep., CySIMS (2017)
37. NevilleClarke: Taking-off with BowTie (2013), <http://www.nevilleclarke.com/indonesia/articles/topic/52/title/>
38. Ni, H., Chen, A., Chen, N.: Some extensions on risk matrix approach. *Safety Science* 48(10), 1269–1278 (2010)
39. Nielsen, D.S.: The cause/consequence diagram method as a basis for quantitative accident analysis. Tech. rep., Danish Atomic Energy Commission (1971)
40. Phillips, C., Swiler, L.P.: A graph-based system for network-vulnerability analysis. In: Proceedings of the 1998 workshop on New security paradigms. pp. 71–79. ACM (1998)
41. Piètre-Cambacédès, L., Bouissou, M.: Cross-fertilization between safety and security engineering. *Reliability Engineering & System Safety* 110, 110–126 (2013)
42. Raspotnig, C., Karpati, P., Katta, V.: A Combined Process for Elicitation and Analysis of Safety and Security Requirements, pp. 347–361. Springer Berlin Heidelberg, Berlin, Heidelberg (2012), [http://dx.doi.org/10.1007/978-3-642-31072-0\\_24](http://dx.doi.org/10.1007/978-3-642-31072-0_24)
43. Ruijters, E., Stoelinga, M.: Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools. *Computer science review* 15, 29–62 (2015)
44. Santamarta, R.: A wake-up call for satcom security. Technical White Paper (2014)
45. Schneier, B.: Attack trees. *Dr. Dobbs journal* 24(12), 21–29 (1999)
46. Sha, L., Gopalakrishnan, S., Liu, X., Wang, Q.: Cyber-physical systems: A new frontier. In: Sensor Networks, Ubiquitous and Trustworthy Computing, 2008. SUTC'08. IEEE International Conference on. pp. 1–9. IEEE (2008)
47. Shostack, A.: Threat modeling: Designing for security (2014)
48. Simon, H.A.: The sciences of the artificial. MIT press (1996)
49. Sindre, G., Opdahl, A.L.: Eliciting security requirements with misuse cases. *Requirements engineering* 10(1), 34–44 (2005)
50. Sun, M., Mohan, S., Sha, L., Gunter, C.: Addressing safety and security contradictions in cyber-physical systems. In: Proceedings of the 1st Workshop on Future Directions in Cyber-Physical Systems Security (CPSSW09) (2009)



51. Viscusi, W.K., Aldy, J.E.: The value of a statistical life: a critical review of market estimates throughout the world. *Journal of risk and uncertainty* 27(1), 5–76 (2003)
52. Winther, R., Johnsen, O.A., Gran, B.A.: Security assessments of safety critical systems using hazops. In: *International Conference on Computer Safety, Reliability, and Security*. pp. 14–24. Springer (2001)
53. Zalewski, J., Drager, S., McKeever, W., Kornecki, A.J.: Towards experimental assessment of security threats in protecting the critical infrastructure. In: *ENASE 2012-Proceedings of the 7th International Conference on Evaluation of Novel Approaches to Software Engineering*, Wroclaw, Poland (2012)

**D: ‘Facing uncertainty in cyber insurance policies’**

A license to reproduce the published material [20] for inclusion in this thesis has been obtained from Springer Nature.

D

# Facing Uncertainty in Cyber Insurance Policies

Per Håkon Meland<sup>1,2(✉)</sup>, Inger Anne Tøndel<sup>1,2</sup>, Marie Moe<sup>2</sup>,  
and Fredrik Seehusen<sup>2</sup>

<sup>1</sup> Norwegian University of Science and Technology, Trondheim, Norway  
{per.hakon.meland,inger.anne.tondel}@ntnu.no

<sup>2</sup> SINTEF Digital, Trondheim, Norway  
{per.h.meland,inger.a.tondel,marie.moe,fredrik.seehusen}@sintef.no

**Abstract.** Cyber insurance has gained less ground in Europe than in the U.S., but with emerging laws and regulations, the prospect of considerable fines for security breaches is pushing many organisations into this market. A qualitative interview study in Norway reveals the main uncertainty factors for organisations that have little experience with the cyber insurance consideration process, and how they perceive the products, process and expected support in case of a cyber incident. These uncertainty factors can be reduced by being aware of typical coverage gaps, exclusions and loss types that are commonly found in cyber insurance products.

**Keywords:** Cyber insurance · Risk management · Gap analysis · Exclusions · Coverage · Negotiation

## 1 Introduction

Cyber insurance is an expanding market, fuelled by the growing number of cyber threats as our society becomes increasingly dependent on interconnected digital technology. In fact, Lloyd's City Risk Index [16] and the World Economic Forum [28] both consider cyber attacks to be one of the top risks facing the world today. Cyber insurance can be defined as the “transfer of financial risk associated with network and computer incidents to a third party” [5], and is meant to take care of incidents that have low frequency and high impact.

In the U.S., there is and has been a considerable up-take of cyber insurance. A recent survey by Hiscox [14] reports that 55% of U.S. respondents state they have cyber insurance. Looking at Europe, the situation is a bit different. According to a survey by Marsh & McLennan Co, only 13% of European companies have purchased this [19]. Why nine out of ten cyber insurance policies in the world are in the U.S., can probably be explained by more than ten years of state breach notification laws [7]. The situation is likely to become more similar in Europe, when emerging data protection regulations take effect in the near future [9]. For this reason, many organisations are now preparing to enter this market, but this is a new and challenging task for them, since there are not well-established practices for considering cyber insurance.

© Springer International Publishing AG 2017  
G. Livraga and C. Mitchell (Eds.): STM 2017, LNCS 10547, pp. 89–100, 2017.  
DOI: 10.1007/978-3-319-68063-7\_6

The main contribution of this paper is a study of the demand side view of cyber insurance, driven by the following research questions:

1. What are the main uncertainty factors in the consideration phase as perceived by the demand side?
2. How can these uncertainties be reduced?

Section 2 gives an overview over the related work for this topic. The former research question is studied in Sect. 3 through qualitative interviews with Norwegian organisations, who only have very little experience with this new type of product. For the latter research question, we analyse and discuss these uncertainties in Sect. 4 with experiences found in a more global perspective to see whether or not they are well-founded, and what can be done to reduce them. Section 5 provides a conclusion to the work.

## 2 Related Work

There have already been several publications covering various challenges for the demand side of cyber insurance. Bandyopadhyay [2] have developed nine hypotheses on adoption of cyber insurance by organisations. He claims that organisations likely to adopt and utilise cyber insurance are recognized by high intensity of state of the art technology, business critical information systems, central management of cyber risks, efficient intra-organisational communication and collaboration, and imposed regulations. Those who are less accommodating typically have high security experience, high risk appetite, and a volatile business environment.

A survey by the Ponemon institute [21] provides some more empirical insight in which factors are most important when deciding whether or not to buy cyber insurance. For instance, 70 % of their respondents reported increasing interested in cyber-insurance policies after experiencing an incident. Among those that do not plan to buy insurance, the following main reason were given: “Premiums are too expensive” (52 %) and “Too many exclusions, restrictions and uninsurable risks” (44 %). Bandyopadhyay [3] has also argued that overpricing due to information asymmetry has been the primary reason for the limited growth of the cyber insurance market seen from the demand side. Additional barriers have been explained in separate studies by ENISA [12], U.S. Department of Homeland Security [23] and MARSH [17], such that firms already think they are covered by their existing general business interruption policies. Mr. Brew from Liberty International Underwriters [22] lists the following reasons why more customers do not buy cyber insurance:

- Cost and revenue concerns: Some see cyber security as a luxury purchase.
- Uncertainty: Will they actually pay out if there is an event? Untested market.
- High risk appetites: Technology entrepreneurs are risk takers, and do not see insurance as a necessary investment.
- Maturity: Companies are unaware of the availability of cyber insurance (and also about cyber security risk exposure).

A recent joint global study [20] by Swiss RE and IBM Institute for Business Value concluded with a very simple reason why companies were not buying cyber insurance; *they simply had not explored it*. This study included 1005 organisations from 15 industries in over 50 countries.

As can be seen from the literature, there can be many reasons why cyber insurance is still regarded as somewhat “immature, with room for improvement” [15]. The policies themselves tend to have varying form, content and vocabulary, which makes it difficult to grasp coverage and terms, as well as compare policy offerings from different insurers [18]. Though many organisations presumably seem to have taken an informed decision when deciding upon cyber insurance, a significant portion is also sitting on the fence because they do not feel competent to make any decision due to *uncertainty*. In the next section, we explore some of these uncertainty aspects in more detail.

### 3 Interview Study

#### 3.1 Method

During the autumn of 2016, we conducted a series of ten in-depth interviews with representatives from Norwegian organisations. Since only a very few Norwegian organisations currently have cyber insurance, the limited market made it difficult to design a larger empirical study. Still, we were able to obtain representation from different industries, such as finance, media, retail, critical infrastructure and IT. Most of these organisations are large by Norwegian standards, but a few were also medium size in the range of one hundred employees. Six out of the ten organisations had experience with a cyber insurance consideration process. Out of these, one organisation had acquired, two were still considering and three had decided not to invest in this option. The remaining four expressed their needs and thoughts if they were to start such a process.

We consider this setting to be representative for the Norwegian market and similar areas. Norway is considered to be technologically advanced and an example of a society that depends heavily on information systems, and thus, a society exposed to cyber threats. For instance, Norwegians use digital services to a large extent, well above EU average, and companies have a high on-line presence [10]. There is also a steady course towards a cashless economy where almost all transactions are done electronically [26]. Figure 1 illustrates a sample of digital maturity factors compared to the rest of Europe.

Each of the interviews lasted about one hour, and had a semi-structured form where one researcher asked the questions, and another made notes and additional remarks. All the results were also digitally recorded, transcribed and coded in a set of a priori main categories with emerging sub-categories. The complete results of the interviews are out of scope for this paper, but we have extracted the main uncertainty aspects with respect to *products*, *consideration process* and expected *support* in the case of an incident.

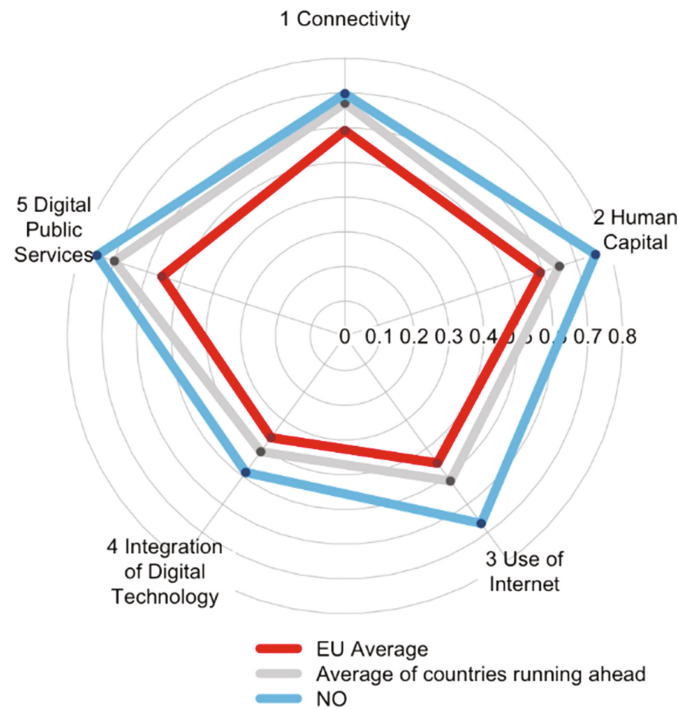


Fig. 1. The digital agenda scoreboard for Norway (2016) [10].

### 3.2 Results

**Products.** In general, the cyber-insurance products and market are perceived as immature by those organisations that have considered to buy cyber-insurance. Characteristics put out by the informants include “there are different definitions to the term cyber risk”, “the market is premature”, “products are not prepared thoroughly”, “there’s lot of fancy words that we don’t know the real meaning of”. One informant had asked if their regular insurance company could provide this product for them, but they did not have anything readily available. In this case, the insurance company made one on-the-fly especially for this organisation.

The most important thing that make cyber insurance interesting seem to be coverage and limit. Price is less important. The informants seem to all agree that insurance is for catastrophes, that is incidents with high consequence and low likelihood. With today’s cyber insurance products, coverage is perceived as low and not enough for to cover catastrophe costs. In addition, a cyber insurance will only cover parts of the real incident cost. Many of the companies we have talked to are mostly worried about reputation loss and loss of market position, and thus future income. Their impression was that these types of costs were not covered by an insurance. Many express that the cyber insurance products are difficult to understand, and that many aspects are unclear, illustrated by quotes such as “for the time being, there is a lot of promise-ware” and “it is a product

where it is not easy to get a concrete feeling of what is covered and not”. Also, some informants were critical to the competence of the insurance companies in this field, mentioning: “When we asked technical questions about security, they could not really answer” and “... they don’t know what they are selling”.

There was a clear trend that the informants seemed unsure about the real benefits of the existing products. In addition, the products are perceived as expensive compared to other insurance products. One informant characterised the premium as “random”, meaning it seems arbitrary what price you get based on the risk and the security measures of the company. This can be summed up by the following statement from one informant: “It is not everything that appears attractive and realistic for us to use. And the extent of coverage you will get in case of a break-in or an incident is a bit diffuse. They [insurance agents] say media support and so on, but what do they mean by that? It is very difficult to know the extent of that. In my opinion, the whole concept of cyber insurance is a bit vague and hard to grasp. The only thing that is concrete is the annual premium you have to pay”.

**Process.** When the organisations started the process of considering cyber insurance, the natural first step for them had been to assess their own cyber risk. Though most of the informants explain that their organisation already has some form of risk assessment practice that includes cyber, this does not seem to be enough to serve as a foundation for making decisions on whether or not to buy cyber coverage. Many of the organisations we talked to were still in the process of performing a more thorough risk assessment of their cyber risk, and a decision to buy cyber insurance was still pending from that assessment. However, as of now, they were still uncertain about their needs. The process of evaluating products was perceived complex and challenging for several reasons. First, as this is a new product, there is a general risk that no one picks up on it and takes responsibility for evaluating its relevance for the organisation (“it could easily fall between two stools”).

Second, risk managers or similar roles that handle other types of insurance products do not know that much about cyber. Thus, they need more support from brokers than what is the case with most other types of insurance products. They also need to interact with IT people internally, something they are not used to, and this exposes them to a field very different from their own main competence. A few notable quotes from the interviews:

- “... it is a new area, and vague because you do not know enough about computers and do not have the fantasy to understand what is happening”
- “...sounded a bit like science fiction the first time I heard about it”
- “...you suddenly enter a technological world that is much more complicated than sitting and reading nice contracts”

Third, as explained before, products are perceived to be immature and terms are often unclear. It was stated: “Terms should be clearer than they are right now. It seems that the insurance guys have just put up a list of things that would be nice to have. It does not say anything about at what level, and if



there are any requirements on proof. Do we have to document all our security measures?” and “what does it mean to have a firewall or antivirus? What are the requirements to the firewall or antivirus? Does it e.g. have to be patched? What about gathering evidence after an incident? The policy does not say anything about this”. Additionally, those that claim to know the cyber insurance market well, stated that this is developing rapidly, both when it comes to products and terms, and as a result, it is challenging to keep up to date.

Those companies with a lot of internal competence on insurance would actually prefer cyber as part of existing coverage, and not as a stand-alone product. One informant stated: “Then you can work with insurance companies that already know you, and it is cheaper”. Another argued the following: “It is a small extension you do in an existing program, while buying a stand-alone product, which is offered on a broad scale, is a totally different scenario. There is extra work to for us to support them with their analysis, I’d rather work with those that already know our risk exposure”.

As part of the negotiation with insurance companies, self-evaluation forms and questionnaires are frequently used. The organisations that have experience with these consider them to be relevant, but with the following remarks:

- “The form seems very high level, maybe because the policies are only meant to cover low pay-outs.”
- “These forms are not suitable for complex, heterogeneous organisations, such as ours, with many locations for our different offices. There must be a dialogue.”

One of the informants emphasised that their key success factor was obtaining a better understanding of the total risks that the organisation faced, and existing insurance coverage. This was stated as: “The most important thing we did in the beginning was this gap-analysis: what do we have, what do we lack when it comes to insurance”. This was an activity in which they invested a lot of time together with their broker.

**Support.** Though practical support from insurance companies in the case of an incident was not something most informants talked much about, there was an agreement on the following two things:

- It would be interesting to them if they would get access to highly specialised competence on the specific technology they are using.
- If such help should be useful, there must be a close relationship between the insurer and insuree over time, and an openness, “so they will know us and know how things are. They should not have to do a lot of research to understand us before they can start implementing countermeasures and limit damages”.

Access to specialised competence and ability to have a close relationship were not something that the informants necessarily perceive to be part of current products, but something that would make the products more interesting. As of

now, they are not sure if this is the kind of support that is offered. Additionally, there are uncertainties related to pay-out. This was related to lack of experience and unclear products (as explained above). One informant explained that they consider cyber insurance products to stem from the U.S. These [insurance] companies are perceived to have other ranges of pay-outs than what's common in Norway. This can impact the trust towards the product and process effort in case there is an incident.

## 4 Reducing Uncertainty

A cyber insurance is not a silver bullet, and can never be a complete replacement for risk modification as a part of a risk management plan. Any organisation considering cyber insurance should focus on what kind of coverage they need to address their residual risk, and harmonise this with other insurances [13, 25]. But in order to do this, a lot of the uncertainty aspects from the previous section must be overcome. There is a lot of uncertainty related to the products themselves. Besides the novelty of the product, this is also caused by the fact that such policies are not standard products, but a result of a negotiation between the insuree and insurer. The negotiation phase is used to tailor standard products to more specific coverage and establish a price for individual insurees [15]. This includes defining exclusions, carve-backs, premium, payouts or support actions in the case of cyber events, cover limits (or caps), etc. To quote Siemens and Beck [25]; “buying an off-the-shelf policy can result in disaster”. A negotiation would also be used when renewing policies, but for cyber insurance in particular, many organisations are doing this for the first time. The products themselves are therefore very much reliant on the process, and the support is a result of what has been agreed upon.

In the following sub-sections, we show what to be aware of when negotiating coverage of gaps, exclusions and loss.

### 4.1 Closing the Gaps

A gap analysis for information security is usually performed to discover potential gaps between what level of security you have in place and requirements from regulations and standards, or in simple terms, *comparing where you are against where you want to be*. We noted from our interviews with Norwegian organisations, that when they were mentioning gap analysis, this was mainly about *determining whether or not the organisation was under-insured for cyber events*.

Most organisations already have a portfolio of insurance products in place, and general liability, property and crime insurances can in many cases cover a number of cyber events. However, they are not designed to fully cover all the potential costs and losses related to cyber risk [15]. In fact, there are significant cyber-related risks that remain largely uninsurable or the coverage is modest compared with the overall exposure [27]. With little experience on claim from traditional insurances and cyber policies, there is a lot of uncertainty about

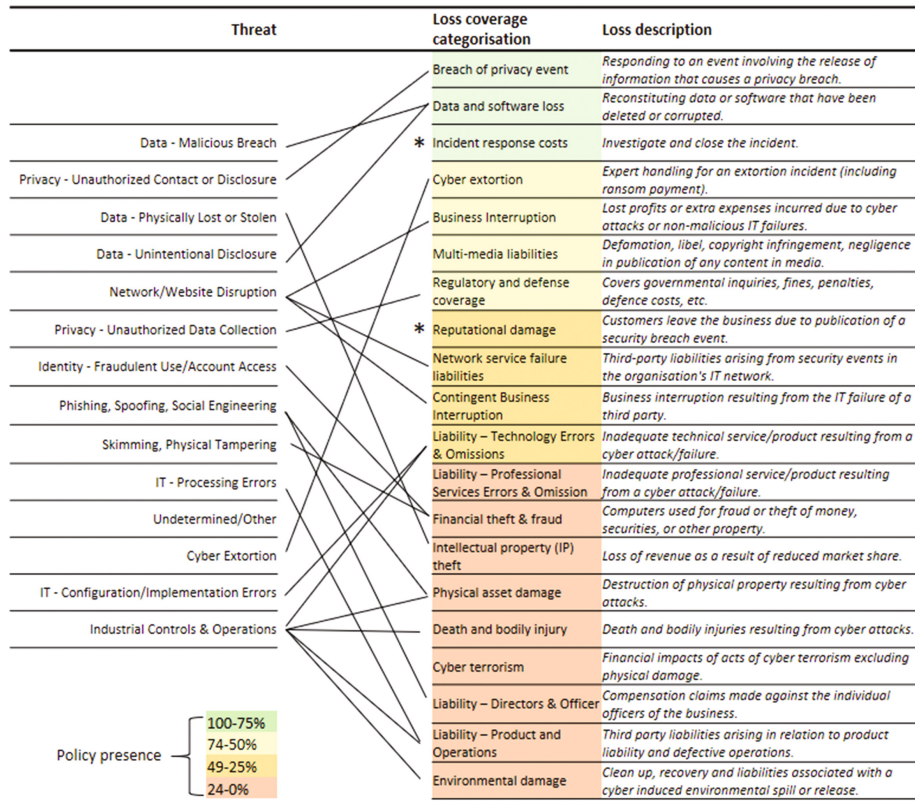


Fig. 2. Mapping between threats and loss coverage

loss coverage gaps. Therefore, it is important to have an idea of what risks are typically insurable and non-insurable, sort out the ones that can cover the needs, and prepare clarifying questions for the negotiation table.

In Fig. 2, we have combined two datasets to illustrate how cyber threats can be mapped to insurance coverage. The column to the left contains a threat categorisation from Advisen<sup>1</sup> ordered by registered loss amount. For instance, “Data - Malicious Breach” accounts for 622 cases with a total loss amount of \$5,311,075K, while “Industrial Controls and Operations” accounts for merely two cases with a total of \$85K. The rightmost two columns show typical loss coverage categories as defined in a study by Cambridge Centre for Risk Studies [6]. These 19 categories extend an original cyber loss categorisation scheme developed by a steering group of 15 insurance companies, several industry organisations and government agencies [17]. There was quite a variation on coverage in the

<sup>1</sup> The dataset we have received from Advisen is dated November 2016 and contains 33023 world-wide cyber loss events. Romanosky has described the origins of this data in [24].

26 UK insurance products that was examined (two-thirds of what was estimated to be on the market). The colour scheme in Fig. 2 indicates how commonly the losses were part of the policies. Due to the lack of an official vocabulary for cyber threats and losses, there is a significant degree of interpretation in this mapping, especially for the lower coverage segment. Also, note that a single threat category can lead to more than one type of loss. Especially “Incident response costs” and “Reputational damage” would have so many threat links that we did not include them in the figure.

In an ideal world, the most expensive threats would normally be present in cyber insurance policies, but as the figure shows, this is currently not the situation. It may also be that a policy contains coverage that is not relevant or necessary for the organisation that considers the insurance. It is therefore recommended to create an individual risk profile that can be used to compare expected threat exposure with what the policy offers to cover.

## 4.2 Checking for Exclusions

It is typically in the lower coverage segment in Fig. 2 that you will run into a world of exclusions that organisations must review, both for their existing policies and those under consideration. For instance, “cyber terrorism” is an ambiguous term, and probably more related to the people or group behind the threat, along with the associated motivation (e.g. political, religious, ideological or similar purposes), rather than the action itself. Many organisations would assume that any DDOS attack would be covered by Business Interruption, but according to [8], such claims could be rejected on the basis of a terrorism exclusion if there is a hacktivist group behind.

Besides war and terrorism exclusions, that are typically found in any type of insurance policy, there are exclusions that are particular for cyber insurance. The following check-list is based on reports from the Association of British Insurers [1] and Thomas Bentz from Holland & Knight [4]:

- **Court jurisdiction** - The territories of U.S. and Canada tend to be excluded from cyber insurance policies purchased in Europe.
- **Claims by related entities** - Claims related to loss of data belonging to employees (personal data), contractors and partial owned subsidiaries are not normally included.
- **Bodily injury and property damage** - As can be seen from the loss coverage categories in Fig. 2, tangible assets tend to be excluded. General liability policies may already cover the direct expenditures, but probably not subsequent lawsuits.
- **Crime vs cyber insurance** - Consequences that are meant to be covered by a crime insurance policy, such as attacks leading to theft of money, will not be reimbursed by a cyber insurance (“Financial theft & fraud” loss coverage category).
- **Mechanical/electronic failure** - Claims due to computers that stop working. Should be limited to malicious acts causing the computers to fail for the policy to respond.

- **Laptop exclusions** - Coverage for portable electronic devices tends to be excluded, especially if they do not encrypt their contained data.
- **Patent, software, copyright infringement** - We have already seen that IP theft belong to the lower coverage segment. Carve-backs (exclusion overrides) can be negotiated to cover claims caused by non-management employees and third parties.
- **Employment practices** - Incident arising from poor or insecure employment processes are often excluded or can shrink the policy's limits.
- **Employee benefit plan breaches** - Often referred to ERISA exclusions in the U.S., breach of data found in e.g. pension plans and health benefit plans, can be a special condition that is not covered.
- **Prior acts** - Since there may be a long time between time of breach and time of discovery, exclusions can limit the covered incidents originating from before policy inception and long tailed consequences.
- **The insured vs insured** - Such exclusion state that a claim made by one insured against another insured is not covered, however, there can be carve-backs for various reasons such as violation of privacy.

### 4.3 Clarifying Loss

It is also useful to clarify what costs are covered for different types of cyber events. The data material from Advisen divides this into the following four categories, which we have detailed using definitions from Allianz [11]:

- **Response costs** - E.g. forensic investigations, identifying and preserving lost data, advice on legal and regulatory duties, notification costs according to legal and regulatory requirements, determining the extent of indemnification obligations in contracts with third party service providers, credit monitoring services and other remedial actions required after a loss of data, public relations expenses to handle negative publicity.
- **Economic loss** - E.g. loss of business income caused by a targeted attack, indemnity for stolen funds, indemnity for cyber extortion.
- **Litigated cases** - Defense costs and damages for which the insured is liable.
- **Fines and penalties** - Monetary fines and penalties levied by regulators arising from a loss of data.

Considering these categories, the Advisen data show that *response costs* has the highest average cost, while *economic loss* has the lowest, averaging about one third of response costs. Any organisation should during the negotiation get a clear definition about what kind of costs are covered for different types of incidents, and check these caps.

## 5 Conclusion

Cyber insurance has gained less ground in Europe than in the U.S., but with emerging laws and regulations, the prospect of considerable fines for security

breaches is pushing many organisations into this market. What remains to see is: Can these organisations properly navigate through the still immature and obscured maze of cyber insurance products, or will they be easy prey for insurance companies offering policies that will not be worth much in the case of cyber events?

We have shown that the demand side struggles with several uncertainty factors when it comes to cyber insurance, and this has hindered the confidence in the product and market adoption process. Our qualitative interview study was based in Norway, but we believe that the same observations are found wherever regulations have not been a strong driving force yet. With an expected increase in this market, there is a need for better guidance in the consideration processes, as well as clearly defined and understandable terms and conditions for the product. This especially includes the identification of security gaps within the organisation, and coverage gaps, exclusions and loss types for the cyber insurance policy.

It was also found during the interview studies, that even for organisations that did not end up buying an insurance, there were still positive effects from the consideration process, since it brought attention and awareness of cyber security to the management level and across the organisation.

**Acknowledgments.** This research has been performed as part of the inSecurance project funded by SINTEF Digital. We would like to thank the representatives from all the organisations that participated in the interviews for sharing their experiences with us, and discussions with representatives from brokers and insurance companies. A final gratitude to Professor Guttorm Sindre at NTNU for feedback and comments.

## References

1. Association of British Insurers: Making sense of cyber insurance: a guide for SMEs. Technical report, ABO (2016)
2. Bandyopadhyay, T.: Organizational adoption of cyber insurance instruments in it security risk management: a modeling approach, Proceedings, P. 5 (2012)
3. Bandyopadhyay, T., Mookerjee, V.S., Rao, R.C.: Why IT managers don't go for cyber-insurance products. *Commun. ACM* **52**(11), 68–73 (2009)
4. Bentz, T.: Negotiating key cyber exclusions. Insurance Day (2015). [https://www.insuranceday.com/news\\_analysis/legal\\_focus/negotiating-key-cyber-exclusions.htm](https://www.insuranceday.com/news_analysis/legal_focus/negotiating-key-cyber-exclusions.htm)
5. Böhme, R., Schwartz, G.: Modeling cyber-insurance: towards a unifying framework. In: Workshop on the Economics in Information Security (WEIS) (2012)
6. Cambridge Centre for Risk Studies: Managing cyber insurance accumulation risk. University of Cambridge, Technical report (2016)
7. Cohn, C., Barlyn, S.: European, Asian companies short on cyber insurance before ransomware attack (2017). <http://www.reuters.com/article/us-cyber-attack-insurance-idUSKCN18B00H>
8. CRIF: Cyber insurance and the terrorism exclusion (2014). <http://www.cyberriskinsuranceforum.com/content/cyber-insurance-and-terrorism-exclusion>
9. DG Justice and Consumers: Reform of EU data protection rules (2016). [http://ec.europa.eu/justice/data-protection/reform/index\\_en.htm](http://ec.europa.eu/justice/data-protection/reform/index_en.htm)

10. Digital Single Market: Digital scoreboard (2016). <https://ec.europa.eu/digital-single-market/digital-scoreboard>
11. Dobie, G., Collins, S.: A guide to cyber risk - managing the impact of increasing interconnectivity. Technical report, Allianz (2015). <http://www.agcs.allianz.com/assets/PDFs/risk%20bulletins/CyberRiskGuide.pdf>
12. ENISA, Robinson, N.: Incentives and barriers of the cyber insurance market in Europe. Report 28th June 2012. [http://www.enisa.europa.eu/activities/Resilience-and-CIIP/national-cyber-security-strategies-ncsss/incentives-and-barriers-of-the-cyber-insurance-market-in-europe/at\\_download/fullReport](http://www.enisa.europa.eu/activities/Resilience-and-CIIP/national-cyber-security-strategies-ncsss/incentives-and-barriers-of-the-cyber-insurance-market-in-europe/at_download/fullReport)
13. Gordon, L.A., Loeb, M.P., Sohail, T.: A framework for using insurance for cyber-risk management. *Commun. ACM* **46**(3), 81–85 (2003)
14. Hiscox: The hiscox cyber readiness report (2017). <https://www.hiscox.co.uk/cyber-readiness-report/docs/cyber-readiness-report-2017.pdf>
15. Hurtaud, S., Flamand, T., de la Vaissiere, L., Hounka, A.: Cyber insurance as one element of the cyber risk management strategy February 2015. <https://www2.deloitte.com/lu/en/pages/risk/articles/cyber-insurance-element-cyber-risk-management-strategy.html>
16. Lloyd's, Cambridge Centre for Risk Studies: Lloyds City Risk Index 2015–2025 (2015). <http://hwww.lloyds.com/cityriskindex/>
17. Maude, F.: The role of insurance in managing and mitigating the risk (2015). <https://www.marsh.com/uk/insights/research/uk-cyber-security-role-of-insurance-in-managing-mitigating-risk.html>
18. Meland, P.H., Tøndel, I.A., Solhaug, B.: Mitigating risk with cyberinsurance. *IEEE Secur. Priv.* **13**(6), 38–43 (2015)
19. Nikolaeva, M., Rivet, M.: French central bank chief urges insurers to step up cyber risk coverage (2017). <http://www.reuters.com/article/us-france-insurance-idUSKBN1591Q9>
20. Pain, L.D., Anchen, J., Bundt, M., Durand, E., Schmitt, M.: Cyber: In search of resilience in an interconnected world (2016). [http://www.swissre.com/library/archive/Demand\\_for\\_cyber\\_insurance\\_on\\_the\\_rise\\_joint\\_Swiss\\_Re\\_IBM\\_study\\_shows.html](http://www.swissre.com/library/archive/Demand_for_cyber_insurance_on_the_rise_joint_Swiss_Re_IBM_study_shows.html)
21. Ponemon: Managing cyber security as a business risk: Cyber insurance in the digital age. Report, Ponemon Institute, August 2013. <http://www.ponemon.org/blog/managing-cyber-security-as-a-business-risk-cyber-insurance-in-the-digital-age>
22. Protection, National, Directorate, Programs: Cyber risk culture roundtable readout report, Technical report. U.S. Department of Homeland Security (2013)
23. Protection, National, Directorate, Programs: Cybersecurity insurance workshop readout report, Technical report. U.S. Department of Homeland Security (2012)
24. Romanosky, S.: Examining the costs and causes of cyber incidents. *J. Cybersecur.* **2**(2), 121–135 (2016)
25. Siemens, R., Beck, D.: How to buy cyber insurance. *Risk Manag.* **59**(8), 40 (2012)
26. Svanemyr, S.: Kontantene forsvinner i butikkene (Norwegian) (2016). <https://tinyurl.com/j7qaq9>
27. Swiss Re Institute: Cyber: getting to grips with a complex risk. Technical report, Swiss Re (2017). <http://www.swissre.com/library/sigma.01.2017.en.html>
28. World Economic Forum: The global risks report 2016, 11st edn. (2016). [http://www3.weforum.org/docs/GRR/WEF\\_GRR16.pdf](http://www3.weforum.org/docs/GRR/WEF_GRR16.pdf)

**E: ‘When to treat security risks with cyber insurance’**

Included is the published material [21], following the Creative Commons Attribution 4.0 International (CC BY 4.0) licensing arrangement used by C-MRiC.ORG.



E

# When to Treat Security Risks with Cyber Insurance

---

Per Håkon Meland<sup>\*,+</sup> and Fredrik Seehusen<sup>\*</sup>

<sup>\*</sup>*SINTEF Digital, Norway*

<sup>+</sup>*Norwegian University of Science and Technology, Norway*

## **ABSTRACT**

Transferring security risk to a third party through cyber insurance is an unfamiliar playing field for a lot of organisations, and therefore many hesitate to make such investments. Indeed, there is a general need for affordable and practical ways of performing risk quantification when determining risk treatment options. To address this concern, we propose a lightweight, data-driven approach for organisations to evaluate their own need for cyber insurance. A generic risk model, populated with available industry averages, is used as a starting point. Individual organisations can instantiate this model to obtain a risk profile for themselves related to relevant cyber threats. The risk profile is then used together with a cyber insurance profile to estimate the benefit and as a basis for comparing offers from different insurance providers.

*Keyword: Cyber insurance, risk quantification, risk profile, threats, decision making.*

---

## **1 INTRODUCTION**

Many organisations are now in the process of determining whether or not they should invest in cyber insurance. This is a new and challenging task for them, since there are not many established practices seen from the demand side. Though stand-alone cyber insurance products have been around for a couple of decades, they are still regarded as "*somewhat immature, with room for improvement*" (Hurtaud, Flamand, Vaissière, & Hounka, 2015).

For instance, varying form, content and vocabulary make it difficult to grasp coverage and terms, as well as compare policy offerings from different insurers (Meland, Tøndel, & Solhaug, 2015). Additional barriers have been explained by ENISA (ENISA, 2012), such that firms already think they are covered by their existing general business interruption policies. This optimistic belief of coverage was confirmed by a later UK study that MARSH published in 2015 (Maude, 2015). In 2016, a global study (Pain, Anchen, Bundt, Durand, & Schmitt, 2016) by Swiss RE and IBM Institute for Business Value concluded that the main reason why companies were not buying cyber insurance, was that they simply had not explored it.

**E** The main contribution of this paper is a proposed assessment approach for organisations considering to buy cyber insurance. This is an investment decision that requires an understanding of cyber risk, but quantifying cyber risk is very challenging, even for large organisations with in-house security competence. Insurances are meant to take care of incidents that have low frequency and high impact, and single organisations are lacking historical data they can base their cost/benefit analysis on. At the same time, the technology, insurance market and threat picture are in constant development, making past experiences and data less valuable.

There have already been several publications covering various aspects for the demand side of cyber insurance. For instance, Gordon et al. (Gordon, Loeb, & Sohail, 2003) and Wang (Wang, 2017) provide frameworks for cyber risk management, where insurance is one of the means for risk reduction. Yannacopoulos et al. (Yannacopoulos, Lambrinouidakis, Gritzalis, Xanthopoulos, & Katsikas, 2008) discuss the level of coverage a firm should consider for privacy breaches given that the premium levels are set. Grossklags et al. (Grossklags, Christin, & Chuang, 2008) use game-theoretic models for shifting between investments in protection and self-insurance. They have showed that self-insurance may be more advantageous, especially when there are other firms that are more likely to be attacked due to weaker security. This model has been extended to also include market insurance by Johnson et al. (Johnson, Böhme, & Grossklags, 2011). Pal and Golubchik (Pal & Golubchik, 2010) have proposed a mathematical framework that co-operative and non-co-operative Internet users can exploit to balance defence investments with partial and full coverage insurance models. Böhme and Schwartz (Böhme & Schwartz, 2010) have developed a framework for modelling cyber insurance markets, which includes various attributes for cyber risk in relation to cyber insurance. Böhme and Schwartz also present a literature survey where both the demand and supply side are considered in this context. Later on,

Mukhopadhyay et al. (Mukhopadhyay, Chatterjee, Saha, Mahanti, & Sadhukhan, 2013) proposed another model to help firms decide upon cyber insurance, but with focus on utility for both the insurer and insured. A cyber risk profile for individual organisations is denoted by a unique utility function, and a Copula-aided Bayesian Belief Network (CBBN) model is used for assessing and quantifying the cyber risk.

Unlike the existing approaches, ours is initiated by a generic risk model that individual organisations can specialise to obtain a more optimal and tailored risk profile model for themselves. We assume that the organisations already have protection mechanisms in place, but want to reduce residual risk of rare events through cyber insurance. To evaluate the benefit of insuring, the risk profile is evaluated with and without a suitable insurance profile. The main advantages of this approach are that it makes use of available data concerning threats, likelihood and loss, and that it does not require the organisation to share information about their risks and incidents with external parties during the consideration phase. This should in turn make the organisation better equipped for negotiations with insurance agencies or agents.

This paper is structured as follows. The background and details of the approach are explained in Section 2. Section 3 discusses strengths and weaknesses, and section 4 provides a conclusion.

## **2 DESCRIPTION OF THE APPROACH**

The development of the approach has been motivated by a Norwegian study (Meland, Tøndel, Moe, & Seehusen, 2017) on current practices for cyber insurance decision making. This study showed that obtaining a good understanding of cyber risk exposure is considered to be a critical, but also a very complex and challenging necessity. Risk managers and people with similar roles that already handle other types of insurance products within a company, typically do not know that much about cyber. Therefore, they find it difficult to perform cost/benefit analysis for cyber security, and to have a good and dynamic overview of the relevant threats. Another significant observation was that not all organisations are willing to share a lot of information about their security procedures, controls and incidents with arbitrary insurance agents, since they fear that this information could be leaked and damage their reputation or be exploited for attacks. There was also a general concern on how smaller organisations, lacking security competence and resources, will be able to make proper judgement on whether to buy cyber insurance or not.

The main target group of our approach is therefore organisations with limited in-house security expertise, that are considering investing in or renegotiating a cyber insurance policy. It has been the goal of our approach to be affordable, directly applicable for practitioners, and also to take advantage of available information. It is meant to accommodate specific industrial domains and improve over time as quantitative data becomes more reliable. We have used previous work from practical risk assessment (Tran, Solhaug, & Stølen, 2013), and adapted this to specifically address cyber insurance decision making.

The approach follows the steps as illustrated in Figure 1. The creation of a *generic risk model* is the first step, and is a collaborative task between security professionals, researchers and cyber underwriters. This risk model represents the typical threat events that a cyber insurance can cover, and what impact/consequences such events can lead to. The model includes sets of baseline data to be used as a starting point. The second step is performed by individual organisations to create a *risk profile* for themselves. The final step is the creation of a *cyber insurance profile*, which indicates cost reductions per threat in combination with the premium. The next sections explain these steps in more detail and with examples.

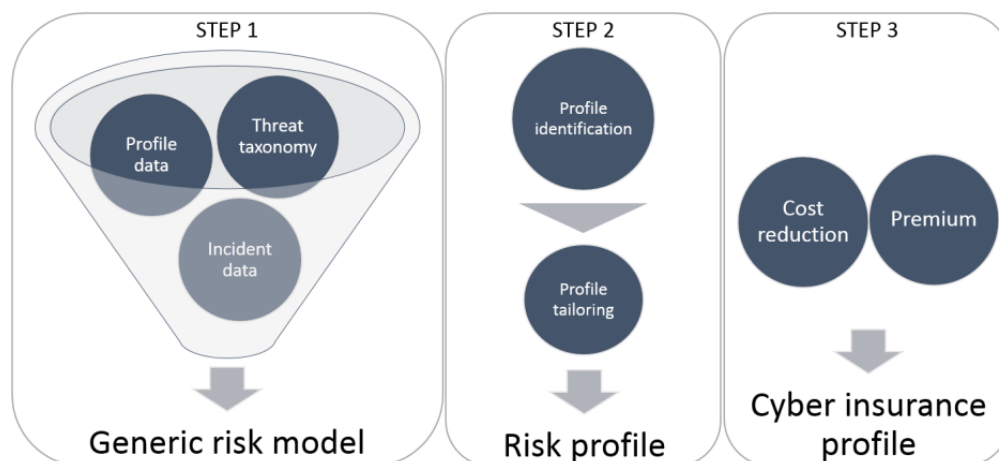


Figure 1. Approach overview.

### Assembling a generic risk model

The purpose of this step is to define a risk model which is generic in the sense that it is parametrised by *company profiles*. A company profile is set of values such as size, location and industry that can be used in order to

categorise a given company. More precisely, a *generic risk model* is a triple  $(T, f, c)$  consisting of:

- A set of threats  $T$ ;
- A generic frequency function  $f$  which takes a company profile  $cp$  as input and yields a mapping that for each threat  $t$  in  $T$  yields a frequency estimating how often incidents caused by  $t$  occur per year.
- A generic cost function  $c$ , which takes a company profile  $cp$  as input and yields a mapping that for each threat  $t$  in  $T$  yields the estimated cost of incidents caused by  $t$ .

The company specific *risk profile*, obtained from the generic risk model  $M = (T, f, c)$  for a given company profile  $cp$ , is a triple  $(T_s, f_s, c_s)$  whose threats  $T$  are equal to the threats of  $M$ , and whose frequency and cost functions defined by  $f_s = f(cp)$  and  $c_s = c(cp)$  respectively.

We let  $rv$  be a function, which takes a frequency  $v_f$  and a consequence  $v_c$  and yields their *risk value* defined by their product, i.e.  $rv(v_f, v_c) \triangleq v_f \cdot v_c$ . The risk value of a given threat can be viewed as the *annual expected loss* due to incidents caused by this threat since we assume that frequencies estimate number of threat incidents *per year*.

In order to use the risk profile for determining whether or not to buy cyber insurance, we need to compute the total aggregated risk value, or the total annual expected loss due to all threats. We do not make any assumptions about overlap between threat incidents, i.e. whether the occurrence of one threat incident counts as an occurrence of an incident caused by another threat. For this reason, the total aggregated risk value is described by an interval, where the minimum interval value corresponds to the aggregated risk value in the case where there is the maximum possible overlap, and the maximum interval value corresponds to the case where there is the minimum possible overlap. More precisely, we define the total risk value of a risk profile  $(T_s, f_s, c_s)$  by the interval:

$$[rv(f_{min}, c_{min}), rv(f_{max}, c_{max})]$$

where

- $f_{min}$  is the frequency for the case where there is a maximum possible overlap defined by the maximum frequency value  $\max(\{f_s(t) \mid t \in T_s\})$ ;
- $f_{max}$  is the frequency for the case where there is no overlap defined

by the sum of all frequency values  $\sum_{t \in T_s} f_s(t)$ ;

- $c_{max}$  is the cost estimate of an arbitrary threat incident for the case where no overlap, defined by  $(\sum_{t \in T_s} rv(f_s(t), c_s(t)))/f_{max}$ ;
- $c_{min}$  is the cost estimate of an arbitrary threat incident for the case with overlap. The definition of  $c_{min}$  may depend on the risk profile. In this paper, we let  $c_{max}$  be an approximation of this cost, i.e. we let  $c_{min} = c_{max}$ . However, other definitions should be considered if this is not a reasonable approximation for the given risk profile.

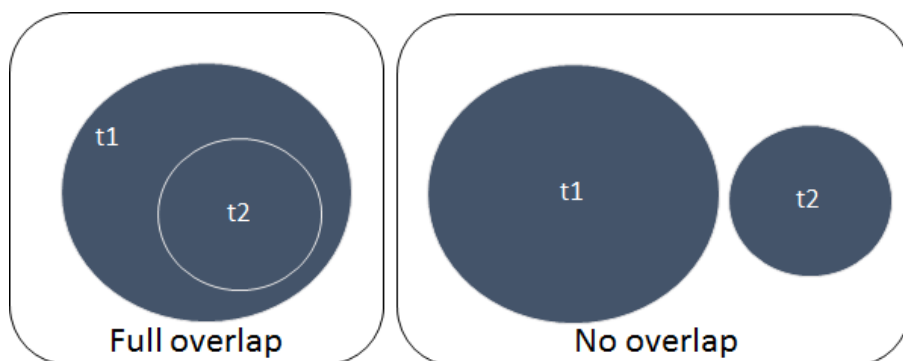


Figure 2 Full overlap: each incident caused by threat  $t_2$  also count as an incident caused by  $t_1$ . No overlap: No incident caused by  $t_2$  count as an incident caused by  $t_1$ .

Figure 2 illustrates what is meant by overlap. On the left-hand side, all incidents caused by threat  $t_2$  also count as incidents caused by threat  $t_1$ . This is the case of maximum possible overlap which is assumed in the definition of  $f_{min}$ . Here, the combined frequency of the incidents caused by threats  $t_1$  and  $t_2$  is equal to the frequency of  $t_1$ , i.e.  $f_{min} = \max\{f_s(t_1), f_s(t_2)\} = f_s(t_1)$ . On the right-hand side, there is no overlap, i.e. no incident caused by  $t_2$  counts as an incident caused by  $t_1$ . This is the case of minimum possible overlap which is assumed in the definition of  $f_{max}$ . Here, the combined frequency of the incidents caused by threats  $t_1$  and  $t_2$  is equal to the sum of the frequencies of  $t_1$  and  $t_2$ , i.e.  $f_{max} = f_s(t_1) + f_s(t_2)$ . In this case, the total risk value is also equal to the sum of the risk value for each threat. This allows us to calculate the value of  $c_{max}$  as follows due to the definition of  $rv$ :

$$\begin{aligned} rv(f_{max}, c_{max}) &= rv(f_s(t_1), c_s(t_1)) + rv(f_s(t_2), c_s(t_2)) \\ f_{max}c_{max} &= rv(f_s(t_1), c_s(t_1)) + rv(f_s(t_2), c_s(t_2)) \\ c_{max} &= (rv(f_s(t_1), c_s(t_1)) + rv(f_s(t_2), c_s(t_2)))/f_{max} \end{aligned}$$

In the continuation of this step, we include the following two activities for defining threats and profile type, based on data sources containing threat information, and for defining frequency and cost functions mapped to threats.

### *Define threats and profile type*

The purpose of this activity is to define the set of threats and the possible company profile attributes of the generic risk model. The activity starts by identifying data sources that contain threat categories or taxonomies and statistics about threat occurrences and cost. After this has been done, a threat categorisation is selected from the data sources or created based on the data sources.

In our experience, nearly all data sources use different threat categorisations, thus creating a new unified categorisation is not straight forward. The data sources also tend to vary with respect to the detail and the kind of statistics they contain. Table 5 in the Data appendix gives an example of different threat categories from four different data sources.

As an example for this paper, we have chosen a threat categorisation from Advisen<sup>1</sup> as a basis for the generic model. This is not because we think its categorisation is the best, but because it seems to have the most detailed cost data. In addition to this, we have used data from Klahr et al. (Klahr, Amili, Shah, Button, & Wang, 2016) for estimating the likelihood of threat incidents. Advisen also contains data about how events (threat incidents) are distributed on different *industries*, and Klahr et al. contains data about the frequency of cyber breaches based on *company size*. In this paper, we will therefore consider company profiles based on size and industry. The Data appendix contains all the data material that we will use in this paper. Table 1 gives a definition of the threats  $T$ , as well as size  $S$  and industry  $I$  values considered in this paper.

---

<sup>1</sup>The dataset we have received from Advisen is dated November 2016 and contains 33023 cyber loss events. Romanosky has described the origin of this data in (Romanosky, 2016).



Table 1. Definition of threats  $T$ , size values  $S$ , and industry values  $I$ .

Name	Definition
$T$	Data - Malicious Breach; Privacy - Unauthorized Contact or Disclosure; Data - Physically Lost or Stolen; Data - Unintentional Disclosure; Network/Website Disruption; Privacy - Unauthorized Data Collection; Identity - Fraudulent Use/Account Access; Phishing, Spoofing, Social Engineering; Skimming, Physical Tampering; IT - Processing Errors; Undetermined/Other; Cyber Extortion; IT - Configuration/Implementation Errors; Industrial Controls & Operations
$S$	Micro; Small; Medium; Large
$I$	Services; Finance, Insurance and Real Estate; Public Administration; Wholesale and Retail Trade; Manufacturing; Transportation, Communications and Utilities; Mining and Construction; Agriculture, Forestry and Fishing

### *Define frequency and cost functions*

The purpose of this activity is to define frequency and cost functions that map threats and company profiles to frequency and cost estimates. The definition should be made on the basis of the data that have been identified in the previous activity (which in our case is summarised in the *Data* appendix).

The definition of the frequency function  $f$  and the cost function  $c$  of our generic risk model are given in Table 2. Note that the available data material is not 100% applicable for defining the frequency and cost function that we need. For instance, the estimation of percentage of companies that have been breached due to a cyber threat is based on a survey in the UK (Klahr et al., 2016), and it may not be applicable for companies outside the UK. Another example is related to the cost data from Advisen, where it is unclear whether the data basis is a good representation of the entire population, and not for instance skewed to data mostly from the US, to big companies or to cyber events that are particularly costly. Indeed, the cost of

cyber events is estimated to be significantly higher in Advisen than in other studies from e.g. Kasperksy (Kaspersky, 2015) and particularly Klahr et.al. In the definitions, we implicitly assume that the data material used is applicable.

*Table 2. Definition of the frequency and cost functions  $f$  and  $c$  (and helper functions) for the generic risk model.*

<b>Name</b>	<b>Definition</b>	<b>Description</b>
$ei(i) \triangleq$	$ev(i)/26872$	Proportion of cyber incidents/events that occur in industry $i$ . Here, $ev(i)$ denotes events recorded for industry $i$ (the “Events” column in Table 7) and 26872 denotes the total number of recorded events (last row of Table 7).
$tp(t) \triangleq$	$ev(t)/33023$	The proportion of incidents that are caused by threat $t$ . Here $ev(t)$ denotes the number of events/incidents recorded with respect to threat $t$ according to Advisen (the “Events” column in Table 6) and 33023 is the total number of events recorded according to Advisen (last row of Table 6).
$b(s, i) \triangleq$	$\frac{b_s(s)ei(i)}{b_s(s)ei(i)+(1-b_s(s))si(i)}$	The likelihood of experiencing a threat incident within a one year period for a company with size $s$ in industry $i$ . Here, $b_s(s)$ denotes the proportion of companies of size $s$ breached within a one year period (column “Proportion breached” in Table 9) and $si(i)$ denotes the relative size of the industry $i$ (column “Relative size” in Table 10). We make the simplifying assumption that those companies that were breached, were breached only once, and that this breach counts as a single threat incident.
$f((s, i))(t) \triangleq$	$b(s, i) \cdot tp(t)$	The number of times per year that an incident caused by threat $t$ occurs under the profile $(s, i)$ , i.e. for a company with size $s$ in industry $i$ .
$c((s, i))(t) \triangleq$	$evt(t)/tl(t)$	The expected cost of an incident caused by threat $t$ under company profile $(s, i)$ . Here, $evt(t)$ denotes the number of events with recorded loss for threat $t$ (corresponding to the “Events with loss” column in Table 6) and $tl(t)$ denotes total recorded loss (corresponding to the “Total loss” column in Table 6).

In Table 2,  $f$  and  $c$  are the generic frequency and cost mappings that given the company profile  $(s, i)$  yields a frequency and a cost mapping that is

specific to companies in industry  $i$  having size  $s$ . Both  $f$  and  $c$  are defined in terms of the other helper functions in Table 2. Of these, the definition of function  $b(s, i)$  may not be immediately clear. This function is based on the function  $b_s(s)$  which gives us the proportion of companies of size  $s$  breached within a one year period. However, what we want, and which is given by  $b(s, i)$ , is the proportion of companies of size  $s$  that were breached *provided* that they are in industry  $i$ . This is defined to be equal to the ratio of companies of size  $s$  in industry  $i$  that were breached ( $b_s(s)ei(i)$ ) to companies of size  $s$  in industry  $i$  that were both breached *and not breached* ( $b_s(s)ei(i) + (1 - b_s(s))si(i)$ ).

### Tailoring an individual risk profile

The purpose of this step is to adapt the generic risk model to a particular organisation. Unlike the previous step, the intended user is a company or organisation that considers cyber insurance. The step has two activities, first a profile is identified and a corresponding risk profile is derived from the generic risk model. Then, this risk profile is manually refined by tailoring the frequency and cost values.

In the following, we will illustrate the step in an example for a fictive company we refer to as Acme, which is a medium sized company that provides an online marketplace where users can buy and sell goods and services from each other.

*Table 3. The derived risk profile (second and third column) and the manually refined profile (fourth and fifth columns). Only the calculated risk value for the latter profile is shown.*

Threat	Frequency	Cost	Frequency	Cost	Risk value
Data - Malicious Breach	0.217	8538707	0.217	8538707	1856717
Privacy - Unauthorized Contact or Disclosure	0.116	5191220	0.116	5191220	601702
Data - Physically Lost or Stolen	0.076	983992	0.000	983992	0
Data - Unintentional Disclosure	0.074	1547339	0.074	1547339	114929
Network/Website Disruption	0.032	1327197	1.000	100000	100000
Privacy - Unauthorized Data Collection	0.012	1770338	0.012	177033	21466
Identity - Fraudulent Use/Account Access	0.012	3167541	4.000	100000	400000
Phishing, Spoofing, Social Engineering	0.011	40435298	0.011	40435298	447775
Skimming, Physical Tampering	0.011	1973479	0.000	1973479	0

IT - Processing Errors	0.007	92043291	0.000	92043291	0
Undetermined/Other	0.003	0	0.000	0	0
Cyber Extortion	0.003	92615	0.003	92615	278
IT - Configuration/Implementation Errors	0.003	12427442	0.000	12427442	0
Industrial Controls & Operations	0.001	42655	0.000	4265	0
<b>Total risk value/expected loss per year</b>					<b>[2608007, 3542866]</b>

### *Instantiate generic risk model*

Based on the Acme profile (Size: *Medium* and Industry: *Services*) and the definition of our generic model, we can automatically derive the corresponding risk profile. This risk profile is shown in Table 3 in the first frequency and cost columns (column two and three). Here, the frequency value represents occurrences of the given threat incidents *per year* and the cost value represent the cost of threat incidents in USD.

The frequency and cost of each threat incident are calculated by the frequency function  $f$  and the cost function  $c$  defined in Table 2. For instance, the frequency for the threat  $t = \text{"Data - Malicious Breach"}$ , for company size  $s = \text{"Medium"}$  in industry  $i = \text{"Services"}$  is

$$f((s, i))(t) = b(s, i) \cdot tp(t) = 0.578 \cdot 0.376 = 0.217$$

### *Update metrics with own data (if any)*

In this step, a domain expert can manually tailor the risk profile to her organisation. The procedure for this step is as follows: Walk through each threat, classify into one of the three categories described below, and adjust the frequency and cost accordingly. The three categories are:

- **Irrelevant threats**, i.e. threats that do not apply to the company, threats that are negligible, or threats that the company is not interested in insuring. Since the generic risk model is intended to capture all possible cyber threats, it will typically be the case that many of the threats are not relevant. For these threats, the frequency should be set to zero.
- **Familiar threats**, i.e. threats that have occurred in the past and/or occur on a regular basis. For these threats, the frequency should typically be increased, but the cost estimate should often be decreased, since the prior experience in dealing with these kinds of

threats contributes to lowering the cost. To avoid too much disalignment, adjustments in either direction can be based on the general prediction approach by Kahneman (Kahneman, 2011), which is used to adjust reference class averages with non-regressive intuitive predictions. In practice, the correlation between the risk profile attributes and the more specific attributes of the organisation can be used as a basis for this.

- **Unfamiliar threats**, i.e. threats for which there is no prior experience, but that could potentially occur. For these threats, the likelihood of the risk profile derived from the generic model provides a good starting point for frequency estimation, and should be kept unchanged if the company has no information about this threat. The same applies for the cost.

Continuing the example, we have shown the refined risk profile for the company in question in the second frequency and cost columns (the fourth and fifth columns) of Table 3. For Acme, physical attacks or unintended incidents are considered out of scope. The frequencies for threats in rows 3,9-11,13-14 have therefore been set to 0. Acme users buy and sell goods and services from each other, and fraudulent use of the service happens regularly i.e. about every four months. The typical attack vector is that the attacker is able to obtain the credentials of an end-user to the site by hacking the end-user directly. Hence, Acme is not directly responsible, but it could be *perceived* that way by the market. The cost is therefore not negligible, but not as high as in the derived risk profile. Acme experiences "Network/Website Disruption" from time to time, but these issues are covered by the service level agreement with the company that hosts the online marketplace, and the cost of these kind of incidents have been lower than in the derived profile. The frequency and cost of the remaining threats have been left unchanged. In Table 3, we can see that the threat with the highest risk value is "Data - Malicious Breach". This gives an indication of the types of threats that should be in focus when considering risk transfer to cyber insurance.

### Creating a cyber insurance profile

Central to our approach, the decision on risk transfer should be based on a *cyber insurance profile*. A cyber insurance profile is a pair  $(cc, p)$  consisting of a cost cover function  $cc$  that takes a threat  $t$  as input and yields an estimate of how much of the cost of incidents caused by  $t$  will be covered by the insurance if they occur, and a cost  $p$ , the insurance premium, estimating the cost of insurance per year. These estimates must be determined based on a given insurance policy. The insurance premium is

often easy to determine, but the cost coverage can be more difficult to estimate as it also requires an understanding of the exclusions of the insurance policy. For a discussion of the kind of threats that are usually covered by cyber insurance, the reader is referred to (Meland et al., 2017).

Given a risk profile and an insurance profile, the *residual cost* of each threat  $t$  is obtained by subtracting the cost cover for  $t$  from the cost of  $t$  as specified in risk profile (setting the value to zero if the cost cover happens to be greater than the cost). We can then calculate a new (residual) total risk value based on the residual cost values. An insurance profile is *beneficial* if the total residual risk value with insurance plus the insurance premium is lower than the total risk value without insurance.

Table 4 gives an example of an insurance profile (column one and two), where we have assumed that the insurance covers the cost of each threat incident by 2500000 USD if they occur. Columns three and four in Table 4 shows the risk profile of our running example (Table 3) under this insurance profile. Here the cost of each threat incident has been reduced by 2500K USD, nullifying most of the residual risk values. The total residual risk value ranges from 1831796 USD to 2045122 USD. In the worst case, the benefit of this insurance profile is 562885 USD, i.e. the minimum total risk value of the risk profile without insurance minus the maximum total risk value of the risk profile with insurance (Table 4). Hence, in this case, the insurance would be beneficial if the premium is below 562885 USD per year.

*Table 4. Example of a cost cover function (column one and two) of an insurance profile and a risk profile under this insurance profile (columns one, three, and four).*

Threat	Cost cover	Frequency	Cost (res.)	Risk value (res.)
Data - Malicious Breach	2500K	0.217	6038707	1313099
Privacy - Unauthorized Contact or Disclosure	2500K	0.116	2691220	311933
Data - Physically Lost or Stolen	2500K	0.000	0	0
Data - Unintentional Disclosure	2500K	0.074	0	0
Network/Website Disruption	2500K	1.000	0	0
Privacy - Unauthorized Data Collection	2500K	0.012	0	0
Identity - Fraudulent Use/Account Access	2500K	4.000	0	0

Phishing, Spoofing, Social Engineering	2500K	0.011	37935298	420090
Skimming, Physical Tampering	2500K	0.000	0	0
IT - Processing Errors	2500K	0.000	89543291	0
Undetermined/Other	2500K	0.000	0	0
Cyber Extortion	2500K	0.003	0	0
IT - Configuration/Implementation Errors	2500K	0.000	9927442	0
Industrial Controls & Operations	2500K	0.000	0	0
<b>Total residual risk value/expected loss per year</b>			<b>[1831796, 2045122]</b>	

### 3 DISCUSSION

We have designed this approach to aid cyber insurance decision making based on the identified needs from a specific country (Meland et al., 2017). Still, we argue that this is transferable to other regions as well, since cyber threats are global, technology is converging and organisations seem to be facing the same barriers when dealing with cyber insurance.

Our notion of a risk model and risk profile is quite simple compared to other risk models we have already mentioned in the literature. First, we model likelihoods as real values representing frequencies of occurrence, but other notions of likelihood could have been possible. For instance, probabilities, intervals of probabilities or frequencies, or probability distributions. Second, we only model the likelihood that an incident caused by a threat occurs, and the cost of this incident if it occurs. However, it would also have been possible to model how often threat attacks occur, how likely it is that they succeed if they are carried out, what vulnerabilities could be exploited, what barriers are in place, etc.

There are three reasons why we have chosen to use the simple risk model. First, we are interested in defining a generic model which could apply to a large number of organisations, and we cannot rely on experiences from a particular organisation when estimating the likelihood and the cost of this model. We must therefore rely on threat statistics from available data sources in order to do the estimation. These statistics is often provided at a general level, and there is no unique and accepted source of information about e.g. economic magnitude (Eling & Wirfs, 2016), (Armin et al., 2015). For instance, it would be difficult to find an estimate of how often a particular kind of threat attack will succeed if it is carried out. Therefore, we have chosen not to include this in the risk model. The same reason applies for the way we have modelled frequency and cost. Using probability

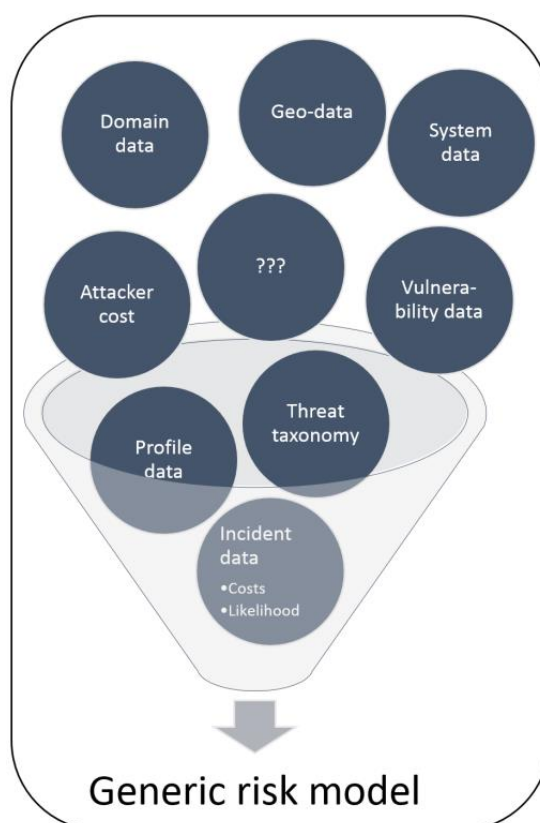
distributions for both of these would have provided more analysis options, but finding the statistical data material available need to derive these distributions is difficult. Also pointed out by (Sigma, 2017), full probabilistic models are still in their infancy, and better cyber risk models will eventually emerge as understanding of the fundamental risk drivers develops and more data about cyber losses become available. Second, we aim to have a lightweight approach which is understandable for a non-expert, and which can be carried out in little time. The tailoring activity is meant to adjust for how the organisations perceive themselves compared to other businesses. Furthermore, the approach may be extended with more advanced utility functions for the demand side of cyber insurance, e.g. as suggested in (Mukhopadhyay, Saha, Mahanti, Chakrabarti, & Podder, 2005), (Böhme & Schwartz, 2010), (Eling & Wirfs, 2016) and (Wang, 2017) if needed. Third, the risk model is not intended to give a completely accurate description of the cyber risks for the organisation. Instead, it is meant to be used as a *guide* for the further steps in deciding whether or not to buy cyber insurance. Although our approach is based quantitative data, high accuracy is not important as long as it informs the decision making process. The approach can also be combined with the cyber insurance decision plan suggested by Gordon et al. (Gordon et al., 2003), which also includes steps for assessing insurance gaps, evaluating available policies, and selecting a specific policy. However, their work was published very early, and does not address that in practice, negotiations are used to tailor policy coverage and price to individual insurees instead of offering standard products (Hurtaud et al., 2015). To quote Siemens and Beck (Siemens & Beck, 2012); "*buying an off-the-shelf policy can result in disaster*".

Regardless whether or not an organisation chooses to go forward with cyber insurance, this can only be part of the solution, and should not lead to negligence of security controls. In fact, there are significant cyber-related risks that remain largely uninsurable or the coverage is modest compared with the overall exposure (Sigma, 2017).

We have not defined as a part of our approach exactly *who* should be involved in the various steps for each organisation, since this will typically vary based on the size and type of the organisation, but an overview of suggested roles related to recommendations and decision is already given in a report from SANS (Filkins, Wright, & Bradford, 2016). A worrying finding from the same report, is that security professionals are rarely (28%) involved in the decision-making process leading to the purchase of cyber insurance.



For the continuation of our work, the generic model must be further developed, preferably with better measurement data, since we found a lot of deviations between different sources. Additional profile attributes and baseline values can be added, such as number of employees, system data (technological dependencies), GDP, geographic location, as well as an indication of risk appetite. Figure 3 illustrates what might be additional data ingredients going into the funnel.



*Figure 3. Possible additions of baseline data for the generic risk model.*

A vast set of rating indicators for cyber insurance have been identified by Innerhofer-Oberperfler and Brey (Innerhofer-Oberperfler & Brey, 2010), but reference datasets must be made available in order to enrich the generic risk model and create more accurate risk profiles. We share the same positive opinion as Biener et al. (Biener, Eling, & Wirfs, 2015), that with increased market development, we can expect better data sources as risk pools grow larger. Platforms for data sharing, organised by national regulators and international associations, should help keep this data accurate and updated to overcome the challenge of rapid technology development

and changing threat pictures. Even so, before increasing the complexity of the actuarial data, more systematic evaluations of the approach itself should be conducted, including a sensitivity study on the use of inaccurate data. We have so far received informal feedback during workshops with the insurance industry, and they appreciate the way risk models can be matched with insurance product to help their customers. Both insurers and insureds clearly share the common goal of better cyber security quantifications based on predictive, dynamic threat models.

## 4 CONCLUSION

We have observed that the demand side would like to have more practical help with deciding whether they need cyber insurance as a risk treatment option. Though several approaches for calculating insurance utility exist in the literature, they rely heavily on good input values for likelihood and costs/loss, and determining this is a great challenge for individual organisations. Our approach utilises available data sources to define a generic risk model, which is again tailored to the risk profile of individual organisations. The caveat here is that cyber event data will be quickly outdated and irrelevant if it is not updated and improved over time. We encourage the security and insurance community to make data about emerging threats and related costs available so that organisations can make informed decisions about risk treatment on a regular basis.

## 5 REFERENCES

- Armin, J., Thompson, B., Ariu, D., Giacinto, G., Roli, F., & Kijewski, P. (2015). 2020 cybercrime economic costs: No measure no solution. *10th international conference on availability, reliability and security (ARES)*, (pp. 701-710). IEEE.
- Biener, C., Eling, M., & Wirfs, J. (2015). Insurability of cyber risk: An empirical analysis. *The Geneva Papers on Risk and Insurance*.
- Böhme, R., & Schwartz, G. (2010). Modeling Cyber-Insurance : Towards A Unifying Framework. *Workshop on the Economics of Information Security*, 1-36.
- Eling, M., & Wirfs, J. H. (2016). Cyber Risk: Too Big to Insure?--Risk Transfer Options for a Mercurial Risk Class. *I. VW Schriftenreihe*, 59.
- ENISA. (2012). *Incentives and barriers of the cyber insurance market in Europe*.
- Fidelity. (2016). *Quarterly Sector Update, forth quarter 2016*. Retrieved from [https://www.fidelity.com/bin-public/060\\_www\\_fidelity\\_com/documents/Q4%202016%20Sector%20Update\\_Fidelity\\_FINAL.pdf](https://www.fidelity.com/bin-public/060_www_fidelity_com/documents/Q4%202016%20Sector%20Update_Fidelity_FINAL.pdf)
- Filkins, B., Wright, B., & Bradford, D. (2016). *Bridging the Insurance/InfoSec Gap: The SANS 2016 Cyber Insurance Survey*.
- Gordon, L., Loeb, M., & Sohail, T. (2003). A framework for using insurance for cyber-risk management. *Communications of the ACM*, 46(3), 81-85.

- Grossklags, J., Christin, N., & Chuang, J. (2008). Secure or insure?: a game-theoretic analysis of information security games. *Proceedings of the 17th international conference on World Wide Web*, (pp. 209-218). ACM.
- Hurtaud, S., Flamand, T., Vaissière, L. d. I., & Hounka, A. (2015, February). Cyber Insurance as one element of the Cyber risk management strategy. *Inside magazine*.
- Innerhofer-Oberperfler, F., & Breu, R. (2010). *Potential rating indicators for cyberinsurance: An exploratory qualitative study*. In *Economics of Information Security and Privacy* (pp. 249-278). Springer, Boston, MA.
- Johnson, B., Böhme, R., & Grossklags, J. (2011). Security games with market insurance. *International Conference on Decision and Game Theory for Security* (pp. 117-130). Springer, Berlin, Heidelberg.
- Kahneman, D. (2011). *Thinking, fast and slow*: Macmillan.
- Kaspersky. (2015). *Damage Control: The Cost of Security Breaches IT Security Risks* Retrieved from <https://media.kaspersky.com/pdf/it-risks-survey-report-cost-of-security-breaches.pdf>
- Klahr, R., Amili, S., Shah, J. N., Button, M., & Wang, V. (2016). *Cyber Security Breaches Survey 2016*. Retrieved from <https://www.gov.uk/government/publications/cyber-security-breaches-survey-2016>
- Maude, F. (2015). *The role of insurance in managing and mitigating the risk*. Retrieved from <https://www.sbs.ox.ac.uk/cybersecurity-capacity/content/uk-hm-government-and-marsh-cyber-security-role-insurance-managing-and-mitigating-risk>
- Meland, P. H., Tøndel, I. A., Moe, M., & Seehusen, F. (2017). *Facing Uncertainty in Cyber Insurance Policies*. Paper presented at the International Workshop on Security and Trust Management, (pp. 89-100). Springer, Cham.
- Meland, P. H., Tøndel, I. A., & Solhaug, B. (2015). Mitigating risk with cyberinsurance. *IEEE Security & Privacy*, 13(6), 38-43.
- Mukhopadhyay, A., Chatterjee, S., Saha, D., Mahanti, A., & Sadhukhan, S. K. (2013). Cyber-risk decision models: To insure {IT} or not? *Decision Support Systems*, 56, 11-26. doi:<http://dx.doi.org/10.1016/j.dss.2013.04.004>
- Mukhopadhyay, A., Saha, D., Mahanti, A., Chakrabarti, B. B., & Podder, A. (2005). Insurance for cyber-risk: A utility model. *Decision*, 32, 153-170.
- Pain, L. D., Anchen, J., Bundt, M., Durand, E., & Schmitt, M. (2016). *Cyber: In search of resilience in an interconnected world*. Retrieved from [http://www.swissre.com/library/archive/Demand\\_for\\_cyber\\_insurance\\_on\\_the\\_rise\\_joint\\_Swiss\\_Re\\_IBM\\_study\\_shows.html](http://www.swissre.com/library/archive/Demand_for_cyber_insurance_on_the_rise_joint_Swiss_Re_IBM_study_shows.html)
- Pal, R., & Golubchik, L. (2010). Analyzing self-defense investments in internet security under cyber-insurance coverage. *30th International Conference on Distributed Computing Systems (ICDCS)*, (pp. 339-347). IEEE.
- Romanosky, S. (2016). Examining the costs and causes of cyber incidents. *Journal of Cybersecurity*, 2(2), 121-135. doi:<http://dx.doi.org/10.1093/cybsec/tyw001>
- Siemens, R., & Beck, D. (2012). How to buy cyber insurance. *Risk Management*, 59(8), 40.
- Sigma. (2017). *Cyber: getting to grips with a complex risk*. Retrieved from [http://www.swissre.com/library/sigma\\_01\\_2017\\_en.html](http://www.swissre.com/library/sigma_01_2017_en.html)
- Tran, L. M. S., Solhaug, B., & Stølen, K. (2013). An Approach to Select Cost-Effective Risk Countermeasures. *Proceeding of 27th IFIP WG 11.3 Conference on Data and Applications Security and Privacy (DBSEC'13)*, (pp. 266-273). Springer, Berlin, Heidelberg.

Wang, S. (2017). Integrated Framework for Information Security Investment and Cyber Insurance.

Yannacopoulos, A. N., Lambrinouidakis, C., Gritzalis, S., Xanthopoulos, S. Z., & Katsikas, S. N. (2008). Modeling privacy insurance contracts and their utilization in risk management for ICT firms. In *European Symposium on Research in Computer Security* (pp. 207-222). Springer, Berlin, Heidelberg.

## Data

Table 5. Examples of threat categories from the different data sources.

ENISA	Advisen	Kaspersky	Klahr et.al.
Malware	Data - Malicious Breach	Malware	Viruses, spyware or malware
Web-based attacks	Privacy - Unauthorized Contact or Disclosure	Phishing attacks	Other impersonating organisation in emails or online
line Web application attacks	Data - Physically Lost or Stolen	Accidental leaks/sharing of data by staff	Denial-of-service attacks
DoS/DDoS	Data - Unintentional Disclosure	Vulnerabilities / flaws in existing software	Access to computers, networks or services without permission (i.e. hacking)
Phishing	Network/Website Disruption	Network intrusion / hacking	Money stolen electronically
Insider threat	Privacy - Unauthorized Data Collection	Denial of service	Breaches from personally-owned devices
Cyber espionage	Identity - Fraudulent Use/Account Access	Loss/theft of mobile devices by staff	Personal information stolen
Ransomware	Phishing, Spoofing, Social Engineering	Intentional leaks / sharing of data by staff	Breaches from externally-hosted web services
Hactivism	Skimming, Physical Tampering	Fraud by employees	Unlicensed or stolen software downloaded
ICS/SCADA hacking	IT - Processing Errors	Theft of mobile devices by external party	Money stolen via fraud emails of websites
Critical vulnerabilities	Undetermined/Other	Cyberespionage	Software damaged or stolen
Physical damage/theft/loss	Cyber Extortion	Security failure by third party supplier	Breaches on social media
Malicious code	IT - Configuration/Implementation Errors	Targeted attacks aimed specifically at our organisation / brand	Intellectual property theft
Botnets	Industrial Controls & Operations		

*Table 6. Occurrences of threat incidents (events) and their loss measured in USD. Data source: Advisen.*

<b>Case Type</b>	<b>Events</b>	<b>Events with loss</b>	<b>Total loss</b>
Data - Malicious Breach	12 410	622	\$5 311 075K
Privacy - Unauthorized Contact or Disclosure	6 615	668	\$3 467 735K
Data - Physically Lost or Stolen	4 347	80	\$78 719K
Data - Unintentional Disclosure	4 239	102	\$157 829K
Network/Website Disruption	1 824	115	\$152 628K
Privacy - Unauthorized Data Collection	692	479	\$847 992K
Identity - Fraudulent Use/Account Access	675	102	\$323 089K
Phishing, Spoofing, Social Engineering	632	52	\$2 102 635K
Skimming, Physical Tampering	623	84	\$165 772K
IT - Processing Errors	390	41	\$3 773 775K
Undetermined/Other	196	0	\$0K
Cyber Extortion	171	153	\$14 170K
IT - Configuration/Implementation Errors	168	19	\$236 121K
Industrial Controls & Operations	41	2	\$85K
<b>Total</b>	<b>33 023</b>	<b>2 519</b>	

*Table 7. Occurrences of threat incidents distributed on industries. Data source: Advisen.*

<b>Industry</b>	<b>Events</b>
Services	11 447
Finance, Insurance and Real Estate	5 633
Public Administration	4 142
Wholesale and Retail Trade	2 668
Manufacturing	1 508
Transportation, Communications and Utilities	1 238
Mining and Construction	202
Agriculture, Forestry and Fishing	34
<b>Sum</b>	<b>26 872</b>

*Table 8. Industry sector size based in S\&P index. Data source: (Fidelity, 2016).*

<b>Sector size</b>	<b>Weight in S&amp;P index</b>
Consumer Discretionary	12.5
Consumer Staples	9.9
Energy	7.3
Financials	15.8
Health Care	14.7
Industrials	9.7

Information Technology	21.2
Materials	2.9
Telecommunication Services	2.6
Utilities	3.3

*Table 9. Proportion of companies in the UK that have been breached in a period of 12 months. Data source: (Klahr et al., 2016).*

Company type	Proportion breached
Micro	0.17
Small	0.33
Medium	0.51
Large	0.65
Overall	0.24

### Derived data

The data in Table 10 has been derived by mapping the industry categorisation of Table 8 to the categorisation of Table 3 as follows:

- {Information Technology, Telecommunication Services, Health Care}  $\mapsto$  Services
- {Financials}  $\mapsto$  Finance, Insurance and Real Estate
- {}  $\mapsto$  Public Administration
- {Consumer Discretionary, Consumer Staples}  $\mapsto$  Wholesale and Retail Trade
- {Industrials}  $\mapsto$  Manufacturing
- {Energy, Utilities}  $\mapsto$  Transportation, Communications and Utilities
- {Materials}  $\mapsto$  Mining and Construction
- {}  $\mapsto$  Agriculture, Forestry and Fishing

Note that the category "Public Administration" is not covered by the categories in Table 8. In the derived Table 10, we have assumed that this industry sector is 15% of the total. The relative size of the other industries in Table 10 are obtained by summing the percentages of their corresponding industries in Table 8 and multiplying by 0.85 (the proportion not covered by the Public Administration sector).

*Table 10. Relative industry size.*

Industry	Relative size
Services	32 %
Finance, Insurance and Real Estate	13 %

Public Administration	15 %
Wholesale and Retail Trade	19 %
Manufacturing	9 %
Transportation, Communications and Utilities	9 %
Mining and Construction	3 %
Agriculture, Forestry and Fishing	0 %

## **BIOGRAPHICAL NOTES**

**Per Håkon Meland** is a senior research scientist at the independent research institute SINTEF in Norway. He obtained his MSc degree in Computer Science at the Norwegian University of Science and Technology in 2002, where he is also a PhD fellow in the intertwined fields of threat modelling and security economics.

**Dr Fredrik Seehusen** a scientist at the Norwegian Defence Research Establishment. He gained his PhD in computer science from the University of Oslo in 2009. He has previously worked at SINTEF in the areas of risk assessment, security testing, and formal methods.

## **ACKNOWLEDGEMENTS**

This research has been performed as part of the inSecurance project by SINTEF Digital. We would like to thank Professor Guttorm Sindre at NTNU for feedback, all the informants that participated in the interviews for sharing their experiences and needs with us, and representatives from brokers and insurance companies with whom we have been discussion this topic. This work has been partially funded by the EU-project WISER (653321).

## **REFERENCE**

**Reference to this paper should be made as follows:** Meland, P.H. & Seehusen, F. (2018). When to Treat Security Risks with Cyber Insurance. *International Journal on Cyber Situational Awareness*, Vol. 3, No. 1, pp. 39-60.

**F: ‘An experimental evaluation of bow-tie analysis for security’**

Included is the published material [22], following the Creative Commons Attribution 4.0 International (CC BY 4.0) licensing arrangement used by Emerald Publishing Limited.



F

# An experimental evaluation of bow-tie analysis for security

Per Håkon Meland

*Department of Digital, SINTEF for Industriell og Teknisk Forskning,  
Trondheim, Norway and Department of Computer Science,  
Norwegian University of Science and Technology, Trondheim, Norway*

Karin Bernsmed and Christian Frøystad

*Department of Digital, SINTEF for Industriell og Teknisk Forskning,  
Trondheim, Norway, and*

Jingyue Li and Guttorm Sindre

*Department of Computer Science,  
Norwegian University of Science and Technology, Trondheim, Norway*

## Abstract

**Purpose** – Within critical-infrastructure industries, bow-tie analysis is an established way of eliciting requirements for safety and reliability concerns. Because of the ever-increasing digitalisation and coupling between the cyber and physical world, security has become an additional concern in these industries. The purpose of this paper is to evaluate how well bow-tie analysis performs in the context of security, and the study's hypothesis is that the bow-tie notation has a suitable expressiveness for security and safety.

**Design/methodology/approach** – This study uses a formal, controlled quasi-experiment on two sample populations – security experts and security graduate students – working on the same case. As a basis for comparison, the authors used a similar experiment with misuse case analysis, a well-known technique for graphical security modelling.

**Findings** – The results show that the collective group of graduate students, inexperienced in security modelling, perform similarly as security experts in a well-defined scope and familiar target system/situation. The students showed great creativity, covering most of the same threats and consequences as the experts identified and discovering additional ones. One notable difference was that these naïve professionals tend to focus on preventive barriers, leading to requirements for risk mitigation or avoidance, while experienced professionals seem to balance this more with reactive barriers and requirements for incident management.

**Originality/value** – Our results are useful in areas where we need to evaluate safety and security concerns together, especially for domains that have experience in health, safety and environmental hazards, but now need to expand this with cybersecurity as well.

**Keywords** Security, Threats, Bow-tie analysis, Misuse case, Controlled experiment

**Paper type** Research paper



## 1. Introduction

There is an increasingly tight coupling between the cyber and physical world, which leads to new forms of risks that have not been considered adequately, such that the cyber element adversely affects the physical environment (Banerjee *et al.*, 2012). This is typically seen in industries that up until now have been running on isolated platforms and networks but through rapid digital transformations find themselves exposed to hostile cyber attacks from new categories of adversaries, as well as unintentional disclosure of sensitive data. For instance, a *Shodan* search conducted by Trend Micro in 2017 found more than 83,000 industry robots exposed on the internet, whereas more than 5,000 of these had no authentication whatsoever (Maggi *et al.*, 2017). These robots were operating in sectors such as automotive, aerospace, defence, food, and beverages. Similarly, the increased connectivity and lack of security awareness in the shipping industry are making stakeholders worried that this will become the *next playground for hackers* (WMN, 2014). A common trait of these industries is that there are already well-established practices for managing safety concerns. If these practices can be extended to encompass security, we might have an easier path than introducing a set of security analysis techniques that are unfamiliar to them and must be used in parallel.

Security models provide a useful basis for security analysis and requirements elicitation, e.g. supporting comparative evaluations of threats and intended security properties (Bau and Mitchell, 2011). Security modelling comes in many different forms and flavours (Bernsmed *et al.*, 2017), and there is not necessarily one single best or correct approach (Shostack, 2008). In many practical situations, this is a choice depending on factors such as available resources, focus area, domain, level of abstraction and personal preferences, but there is currently little empirical knowledge that can guide us when making these trade-offs. Just as with many other tasks within software engineering, there are many techniques and methods that are used because conventional wisdom suggests that they are the best approaches. As a remedy to this, experiments can investigate the situations to validate whether such claims are true (Pfleeger, 1994). According to Tichy (1998), *experimentation can accelerate progress by quickly eliminating fruitless approaches, erroneous assumptions, and fads. It also helps orient engineering and theory into promising directions.*

Our research objective has been to gain empirical knowledge on the use of bow-tie analysis applied for cybersecurity. Bow-tie analysis has a long tradition from the safety and reliability domain, where identified preventive and reactive barriers are used as sources for eliciting requirements. We wanted to evaluate how well the same analysis technique performs in the context of security, and complements to existing security modelling techniques, such as misuse case diagrams (Sindre and Opdahl, 2001). The research hypothesis central to this work is that *the bow-tie notation has a suitable expressiveness for security as well as safety*, and we have performed controlled experiments with both experienced and aspiring security professionals to get a wider range of people who are representative for the techniques. There already exists evidence that bow-tie analysis performs well for safety considerations, but if the hypothesis is falsified, then applying bow-tie analysis in assessments where we need to consider both safety and security in combination would make no sense.

This paper is an extension of previous work in Meland *et al.* (2018) and is structured as follows. We briefly show related work and explain the history and notation of bow-ties and misuse cases in Section 2, as well as how they can be compared to each other. In Section 3, we explain our research method and the details of the experiments at hand. This is followed by a summary of results in Section 4. These results are then interpreted and discussed as a part of Section 5, and the paper is concluded in Section 6.

## 2. Background

### 2.1 Models covering safety and security

There are many examples in the literature of models that allow combinations of safety and security considerations. For instance, [Johnson \(2011\)](#) shows how to build cybersecurity assurance cases for Global Navigation Satellite Systems (GNSS) using Boolean Driven Markov Processes (BDMP), extending conventional fault trees. [Winther et al. \(2001\)](#) include security as part of HAZOP studies, which is a systematic analysis on how deviations from the design specifications in a system can arise and whether these deviations can result in hazards. [Raspotnig et al. \(2012\)](#) make use of UML-based models within a combined safety and security assessment process to elicit requirements. [Kumar and Stoelinga \(2017\)](#) combine fault and attack trees so that both safety and security can be considered in combination. Fishbone diagrams are similar to bow-ties and are mentioned in [Nolan's book \(2014\)](#) on safety and security reviews for the process industries, but examples here only focus on safety incidents. FMVEA (failure mode, vulnerabilities and effect analysis) ([Schmittner et al., 2014](#)) is a safety and security co-analysis method extended from FMEA (failure mode and effect analysis). Like FMEA, FMVEA proposes to use the STRIDE model ([Shostack, 2008](#)) to identify threat modes first, and then analyse the effect of each threat model. Further examples of methods, models, tools and techniques in the intersection of safety and security can be found in the surveys by [Zalewski et al. \(2012\)](#), [Piètre-Cambacédès and Bouissou \(2013\)](#), [Chockalingam et al. \(2016\)](#), as well as [Kriaa et al. \(2015\)](#).

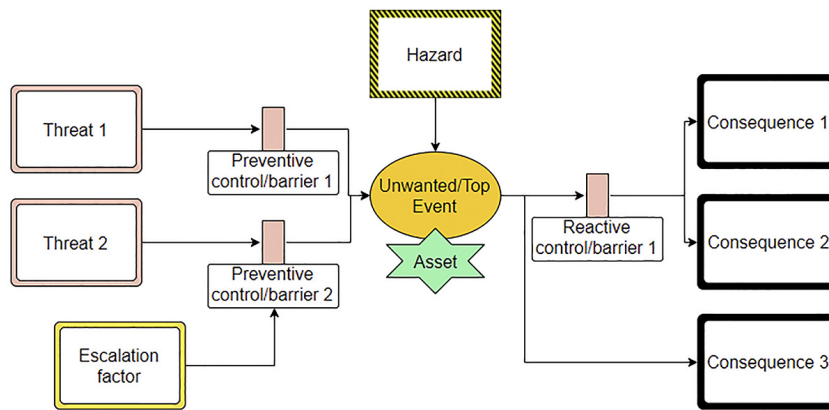
### 2.2 The bow-tie history and notations

Bow-tie analysis has since the 1970s been used by organisations worldwide for risk-management purposes, but primarily to demonstrate control over health, safety and environmental (HSE) hazards ([Lewis and Smith, 2010](#)). For instance, [Khakzad et al. \(2013\)](#) show this application in safety risk analysis in offshore drilling, [Trbojevic and Carr \(2000\)](#), as well as [Mokhtari et al. \(2011\)](#), do the same for safety assessment in international maritime ports, and [Lu et al. \(2015\)](#) apply bow-ties in the context of leakage from natural gas pipelines.

In our modern cybersecurity world, we have to consider the intertwined relationship between safety and security during risk assessment, and make sure that requirements can be traced back to a *source*, such as an intended barrier. As already described by [Bernsmed et al. \(2017\)](#), there have been several efforts at adopting the bow-tie notation for cybersecurity within areas such as engineering environments and maritime operations. This is because these areas are already familiar with the notation from safety assessments, and therefore it is assumed to be easier obtaining community buy-in by evaluating cybersecurity threats in the same way as accident scenarios. [Abdo et al. \(2018\)](#) have also proposed a combined bowtie/attack tree methodology to consider the effect of cyber security on safety risk scenarios. However, we are not aware of any empirical evidence from the literature proving that bow-ties are suitable to cover security concepts in addition to safety.

A central part of bow-tie analysis is the creation of graphical bow-tie diagrams. A bow-tie diagram is something that resembles a fault-tree on the left-hand side with an event-tree on the right ([Lewis and Smith, 2010](#)). [Figure 1](#) gives an overview of the modelling elements that have been included in our experiment, based on [Bernsmed et al. \(2017\)](#). First, the *Hazard* element represents the risky environment in which one or several *Unwanted events* (aka *top event*) can occur but which is also necessary to perform business. Note that we only model one top event per diagram. A *threat* is anything that can potentially cause an unwanted event ([ISO/IEC, 2011](#)), and there can be several types of such threats in a single diagram. To prevent or eliminate threats, we can add *barriers* (aka *controls*) that interfere between threats and the top event. An *Escalation factor* is a specific type of threat that targets a barrier,

**Figure 1.**  
The basic elements of the bow-tie notation with security extension

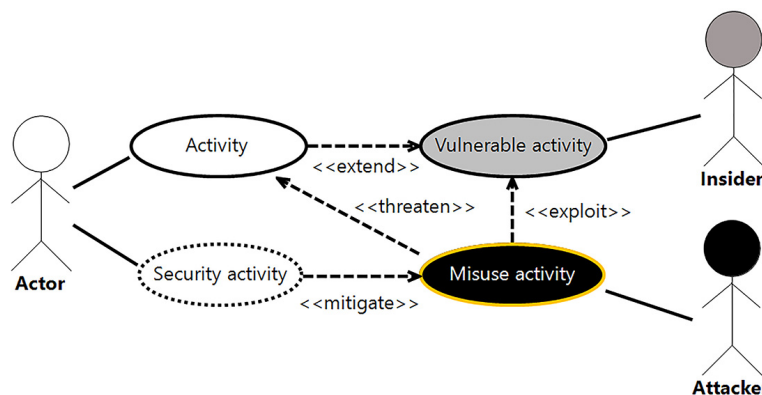


opening up for the original threat. A top event can result in one or several *consequences*. As with threats, we can add *controls/barriers* that can reduce the probability or eliminate the consequences, but these are now of a reactive nature because the top event has already occurred. Finally, and specifically added for security, an *asset* is anything tangible or intangible with value and should be protected. We allow one or more assets to be modelled per diagram.

### 2.3 The misuse case history and notation

Misuse case modelling is a well-known technique for graphical security modelling, and can be summarized as an extension to regular UML use cases (Jacobson, 1993), also covering misuse and used to elicit security requirements (Sindre and Opdahl, 2001). Misuse cases have already been proven useful in different industrial cases when considering security (Matulevicius *et al.*, 2008). They have also been used in controlled experiments to identify safety hazards (Stålhane and Sindre, 2008). Misuse cases are therefore a good basis for comparison with bow-tie diagrams, though one might say that they have an opposite historical path (coming from the security domain and subsequently applied to safety).

The misuse case notation can be summarized as shown in Figure 2. Here, we have included the extensions from Røstad (2008) that also cover vulnerabilities and insiders.



**Figure 2.**  
Overview of the misuse case notation

*Activities* represent the normal behaviour of the system. Normal *actors* instantiate the activities and represent anyone (could be human users but also other systems) interacting with the system as intended, i.e. they do not harm the system either intentionally or unintentionally. *Misuse activities* represent threats towards the system, typically something an attacker would like to perform or achieve. The *attacker actors* have malicious intents towards the system. *Vulnerable activities* are part of the normal behaviour of the system, but represent functionalities that make the system exploitable. *Insider actors* are trusted users of the system that could intentionally or unintentionally cause harm to the system. *Security activities* show what can be done to mitigate misuse activities or vulnerable activities.

Though misuse case models and bow-tie diagrams share some of the same traits, it can be difficult to directly replace one with the other in an analysis. In [Table I](#) we compare them together to show similarities and differences. We would argue that misuse case and bow-tie diagrams are more complementary than competing types of security models, something we have exploited in our bow-tie experiment.

#### 2.4 Related security requirements techniques

Both bow-tie and misuse case diagrams mainly focus on identifying threats, while a key aspect of requirements engineering would then be to specify requirements concerning the necessary level of security in mitigating these threats. An approach closely related to misuse cases would be security use cases ([Firesmith, 2003](#)), which go somewhat further in the direction of requirements rather than threats. Other UML-related approaches that offer more detailed specification of requirements are SecureUML ([Lodderstedt et al., 2002](#)) and UMLsec ([Houmb et al., 2010](#); [Jürjens, 2002](#)) that offer security extensions to several other UML diagrams (e.g. class, activity, sequence diagrams) and not just use case diagrams. Another related approach is the extension to state-transition diagrams proposed by [El-Attar et al. \(2015\)](#).

Bow-tie analysis has been less used in security but could be seen as related to the concept of risk, which is the central focus of the modelling language proposed by [Mayer et al. \(2007\)](#). Other well-known approaches to security requirements modelling include goal-oriented

**Table I.**  
A comparison of misuse case models and bow-tie diagrams

Misuse case models	Bow-tie diagrams
[Both] Defined by a simple to understand graphical notation with an open-ended method, allowing for a lot of creativity by the modeller	
Originate from computer security and requirements engineering, based on UML use case diagrams	Originate from the safety and reliability domain, related to fault analysis
Developed to identify malicious actions (misuse) for a given system	Developed to investigate accident scenarios and define barriers
The misuse activity element represents an unwanted event (something that threatens regular activities)	The top event element represents an unwanted event
Broad scope. Suitable for describing many different misuse activities in a single model	Narrow scope. Focus on a single unwanted top event per diagram
Show actors (attackers, misusers, threat agents) related to misuse activities	Do not represent actors, but in which risky environment (hazard) the top event can occur
Mitigations are modelled as security activities	Mitigations are modelled as barriers, which are clearly defined as either preventive or reactive
Can depict vulnerable activities that a can be exploited	Represent various threats/causes that can lead to the top event
Consequences are not part of the model	Explicitly depict possible consequences following the top event



approaches such as KAOS (Van Lamsweerde, 2004), Secure i\* (Elahi *et al.*, 2010) and Secure Tropos (Mouratidis and Giorgini, 2007). There are also many other security requirements techniques, beyond what can be covered in this section. Good overviews of techniques that existed by 2010 are provided by Fabian *et al.* (2010) and (Mellado *et al.*, 2010) and a more recent mapping study by Souag *et al.* (2016) focusing specifically on reuse of knowledge in security requirements engineering.

### 3. Experiment method

To plan our bow-tie experiment, we adopted and applied the guidelines by Kitchenham *et al.* (2002), originally designed for empirical studies in software engineering. The form of the study is a *controlled experiment*, which is a scientific method for identifying cause – effect relationships (Sjøberg *et al.*, 2005), and a means to *acquire general knowledge about which technology (process, method, technique, language or tool) is useful for whom to conduct which tasks in which environments*. The intervention we introduce is the use of the bow-tie notation for security analysis on two sample population that are both working on the same case.

As there are no random assignments, this should be classified as a *quasi-experiment*, and as a formal experiment because we have a high level of control over the variables that can affect the truth of the hypothesis (Pfleeger, 1994).

One of the sample populations consists of students, and therefore it has been important to make sure that they perceive a value from participation (Carver *et al.*, 2004). By carefully scoping the case of the experiment and having an approach that is new to the student sample and professionals in general, we expect to get relevant results with external validity (Salman *et al.*, 2015).

To have a better basis for experiment evaluation, we present the result from a similar experiment with misuse case modelling, though applied to a different case. This allows us to see whether the phenomena related to the dual populations are generalisable or local to bow-ties. Both cases in focus and experiment setups are described in the sections below.

#### 3.1 Case A: digital exams

For the bow-tie analysis, we chose a security modelling assignment related to use of digital exams, something that is rapidly growing in popularity at universities and other educational institutions. Here, exams are created, solved and graded using online systems. This is meant to be more efficient than traditional exams done on paper, however, relies on technology and opens up to new types of threats that need to be identified and dealt with. For instance, a survey by Chen and He (2013) shows that there is a great diversity of security risks for online exams, nevertheless, security is not considered as a top priority among learning providers and practitioners. Additionally, there is evidence that both digital and “analogue” exams suffer because of new technical ways of cheating. According to *The Guardian* (Marsh, 2017), there has been a 42 per cent rise in cheating cases between 2012 and 2016, involving gadgets such as mini cameras and micro earbuds. London (2017) gives an overview of further inventive and not-so-inventive ways that have been used for cheating on online exams. In some developing countries, such as Algeria, Ethiopia, Syria and Iraq, internet access in the whole country is shut down during the exam period to prevent cheating (Bradbury, 2018). All in all, a case related to digital exams provides an interesting and relevant arena for looking at security issues and possible solutions.

In our case, there are many students participating in the exam in the same confined room and within the same time frame. This is a bit different from other types of digital exams, which can be done from home and at any given time. Furthermore, the students are allowed to use their own personal computers with internet access through WiFi, but are not allowed

---

to use supporting materials, such as curriculum books and notes. A specific Web browser must be installed on their computers, known as the *Safe Exam Browser* (SEB)[1], which regulates access to websites, search engines, other applications and system calls, also referred to as *browser lockdown*. Vegendla *et al.* (2016) report on a case study doing penetration testing on the SEB, identifying some vulnerabilities that could be used for cheating. However, it must be noted that this cheating is less likely today, as the software has since been improved.

### 3.2 Case B: Web shop for digital goods

For the misuse case modelling, we selected a system description that most people can relate to through personal experience. Web shops are virtual marketplaces that are accessed online and used to browse for interesting items and complete purchases. Web shops are suitable for security analysis because there have been plenty of examples of real-life compromises. For instance, a 2018 report by a security firm show that almost 90 per cent of the people logging into some popular retailers' e-commerce sites were hackers using stolen data (Green, 2018). A lot of the Web shops also use the same code base; hence, they share a lot of the same vulnerabilities. If the store owners are lazy updating their software, they quickly become easy prey to attackers looking for known vulnerabilities. In 2016, a Dutch developer reported almost 6,000 Web shops with proven vulnerabilities and that were exploited to steal the credit card details of customers (BBC, 2016). The OWASP Juice Shop Project (Kimminich, 2018) is an example of an intentionally insecure Web shop that is being used to train software developers.

In our case, we limited the Web shop inventory to be digital goods, which are non-tangible items, such as music files, wallpapers, games and other types of software that are directly downloaded from the Web shop. The main assets are the digital goods and customer information such as personal data, order history and credit card information. There could be a number of different threat agents/attackers with different motivation, such as cyberthieves or business competitors.

### 3.3 Bow-tie experiment setup

This experiment engaged two types of populations: a small sample of security experts and larger sample of computer science MSc graduate students. The characteristics of these groups can be described as follows: the students participated in the experiments as a part of a classroom exercise in a course on secure software engineering and were motivated to learn security modelling to apply such techniques for their exercises and final exam. Before the experiment, the students had taken several lectures, including security concepts and principles, OWASP top 10, crypto introduction, multilevel security and multilateral security. The students had limited knowledge of security modelling beforehand and no experience at all from bow-tie modelling. Moreover, the students had significant practical experience related to digital exams, as they had already been exposed to this on several occasions. It is unknown how experienced and reflected they were related to cheating.

The security experts had a great deal of prior knowledge and practical experience in various types of security modelling techniques, and in particular bow-tie for specific domains. In contrast to the students, the experts had limited practical experience of participation in digital exams, though one of them was skilled with setting up exams using the online system. The experts were motivated by the research itself, and the desire to create a good reference model that the student results could be compared to.



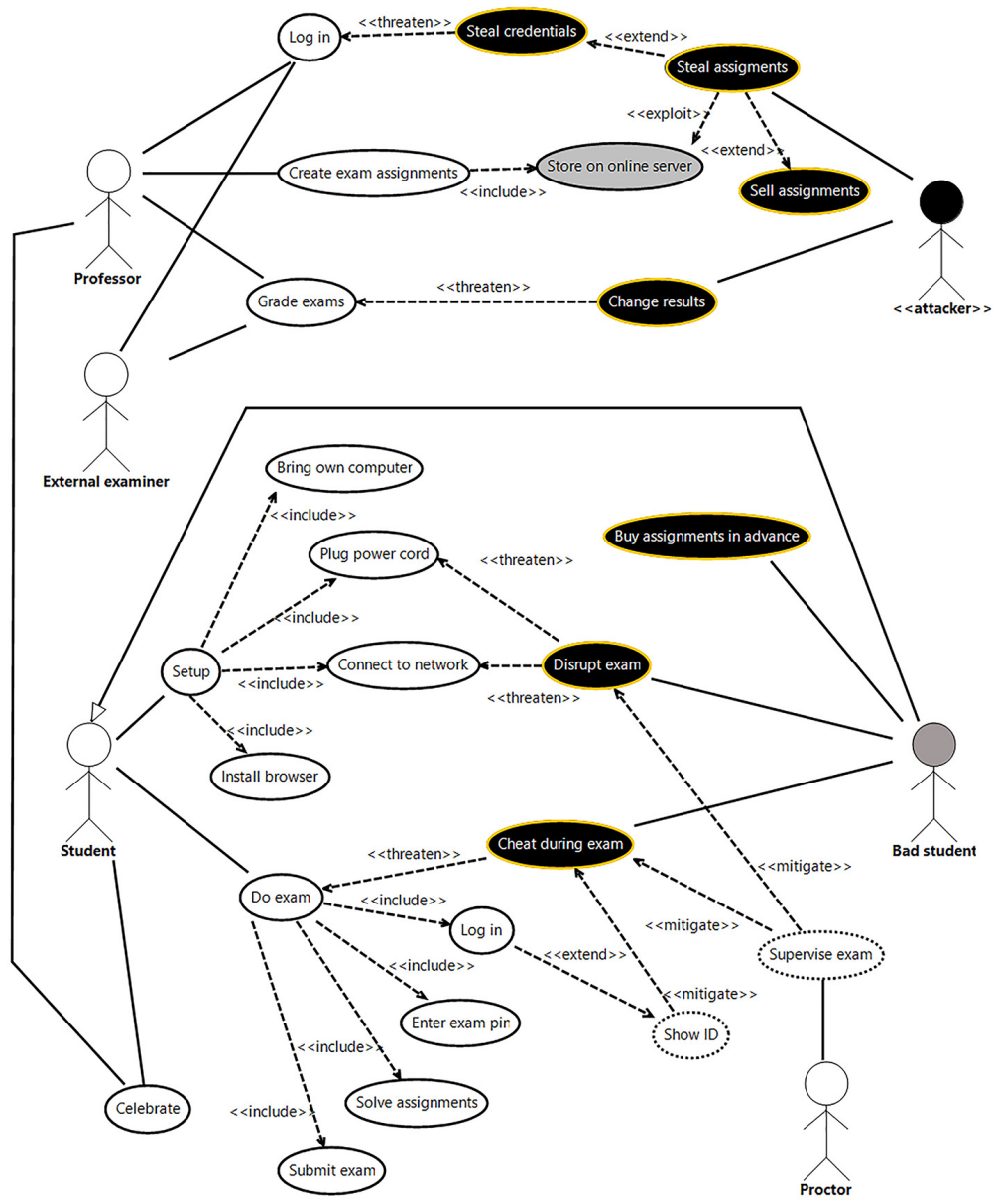
As an introduction, the students were given a lecture on threat modelling, including the misuse case and bow-tie notations. As we know from prior experiences, one of the challenges of bow-tie diagrams, is setting the scope of the unwanted event. Therefore, the students were presented with a misuse case model that we hoped would better define the scope and the relationship between the events. This model is shown in [Figure 3](#) and depicts a number of actors and typical activities related to digital exams, as well as misuse case activities and associated threat agents. For example, the actor *professor* will need to *log in* to the system and *create exam assignments* prior to the examination day. An external *attacker* actor would possibly want to *steal assignments* and maybe sell this online to students who want to cheat. After the examination day, an additional *external examiner* is involved in the process of *grading exams*. The attacker could at this point try to *change the results* of the exam. During the examination day itself, the main legitimate actor is the *student* who needs to *setup* his/her computer, which also involves sub-activities such as *connecting to the network* and *installing the correct SEB software*. To *do the exam*, the student must authenticate by *logging in*, *enter the exam pin* for this particular exam, *solve the assignments* and finally *submit the exam*. On the right side of the diagram, we have depicted a *bad student* insider actor who inherits all the activities from the legitimate student actor. The bad student has a misuse activity mostly relevant prior to the examination day, which is to *buy the assignments in advance*, and two others that threaten the regular activities during the exam. The first one, *disrupt exam*, is basically a way of sabotaging the examination for everyone, possibly motivated by a wish of cancelling/delaying the exam. The second one is *cheat during exam*, which a student would do to illegitimately improve his/her grade. The *proctor* is a type of examination guard that *supervises the exam* and is there to mitigate cheating attempts and disruptions.

The next step of the introduction was to show how a misuse activity can be detailed as a bow-tie top event. This was demonstrated with *disrupt exam* as shown in [Figure 4](#). In this model, there are a number of threats that can lead to a disruption, such as *tampering with the fuse box* to cause power outage, *jamming the wireless network* or performing some other action to *make the online server unavailable*. The assets that needs to be protected are the *network*, the *SEB software* and the *physical premises* themselves. We added some example preventive controls/barriers, such as *locking the fuse box cabinet* and having a system *mirror site* on hot standby. In terms of disruption consequences, computers can stop working and the bad student can be expelled. The only reactive control/barrier shown here is *switching to paper* to complete the exam.

Having introduced the notation, defined the scope and given examples, the populations were now ready to work on their own diagrams. We predefined *digital exam* as the risky environment, *cheat during exam* as the top event and the asset *answers* as a starting point.

Both populations worked on this same case, with access to external information such as SEB documentation and articles about online exams and cheating. The students worked in teams, typically two-three persons per model, spending about 30 min on their task, and were observed by two of the authors of this paper. The experts worked independently of each other for about one hour. Both populations used an online modelling tool[2] to create their models. The tool itself has an intuitive drag-and-drop interface for the basic bow-tie elements, and runs within any Web browser. A screenshot of this tool is shown in [Figure 5](#).

The students were informed that all participation was anonymous and voluntarily, and that we wanted to make use of the result to evaluate the bow-tie notation for security.



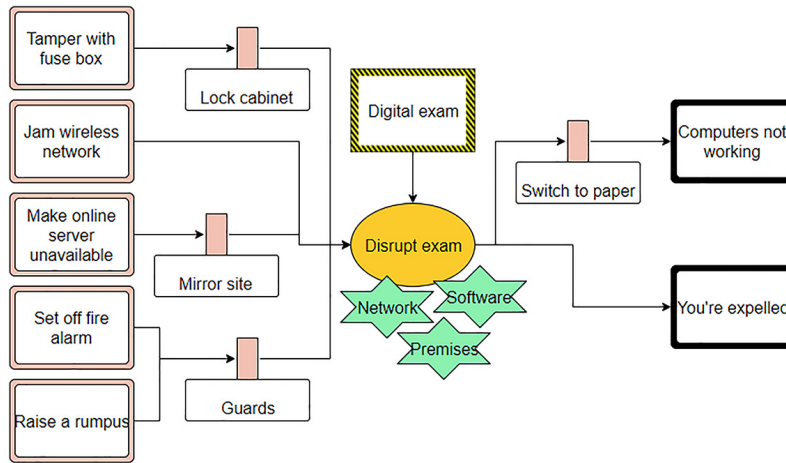
**Figure 3.**  
Defining the scope  
with a misuse  
case diagram

*3.4 Misuse case experiment setup*

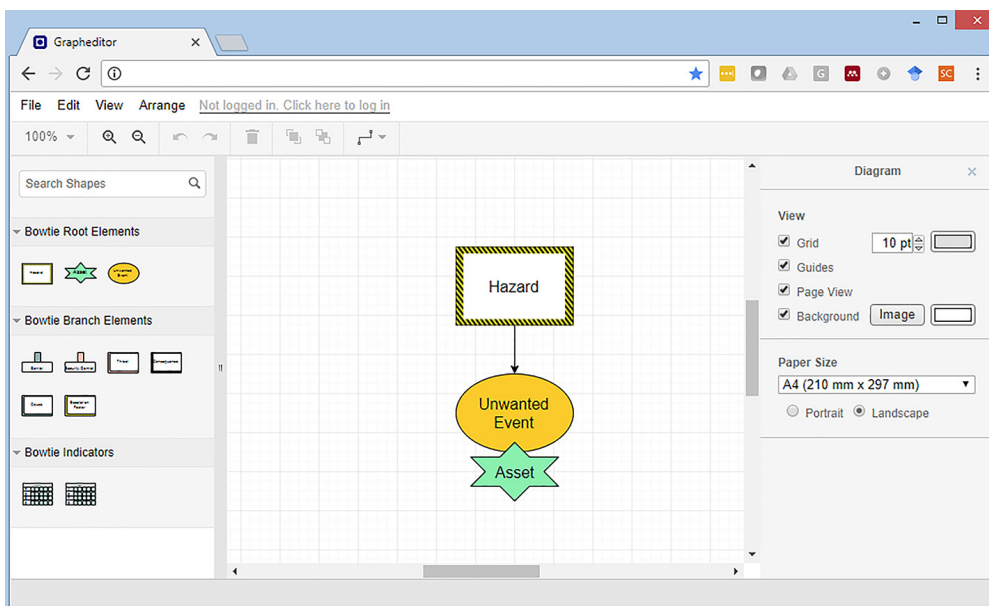
The setup of this experiment was almost identical to the bow-tie experiment, with a few notable exceptions:

- The resulting models had already been created by students taking the same course during three previous years. Hence, it was not the same population of students but a larger set of students with the same characteristics. The experts were the exact same individuals as in the bow-tie experiment.





**Figure 4.** Example model showed as a preparation



**Figure 5.** The online tool used for making the bow-tie diagrams

- The students had some familiarity with UML use case modelling, but no significant experience with misuse case modelling beforehand.
- Both populations used pen and paper to create their models and not an online modelling tool.

The students had been given the same kind of introduction to security modelling, including a walkthrough of misuse case analysis with a few simple example models. When the assignment started, the students were handed out a paper sheet containing a use case template for them to work on in pair-wise groups. This template is shown in Figure 6, and depicts two normal actors, a *customer* and the *Web shop* service. These are associated to a set of



**Figure 6.**  
The Web shop use  
case template



predefined use case activities defining the functional scope of the assignment. An *attacker* actor was also included in the template, but with no associated misuse activities. The use case activities indicate what kind of assets that are involved, such as personal data, order history, shop items, credit card information and product reviews. Both populations were instructed to add misuse case model elements to the template and spent about 30 min on their models. Just as with the bow-tie experiment, all participation and hand-in was voluntarily and anonymous.

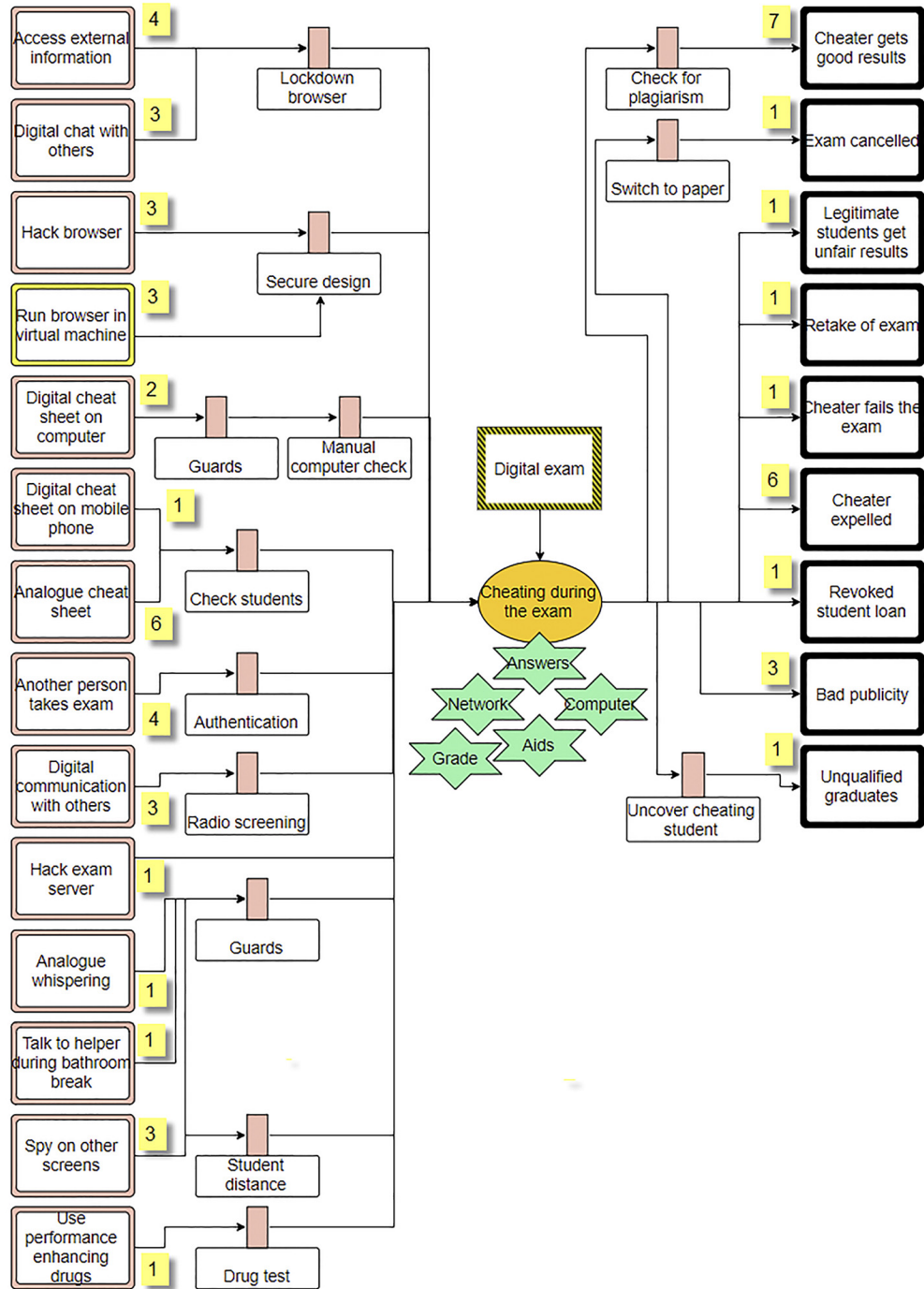
## 4. Results of the bow-tie analysis

### 4.1 Diagrams made by students

A total of 40 students were present in the experiment session, which resulted in 13 different diagrams. Observations from the classroom indicated that approximately 30 students contributed to these diagrams. This estimate is based on the average size of the groups and that we also know that not all diagrams were submitted (this was voluntarily). The diagrams were then analysed, and we created a small taxonomy of threats, controls/barriers and consequences to be able to compare them. Based on this, we developed a combined bow-tie diagram, shown in [Figure 7](#), which also indicates the frequency of the threat and consequence elements found in the diagrams made by the students. As can be seen from the figure, the top threats were:

- *Analogue cheat sheet*, the most popular threat, appeared in 6 out of the 13 models that we collected (6/13). This is probably the most “traditional” way of cheating, and it involves the use of some concealed written material, e.g. paper notes hidden inside the wrapper of a candy bar or somewhere on the body of the student.
- *Access external information* (4/13) encompasses using the computer to search and access information on the internet.
- *Another person takes exam* (4/13) is related to impersonation and not something that is unique to digital exams.
- *Digital chat with others* (3/13) is when the student computer is used to communicate with others in the same room or on the outside.
- *Hack browser* (3/13) is done by somehow modifying the source code or exploiting an existing vulnerability in the SEB software to disable the lockdown functionality.
- *Run browser in virtual machine* (3/13) was represented as a threat in two of the models, and as an escalation factor in a third. In the combined model, we represent it as an escalation factor because this is basically a way of circumventing a preventive barrier by letting the SEB software lockdown the virtual machine instead of the computer itself.
- *Digital communication with others* (3/13) covers all kinds of gadgets besides the student computer that are used for communication with others. This typically includes Bluetooth devices and other radio equipment.
- *Spy on other screens* (3/13), also denoted as “shoulder surfing”, is simply ways of looking at other people’s answers without them noticing it. Peeking at the answers of others is a relevant cheating threat for paper exams too, but may be accentuated for digital exams because screens are nearly vertical, while paper lying horizontally on a desk is harder to read from a distance.

Some additional threats can be found in [Figure 7](#), but these were only present in one or two of the diagrams. Additionally, we discarded three threats that were out of scope for this top event, namely, *Retrieve exam answers beforehand*, *Disrupt exam* and *Blackmail professor*.



**Figure 7.**  
A combination of the  
bow-tie diagrams  
made by the students

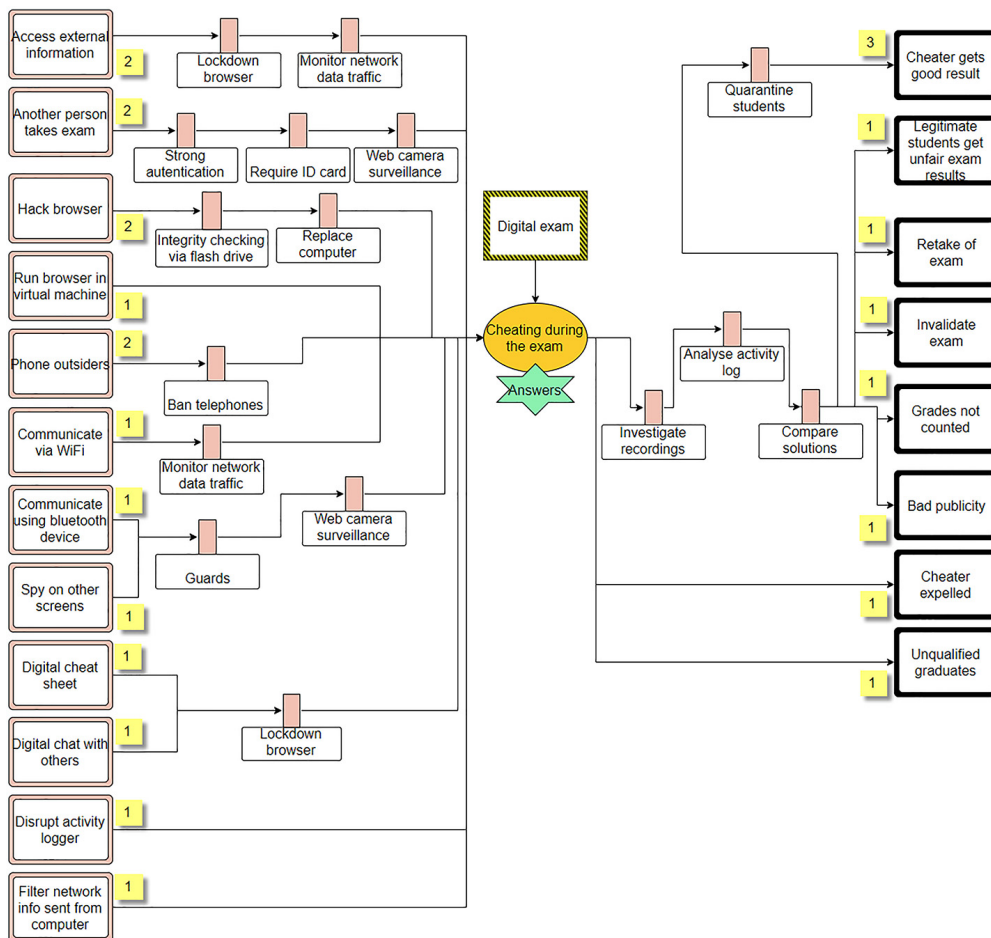


On the consequence side of the diagram, *Cheater gets good results* (7/13) was most prevalent, followed by *Cheater expelled* (6/13) and *Bad publicity* (for the university). Interestingly, these are consequences for both successful cheating as well as consequences for the cheater if he/she gets caught.

The combined diagram does not show the frequency of barriers/controls because a lot of them overlap over more than one threat/consequence. We also noticed that some of the diagrams (4/13) contained additional assets, so we added these to the combined diagram as well (Figure 7).

#### 4.2 Diagrams made by security experts

There were three security experts participating in this experiment, resulting in three independent bow-tie diagrams. These were analysed in the same manner as the student diagrams and aligned using the same taxonomy. The resulting combined diagram from the experts is shown in Figure 8. There were only four threats that had an overlap between the expert diagrams; *Access external information*, *Another person takes exam*, *Hack browser* and *Phone outsiders*. The three first were all present among the top threats from the student diagrams as well, while the latter was not. We discarded one threat from the diagram,



**Figure 8.** A combination of the bow-tie diagrams made by the experts

ICS  
27,4

550

*Introduce vulnerability in SEB OSS project*, as this is something that must be done prior to the exam and hence out of scope for this top event. The expert and student diagrams shared their top consequence, namely, *Cheater gets good result*. Besides from that one, there was little overlap between consequences among the experts. Note that there are several threats and consequences that are without any barriers. It turned out that one of the experts forgot about adding these, and therefore spend more time on finding threats and consequences compared to the others.

Table II shows a numerical comparison of the diagrams created by the two populations. The last row shows how many distinct elements that are common between the combined diagrams from each population. As the level of detail vary, it was not possible to always create direct mappings. Therefore, *Communicate via WiFi* and *Communication using Bluetooth device* in the expert diagram is mapped to the single threat *Digital communication with others* in the student diagram. Likewise, the preventive barrier *Strong authentication* in the expert diagram is mapped towards the less strict *Authentication* in the student diagram. The complete set of original diagrams is openly available from Meland (2018a).

## 5. Results of the misuse case experiment

### 5.1 Models made by students

Because we collected misuse case models made by students from three previous years, the total number of models was increased to 31 in total. The classroom setup had been the same, with two and two students sitting together, so approximately 62 students contributed to these models. During the analysis, we grouped together similar model elements and created a taxonomy of misuse case activities, vulnerable activities and security activities. As misuse case models by nature have a broad scope, we also got a very broad set of misuse case activities (48 distinct), where the top 20 were:

- (1) *Trojan/corrupt code in digital goods* (22/31) targets other customer of the Web shop.
- (2) *Disrupt Web shop service* (13/31) for instance using DDOS/DOS attacks.
- (3) *Phishing* (13/31) is an attack technique used to obtain sensitive information.

Measurement	Experts	Students
Number of participants	3	~ 30
Number of models	3	13
Total number of threats	18	49
Number of distinct threats	12	14
Average number of threats per model	6	3.8
Total number of consequences	10	27
Number of distinct consequences	8	9
Average number of consequences per model	3.3	2.1
Total number of preventive barriers	16	41
Number of distinct preventive barriers	10	9
Average number of preventive barriers per model	5.3	3.2
Total number of reactive barriers	6	6
Number of distinct reactive barriers	4	3
Average number of reactive barriers per model	2	0.5
Common (threats/consequences)/(preventive/reactive) barriers		7/5/3/0

**Table II.**  
A numerical  
summary of bow-tie  
model elements

F



- 
- (4) *SQL/Code injection* (12/31) is an attack technique typically used wherever there is some kind of user input, e.g. Web forms.
- (5) *Write false review* (10/31) can manipulate the user ratings of the digital goods, potentially influencing what the customers buy.
- (6) *Pharming* (10/31) is a general concept where the goal is to steal and collect other customer's payment data.
- (7) *Spoof payment* (9/31) tricks customers into paying at a false payment service.
- (8) *Information theft* (9/31) is a high-level concept where the attacker obtains information that should have been protected.
- (9) *Spoof Web shop* (9/31) tricks the customers into using an imitated, fake version of the Web shop.
- (10) *Steal account/password* (8/31) enables an attacker to impersonate a legitimate user.
- (11) *Malicious input in review form* (7/31) encompasses active code or offensive content inserted into the review form functionality of the Web shop.
- (12) *False/fake signup* (7/31) is when non-existing or impersonated users are registered to the Web shop.
- (13) *Drive-by download* (7/31) is when the information belonging to the digital goods have been replace by or contains malicious code.
- (14) *Man-in-the-middle* (7/31) is a general concept where requests are intercepted and manipulated before they reach their destination.
- (15) *Eavesdropping/sniffing* (7/31) means to tap into some communication. This can be considered a sub-type of information theft.
- (16) *Social engineering* (6/31) is when humans are exploited rather than technical systems. It is mostly associated with the *Contact shop* activity.
- (17) *Manipulate payment* (6/31) threatens the integrity of the payment transaction, for instance the target account or the sum of the payment.
- (18) *Send fake order confirmation* (6/31) targets the customers, typically used for delusion or click-bait.
- (19) *Modify/delete other customer's profile* (6/31) is a broad category, e.g. involving changing personal information and password, as well as replacing the profile picture with explicit photos.
- (20) *Change contact info* (6/31) of the Web shop, subsequently tricking the customers into unknowingly contacting fraudsters.

In addition to these, we would highlight *Pay with stolen credit card* (5/31), *Steal/copy digital goods* (4/31) and *Trolling* (2/31) as particular relevant for the Web shop case. Ten misuse activities only appeared once in the models but were still relevant.

There were noticeable differences in the level of abstraction used to describe the misuse activities. We could have created a smaller taxonomy with more generalized activities, but this would mean loss of some of the more specific information. For instance, *Information theft* is a very broad concept that some of the students used. *Steal account/password* or *Pharming* are more specific and connected to the use case activities at hand, while *SQL/Code injection* is an attack technique that can be used to accomplish information theft (among other things).

The number of distinct vulnerable activities was only seven, making an average of only 0.3 per model. This number is a bit skewed, as 24 of the models had no vulnerabilities at all. The vulnerable activities appearing in more than one model were:

- Retrieve personal information (2/31);
- store private data (2/31); and
- input forms (2/31).

There were 19 distinct security activities, but only 5 that appeared in more than one of the models. These were:

- (1) *Network encryption* (8/31) protects the integrity of the communication through SSL/TLS mechanisms.
- (2) *Checksum for digital goods* (4/31) is used as a mean to verify the integrity or authenticity of the downloaded items from the Web shop.
- (3) *Email verification on profile alterations* (4/31) is used to inform the customer about possible misuse of the account.
- (4) *Sanitize input* (4/31) is used to check user input and mitigate injection attacks.
- (5) *CAPCHA* (2/31) can be used to mitigate bots creating false user accounts or inserting malicious content/spam into the review forms.

Without going too much into details, it is safe to say that that the students mostly focused on misuse activities, and not so much on identifying vulnerable or security activities. [Figure 9](#) shows a combined model with the top 20 misuse case activities and vulnerable/security activities that appear in more than one model. Some relationships have been omitted for readability purposes.

### 5.2 Models made by security experts

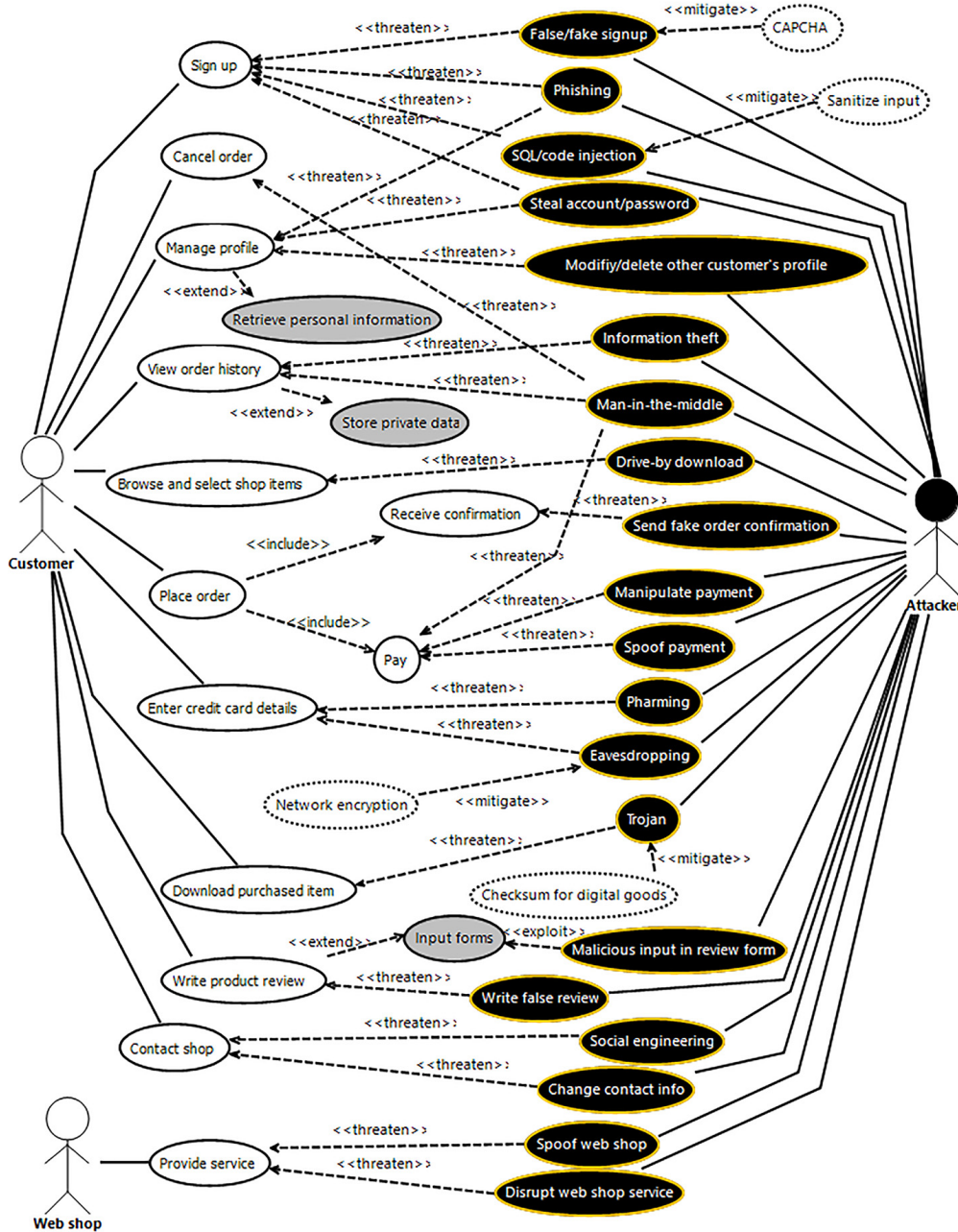
The same three security experts as in the bow-tie experiment participated and created their models individually. The experts identified 18 distinct misuse case activities, whereby 17 of these were also covered by the students. The last one was *Enumerate usernames* (1/3), where the goal is to harvest existing usernames by misusing the *Sign up* activity. These usernames can have a value for an attacker because they give away information about the customer base, and they can also be used in brute force passwords attacks, sending out phishing emails or locking out other users. [Figure 10](#) shows a combined model based on the results from the experts. Some of the relationships have been omitted to increase readability.

The average number of misuse case activities per model was 9, which is only slightly higher compared to the student models. As can be seen in the [Table III](#) summary, it is for the vulnerable and security activities there are significant differences between expert and student models. In fact, the expert had an average of distinct vulnerable activities more than 22 times compared to the students, and more than 7 times for distinct security activities. The complete set of original models is openly available from [Meland \(2018b\)](#).

## 6. Discussion

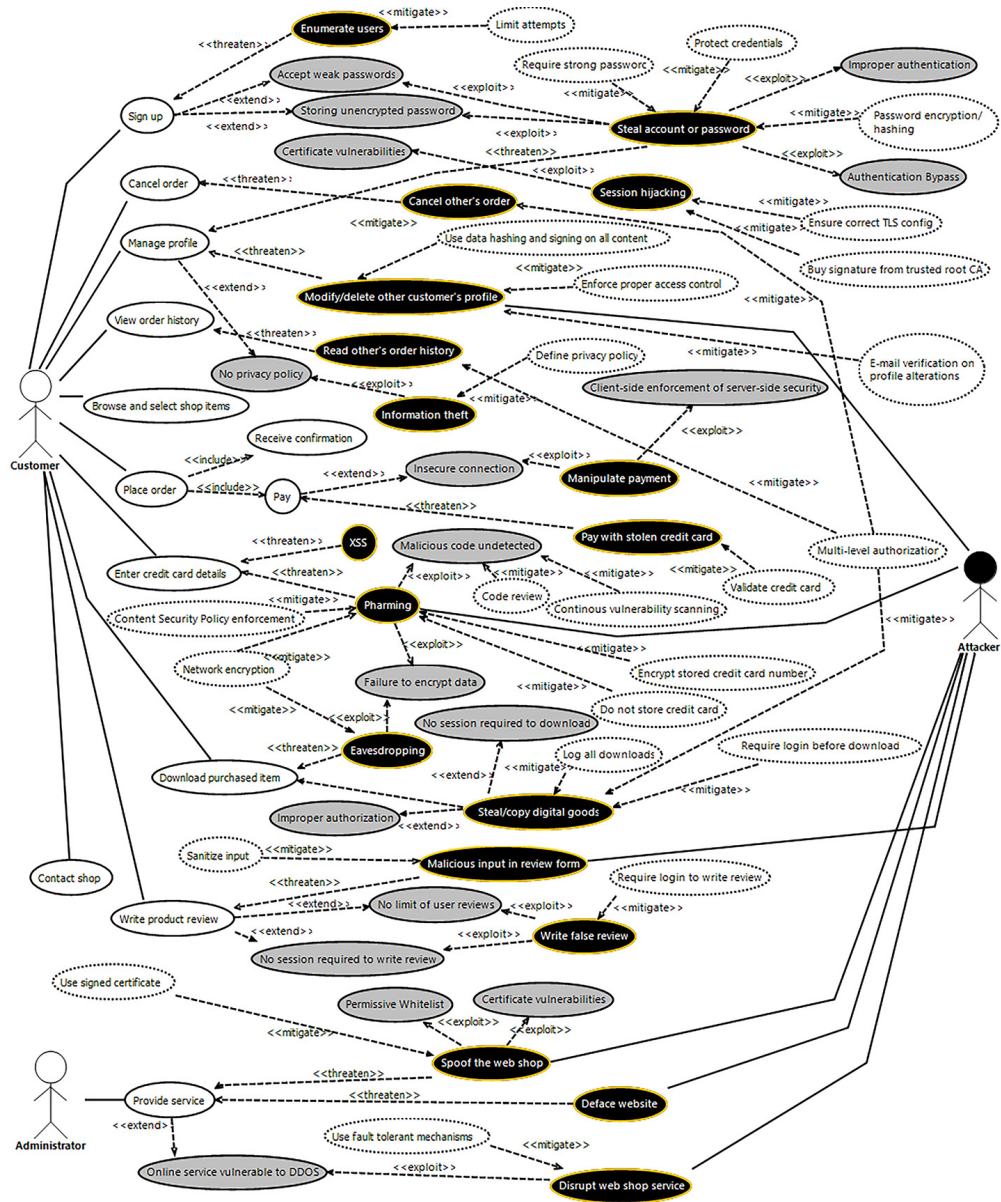
### 6.1 Interpretation of bow-tie experiment results

It was interesting to see how well the students were able to grasp the concepts of bow-tie modelling and apply it to the digital exam case after just a relatively short introduction. There are a few notable differences when comparing results from students with experts, such that the average numbers of threats, preventive barriers and consequences per model



**Figure 9.** A combination of the misuse case models made by the students

are all about 60 per cent higher for the experts. This is to be expected, as the experts had a deeper security knowledge and did also have some additional time for developing their models. The number of reactive barriers was clearly higher for the experts, but this is in line with a general observation that the students tended to focus on the left side of the diagram. In fact, 3 of the 13 models from the students had no elements on the right side whatsoever.



**Figure 10.**  
A combination of the misuse case models made by the experts

Another significant difference was that two of the experts modelled two or three barriers for most of their threats, while this was not observed in any of the student models where all threats had just a single control/barrier. This can be interpreted in two ways; the students did not fully understand that the tool supported adding more than one barrier per threat, or the students did not think that it is necessary to implement more than one barrier per threat in a real system. The last experts did, as mentioned above, not model any barriers, and this skews the average barrier per threat significantly. Identifying a wide range of barriers is

Measurement	Experts	Students
Number of participants	3	~62
Number of models	3	31
Total number of misuse activities	27	251
Number of distinct misuse activities	18	48
Average number of misuse activities per model	9	8.1
Total number of vulnerable activities	20	10
Number of distinct vulnerable activities	16	7
Average number of vulnerable activities per model	6.7	0.3
Total number of security activities	26	37
Number of distinct security activities	24	19
Average number of security activities per model	8.7	1.2
Common (misuse/vulnerable/security) activities		17/1/5

Bow-tie  
analysis for  
security

**555**

**Table III.**  
A numerical  
summary of misuse  
case model elements

considered one of the primary advantages of bow-tie modelling, and we have made a note to encourage this a bit more in later work.

When we consider the students as a collaborative group, the numbers of the distinct threats, consequences and both types of barriers are almost identical to what the experts produced. When we look beyond these numbers and compare the type of elements in the taxonomy, there is a clear tendency for the experts to focus on technical threats and threats that are specific for digital exams, while the students have included more of the traditional ways of cheating. We believe that both these inputs can be important, and advocate for a combination of security experts and end-users (in our case, the students) when developing these kinds of security models, and consequently defining requirement based on barriers.

Our general impression is that the students showed great creativity, covering most of the same threats and consequences as the experts identified, and discovering additional ones as well. The bow-tie notation did not seem like an obstacle for expressing this, which confirms our hypothesis that the bow-tie notation has a suitable expressiveness for security as well as safety issues. The students also identified additional elements on the consequence side that the experts had not thought of, even though it seems like the students spent most of their time on the threat side. The students seemed just as good as the experts at staying inside the scope of the top event, something we believe can be attributed to the misuse case presentation in the introduction of the experiment.

### 6.2 Comparing results from bow-tie and misuse case experiments

Having performed similar experiments with different modelling techniques allowed us to see if some phenomena can be generalized or if they are specific for the technique at hand.

Just as with the bow-tie threats, the combined mass of students was able to identify a broader set of misuse case activities. This might not be a surprise given that the number of students was much greater, especially in the misuse case experiment. However, this observation was not consistent when it comes to the other model elements, such as vulnerable and security activities. This phenomenon of imbalance in the models made by the student population was present in both experiments, but that does not necessarily make the models poor compared to the ones from the experts. In both case A and B, the students had domain knowledge and practical experience as service end users, and were able to imagine lots of “bad stuff”, i.e. malicious or deviant behaviour. The security experts had (and should have) a better repertoire of common pitfalls and security solutions from years developing real-life systems, and this became visible in both cases as well.

---

Looking at the average number of model elements that were added to each model, this number was exactly the same for the student populations in both experiments (9.6). The experts had an average of 16.6 for the bow-tie diagrams and 33.7 for the misuse case models. This is a bit of a surprise because they had more time than the students on the bow-tie experiment, and the same time as the students on the misuse case experiment. This phenomenon can be explained by the fact that the students were almost equally inexperienced to both modelling techniques, while the experts were more experienced to misuse cases than bow-ties. Hence, for inexperienced users, there is no reason to believe that misuse case models outperform bow-tie diagrams in a security context if we consider content generation in isolation.

### 6.3 Limitations and threat to validity

There are several factors to consider regarding the validity of these experiments. Convenience sampling is a threat to a lot of experiments that involve a population consisting of students, as this can come at the cost of low external validity, but we argue that our samples already had taken an interest in security and represent an aspiring group of people that are likely to work with security engineering in their professional careers. According to a survey on controlled software engineering experiments by [Falessi et al. \(2018\)](#), there are pros and cons with both the use of professionals and students, and it is impossible to state that one is always better than the other. Studies by [Salman et al. \(2015\)](#), [Svahnberg et al. \(2008\)](#) and [Höst et al. \(2005\)](#) show that there is little difference in performance between these groups, especially for graduate students ([Runeson, 2003](#)).

Though the participation was voluntarily and anonymously, the students seemed motivated and we did not see any submitted diagrams with frivolous content. Furthermore, it was in their own interest to get some relevant experience in security modelling for their course exercises and final exam.

The time that the students had available for the analysis was very limited. In real life, a thorough bow-tie analysis would include defining a series of top events within the same risky environment, and there would be several iterations on each model to improve their coverage and quality. We have tried to address this by letting the students collaborate directly, and by spending time in the introduction on defining a narrow scope for a single top event. Alternatively, we could also have given different top events to different groups and thus have a wider analysis, but that would impose limitations to the comparison afterwards.

Another limiting factor is that we did not perform any systematic user evaluations. Our evidence is thus solely based on the resulting diagrams, aided by observations and comments received during the experiments. For future work, this can be done in several ways, e.g. with standardised usability surveys or adopt from the *Information Systems (IS) field* Moody's *Method Evaluation Model* (2003) that combines measurable constructs such as effectiveness, perceived usefulness and ease of use, intention to use and actual usage ([Moody, 2003](#)). Another approach could also be to engage participants in interacting focus groups where they more freely discuss their opinions.

As reported in a previous work ([Bernsmed et al., 2017](#)), there are more informal evaluations of situations that combine both safety and security within the same bow-tie diagrams. Though this would have been desirable to try out in this experiment as well, we chose to focus on security issues as we could not find a suitable case where the student would have enough domain knowledge to consider safety, in addition to security.

A final note is related to the different setups of the two experiments. For instance, the modelling cases were not the same, and the bow-ties were made with an online tool, while

the misuse cases made with pen and paper. Such elements make direct comparisons more questionable, but as we pointed out in section 2.3, the modelling techniques have different natures, and we do not intent to prove that one is better than the other. We have rather tried to verify that bow-ties used in a security context do not suffer from significant penalties compared to an established security modelling technique.

#### 6.4 Further research directions

Both misuse case models and bow-tie diagrams are high-level modelling techniques, and in their basic forms they are not concerned about attack sequences, relationships between threats, or attributes such as costs and likelihood. Attack (-defence) trees (Schneier, 1999), (Kordy *et al.*, 2014) can for instance be used to further drill down the details of how the unwanted event/attacker goal can be realised, but there is a need to obtain more practical knowledge about what level of granularity and level of detail to represent with various security modelling techniques, and when we should switch between them.

In both experiments, the students and experts did not attempt to transform the bow-tie barriers or misuse security activities into well-defined security requirements. In addition, prioritisation would be the next step of this process, but that would require quantification of risk and mitigation costs. Both these steps are natural continuations that we would like to follow up.

The bow-tie modelling tool itself was not something we set out to evaluate as a part of this study, but observations and comments suggest that the built-in support for creating and connecting the right elements together was helpful indeed. With the misuse case models, which were drawn by hand, the semantics were less strict and the content more difficult to interpret afterwards.

In our bow-tie experiment, the collaborating students were sitting closely together using the same computer, but it would be interesting to see how well a Web-based tool can facilitate online collaboration. Our tool has already built-in functionality for sharing diagrams between users, as well as getting a quick start by importing templates made by others. During the analysis, it also occurred to us that an online voting mechanism could help create consensus about which threats, consequences and associated barriers should be prioritised.

## 7. Conclusion

Our research hypothesis has been that the bow-tie notation has a suitable expressiveness for security as well as safety, and our results go a long way in verifying this. One of the main strengths of bow-tie analysis is the identification of preventive and reactive barriers, which can be used as traceable sources for the following requirements elicitation process. Naive professionals might tend to focus on preventive barriers, leading to requirements for risk mitigation or avoidance, while experienced professionals seem to balance this more with reactive barriers and requirements for incident management.

Our results are useful in areas where we need to evaluate safety and security concerns together, especially for domains that have experience in HSE hazards, but now needs to expand this with cybersecurity as well. Of course, there should be further studies on a wider range of situations before this can be generalized across domains. The experiment results also advocate for a combination of people involved when creating security models. Our observations show that the security experts were better at finding technical threats and alternative barriers, while the combined mass of students found a wider range of threats (i.e. ways of cheating) and consequences that would affect individuals such as themselves.

## Notes

1. This is an open source tool available and further documented at <https://www.safeexambrowser.org/>
2. Freely available at <https://github.com/KDPRO-SINTEF/BowtieTool>

## References

- Abdo, H., Kaouk, M., Flaus, J.-M. and Masse, F. (2018), "A safety/security risk analysis approach of industrial control systems: a cyber bowtie – combining new version of attack tree with bowtie analysis", *Computers and Security*, Vol. 72, pp. 175-195.
- Banerjee, A., Venkatasubramanian, K.K., Mukherjee, T. and Gupta, S.K. (2012), "Ensuring safety, security, and sustainability of mission-critical cyber-physical systems", *Proceedings of the Ieee*, Vol. 100 No. 1, pp. 283-299.
- Bau, J. and Mitchell, J.C. (2011), "Security modeling and analysis", *IEEE Security and Privacy Magazine*, Vol. 9 No. 3, pp. 18.
- BBC (2016), "Almost 6,000 online shops hit by hackers", [Online], available at: [www.bbc.com/news/technology-37643754](http://www.bbc.com/news/technology-37643754) (accessed 8 November 2018).
- Bernsmed, K., Frøystad, C., Meland, P.H., Nesheim, D.A. and Rødseth, Ø.J. (2017), "Visualizing cyber security risks with bow-tie diagrams", *International Workshop on Graphical Models for Security, Springer*, pp. 38-56.
- Bradbury, D. (2018), *Internet Shut down in Algeria to Stop Exam Cheats, Naked security*, available at: [www.theguardian.com/world/2018/jun/21/algeria-shuts-internet-prevent-cheating-school-exams](http://www.theguardian.com/world/2018/jun/21/algeria-shuts-internet-prevent-cheating-school-exams) (accessed 25 June 2018).
- Carver, J., Jaccheri, L., Morasca, S. and Shull, F. (2004), "Issues in using students in empirical studies in software engineering education", *Software Metrics Symposium, 2003 Proceedings. Ninth International, IEEE*, pp. 239-249.
- Chen, Y. and He, W. (2013), "Security risks and protection in online learning: a survey", *The International Review of Research in Open and Distributed Learning*, Vol. 14.
- Chockalingam, S., Hadžiosmanović, D., Pieters, W., Teixeira, A. and Van Gelder, P. (2016), "Integrated safety and security risk assessment methods: a survey of key characteristics and applications", *International Conference on Critical Information Infrastructures Security, Springer*, pp. 50-62.
- Elahi, G., Yu, E. and Zannone, N. (2010), "A vulnerability-centric requirements engineering framework: analyzing security attacks, countermeasures, and requirements based on vulnerabilities", *Requirements Engineering*, Vol. 15 No. 1, pp. 41-62.
- El-Attar, M., Luqman, H., Karpati, P., Sindre, G. and Opdahl, A.L. (2015), "Extending the UML statecharts notation to model security aspects", *IEEE Transactions on Software Engineering*, Vol. 41 No. 7, pp. 661-690.
- Fabian, B., Gürses, S., Heisel, M., Santen, T. and Schmidt, H. (2010), "A comparison of security requirements engineering methods", *Requirements Engineering*, Vol. 15 No. 1, pp. 7-40.
- Falessi, D., Juristo, N., Wohlin, C., Turhan, B., Münch, J., Jedlitschka, A. and Oivo, M. (2018), "Empirical software engineering experts on the use of students and professionals in experiments", *Empirical Software Engineering*, Vol. 23 No. 1, pp. 452-489.
- Firesmith, D.G. (2003), "Security use cases", *Journal of Object Technology*, Vol. 2 No. 3.
- Green, D. (2018), "If you shopped at these 7 stores in the last year, your data might have been stolen", [Online]. Business Insider Nordic, available at: <https://nordic.businessinsider.com/data-breaches-2018-4> (accessed 8 November 2018).
- Höst, M., Wohlin, C. and Thelin, T. (2005), "Experimental context classification: incentives and experience of subjects", *Proceedings of the 27th international conference on Software engineering, ACM*, pp. 470-478.



- Houmb, S.H., Islam, S., Knauss, E., Jürjens, J. and Schneider, K. (2010), "Eliciting security requirements and tracing them to design: an integration of common criteria, heuristics, and UMLsec", *Requirements Engineering*, Vol. 15 No. 1, pp. 63-93.
- ISO/IEC (2011), *ISO/IEC 27005: 2011 Information Technology – Security Techniques–Information Security Risk Management*, ISO.
- Jacobson, I. (1993), *Object-Oriented Software Engineering: A Use Case Driven Approach*, Pearson Education India.
- Johnson, C. (2011), "Using assurance cases and Boolean logic driven Markov processes to formalise cyber security concerns for safety-critical interaction with global navigation satellite systems", *Electronic Communications of the EASST*, Vol. 45.
- Jürjens, J. (2002), "UMLsec: extending UML for secure systems development", *International Conference on The Unified Modeling Language, Springer*, pp. 412-425.
- Khakzad, N., Khan, F. and Amyotte, P. (2013), "Quantitative risk analysis of offshore drilling operations: a Bayesian approach", *Safety Science*, Vol. 57, pp. 108-117.
- Kimminich, B. (2018), "OWASP juice shop tool project", [Online]. OWASP, available: [www.owasp.org/index.php/OWASP\\_Juice\\_Shop\\_Project](http://www.owasp.org/index.php/OWASP_Juice_Shop_Project) (accessed 8 November 2018).
- Kitchenham, B.A., Pfleeger, S.L., Pickard, L.M., Jones, P.W., Hoaglin, D.C., EL Emam, K. and Rosenberg, J. (2002), "Preliminary guidelines for empirical research in software engineering", *IEEE Transactions on Software Engineering*, Vol. 28 No. 8, pp. 721-734.
- Kordy, B., Mauw, S., Radomirović, S. and Schweitzer, P. (2014), "Attack–defense trees", *Journal of Logic and Computation*, Vol. 24 No. 1, pp. 55-87.
- Kriaa, S., Pietre-Cambacedes, L., Bouissou, M. and Halgand, Y. (2015), "A survey of approaches combining safety and security for industrial control systems", *Reliability Engineering and System Safety*, Vol. 139, pp. 156-178.
- Kumar, R. and Stoelinga, M. (2017), "Quantitative security and safety analysis with attack-fault trees. High assurance systems engineering (HASE)", *IEEE 18th International Symposium on, 2017. IEEE*, pp. 25-32.
- Lewis, S. and Smith, K. (2010), "Lessons learned from real world application of the bow-tie method", *6th Global Congress on Process Safety, American Institute of Chemical Engineers San Antonio, TX*, pp. 22-24.
- Lodderstedt, T., Basin, D. and Doser, J. (2002), "SecureUML: a UML-based modeling language for model-driven security", *International Conference on the Unified Modeling Language, Springer*, pp. 426-441.
- London, M. (2017), "5 Ways to cheat on online exams", [Online]. Inside Higher ED, available at: [www.insidehighered.com/digital-learning/views/2017/09/20/creative-ways-students-try-cheat-online-exams](http://www.insidehighered.com/digital-learning/views/2017/09/20/creative-ways-students-try-cheat-online-exams) (accessed 27 September 2018).
- Lu, L., Liang, W., Zhang, L., Zhang, H., Lu, Z. and Shan, J. (2015), "A comprehensive risk evaluation method for natural gas pipelines by combining a risk matrix with a bow-tie model", *Journal of Natural Gas Science and Engineering*, Vol. 25, pp. 124-133.
- Maggi, F., Quarta, D., Pogliani, M., Polino, M., Zanchettin, A.M. and Zanero, S. (2017), "Rogue robots: testing the limits of an industrial robot's security. Technical report", Trend Micro, Politecnico di Milano.
- Marsh, S. (2017), "More university students are using tech to cheat in exams", [Online]. The Guardian, available at: [www.theguardian.com/education/2017/apr/10/more-university-students-are-using-tech-to-in-exams](http://www.theguardian.com/education/2017/apr/10/more-university-students-are-using-tech-to-in-exams) (accessed 27 September 2018).
- Matulevicius, R., Mayer, N. and Heymans, P. (2008), "Alignment of misuse cases with security risk management", *Availability, reliability and security, ARES 08. Third international conference on, 2008. IEEE*, pp. 1397-1404.

- Mayer, N., Heymans, P. and Matulevicius, R. (2007), *Design of a Modelling Language for Information System Security Risk Management*, RCIS, pp. 121-132.
- Meland, P.H. (2018a), "Bowtie experiment NTNU SINTEF 2018", [Online]. NTNU, available at: <https://doi.org/10.21400/f685ryu2> (accessed 20 November 2018).
- Meland, P.H. (2018b), "Misusecaseexperiments\_SINTEF", [Online]. Zenodo, available at: <https://doi.org/10.5281/zenodo.1492322> (accessed 20 November 2018).
- Meland, P.H., Bernsmed, K., Frøystad, C. and Li, J. (2018), "An experimental evaluation of bow-tie analysis for cybersecurity requirements", in Katsikas, S.K., Cuppens, F., Cuppens, N., Lambrinouidakis, C., Kalloniatis, C., Mylopoulos, J., Antón, A. and Gritzalis, S. (Eds), *ESORICS 2018 International Workshops, CyberICPS 2018 and SECPRE 2018, September 06-07 2018*, Barcelona. Springer.
- Mellado, D., Blanco, C., Sánchez, L.E. and Fernández-Medina, E. (2010), "A systematic review of security requirements engineering", *Computer Standards Interfaces*, Vol. 32 No. 4, pp. 153-165.
- Mokhtari, K., Ren, J., Roberts, C. and Wang, J. (2011), "Application of a generic bow-tie based risk analysis framework on risk management of sea ports and offshore terminals", *Journal of Hazardous Materials*, Vol. 192 No. 2, pp. 465-475.
- Moody, D.L. (2003), "The method evaluation model: a theoretical model for validating information systems design methods", *ECIS 2003 Proceedings*, Vol. 79.
- Mouratidis, H., Giorgini, P. (2007), "Secure tropos: a security-oriented extension of the tropos methodology", *International Journal of Software Engineering and Knowledge Engineering*, Vol. 17, pp. 285-309.
- Nolan, D.P. (2014), "Safety and security review for the process industries: application of HAZOP", *PHA, What-If and SVA Reviews*, Elsevier.
- Pfleeger, S.L. (1994), "Design and analysis in software engineering: the language of case studies and formal experiments", *ACM SIGSOFT Software Engineering Notes*, Vol. 19 No. 4, pp. 16-20.
- Piètre-Cambacédès, L. and Bouissou, M. (2013), "Cross-fertilization between safety and security engineering", *Reliability Engineering and System Safety*, Vol. 110, pp. 110-126.
- Raspotnig, C., Karpati, P. and Katta, V. (2012), "A combined process for elicitation and analysis of safety and security requirements", *Enterprise, Business-Process and Information Systems Modeling*, Springer.
- Røstad, L. (2008), "An extended misuse case notation: including vulnerabilities and the insider threat", *Access Control in Healthcare Information Systems*, Vol. 67.
- Runeson, P. (2003), "Using students as experiment subjects – an analysis on graduate and freshmen student data", *Proceedings of the 7th International Conference on Empirical Assessment in Software Engineering*, Citeseer, pp. 95-102.
- Salman, I., Misirli, A.T. and Juristo, N. (2015), "Are students representatives of professionals in software engineering experiments? Software engineering (ICSE)", *IEEE/ACM 37th IEEE International Conference on, 2015, IEEE*, pp. 666-676.
- Schmittner, C., Ma, Z. and Smith, P. (2014), "Fmvea for safety and security analysis of intelligent and cooperative vehicles", *International Conference on Computer Safety, Reliability, and Security*, Springer, pp. 282-288.
- Schneier, B. (1999), "Attack trees", *Dr Dobb's Journal*, Vol. 24, pp. 21-29.
- Shostack, A. (2008), "Experiences threat modeling at Microsoft", *Modeling security workshop. Department of Computing, Lancaster University, UK*.
- Sindre, G. and Opdahl, A.L. (2001), *Capturing Security Requirments through Misuse Cases*, Norsk Informatikkonferanse 2001.
- Sjøberg, D.I., Hannay, J.E., Hansen, O., Kampenes, V.B., Karahasanovic, A., Liborg, N.-K. and Rekdal, A. C. (2005), "A survey of controlled experiments in software engineering", *IEEE Transactions on Software Engineering*, Vol. 31, pp. 733-753.
- Souag, A., Mazo, R., Salinesi, C. and Comyn-Wattiau, I. (2016), "Reusable knowledge in security requirements engineering: a systematic mapping study", *Requirements Engineering*, Vol. 21 No. 2, pp. 251-283.

- Stålhane, T. and Sindre, G. (2008), "Safety hazard identification by misuse cases: experimental comparison of text and diagrams", *International Conference on Model Driven Engineering Languages and Systems*, Springer, pp. 721-735.
- Svahnberg, M., Aurum, A. and Wohlin, C. (2008), "Using students as subjects-an empirical evaluation", *Proceedings of the Second ACM-IEEE international symposium on Empirical Software Engineering and Measurement*, ACM, pp. 288-290.
- Tichy, W.F. (1998), "Should computer scientists experiment more?", *Computer*, Vol. 31 No. 5, pp. 32-40.
- Trbojevic, V.M. and Carr, B.J. (2000), "Risk based methodology for safety improvements in ports", *Journal of Hazardous Materials*, Vol. 71 No. 1-3, pp. 467-480.
- Van Lamsweerde, A. (2004), "Elaborating security requirements by construction of intentional anti-models", *Proceedings of the 26th International Conference on Software Engineering*, IEEE Computer Society, pp. 148-157.
- Vegendla, A., Sogaard, T.M. and Sindre, G. (2016), "Extending HARM to make test cases for penetration testing", *6th International Workshop on Information Systems Security Engineering*, Ljubljana, Slovenia, Springer, pp. 254-265.
- Winther, R., Johnsen, O.-A. and Gran, B.A. (2001), "Security assessments of safety critical systems using HAZOPs", *International Conference on Computer Safety, Reliability, and Security*, Springer, pp. 14-24.
- WMN (2014), "IMB: Shipping next playground for hackers", [Online], available at: <https://worldmaritimeneews.com/archives/134727/imb-shipping-next-playground-for-hackers/> (accessed 20 November 2018).
- Zalewski, J., Drager, S., McKeever, W. and Kornecki, A.J. (2012), *Towards Experimental Assessment of Security Threats in Protecting the Critical Infrastructure*, ENASE, pp. 207-212.

#### About the authors

Per Håkon Meland is a Senior Research Scientist at the independent research institute SINTEF in Norway. In 2002, he obtained an MSc in Computer Science at the Norwegian University of Science and Technology, where he is also a PhD Fellow in the intertwined fields of threat modelling and security economics. Per Håkon Meland is the corresponding author and can be contacted at: [per.h.meland@sintef.no](mailto:per.h.meland@sintef.no)

Karin Bernsmed is a Senior Research Scientist at SINTEF. She has a PhD from the Norwegian University of Science and Technology, where she also holds an Associate Assistant Professor position. Her research interests include security threat and risk assessment, requirements engineering and design of secure and robust ICT systems.

Christian Frøystad is a Research Scientist at the independent research institute SINTEF in Norway. He obtained an MSc in Computer Science at the Norwegian University of Science and Technology in 2015.

Jingyue Li is an Associate Professor at the Department of Computer Science, Norwegian University of Science and Technology. He obtained a PhD from the Norwegian University of Science and Technology in 2006. His research interests lie in software engineering, software security and system safety.

Guttorm Sindre is a Professor at the Department of Computer Science, Norwegian University of Science and Technology, and is a Leader for the Excited Centre for Excellence in IT Education. He obtained a PhD from the Norwegian Institute of Technology in 1990. His research interests lie in requirements engineering, security requirements and IT education and didactics.

---

For instructions on how to order reprints of this article, please visit our website:

[www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)

Or contact us for further details: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)

**G: ‘Demand side expectations of cyber insurance’**

©2019 IEEE. Reprinted, with permission, from Ulrik Franke and Per Håkon Meland, 2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA), June 2019.

**G**

# Demand side expectations of cyber insurance

Ulrik Franke  
RISE Research Institutes of Sweden  
Kista, Sweden  
ulrik.franke@ri.se

Per Håkon Meland  
SINTEF Digital and  
Norwegian University of Science and Technology  
Trondheim, Norway  
per.h.meland@sintef.no

**Abstract**—Cyber insurance has attracted much attention from both practitioners, policymakers and academics in the past few years. However, it also faces some challenges before it can reach its full potential as a tool for better cyber risk management. One such challenge is the gap between what customers expect and what insurers really offer.

This paper investigates this gap empirically, based on interviews with informant companies in Norway and Sweden considering cyber insurance. The expectations expressed in the interviews are compared to anonymized incident claims reports and claims statistics for 2018 from a global insurance intermediary.

The results show no obvious pattern of discrepancies between different domains. However, informant expectations on business interruption coverage are much greater than one would expect from its share of claims. In this respect, informant expectations on business interruption coverage are more aligned with some recently published scenarios on possible major business interruptions.

**Index Terms**—cyber insurance, company expectations, cyber claims data, cyber coverage, threats

## I. INTRODUCTION

Cyber insurance has reached its age of adolescence, with sporadic growth spurts and a somewhat confusing relationship with more mature siblings such as *crime*, *property* and *liability* insurances. Part of this confusion stems from misaligned expectations. In an analysis of obstacles to more mature cyber insurance, the OECD identified misunderstandings about coverage and unsuitability of the coverage available as two of the main concerns on the demand side [1]. Despite efforts to rectify this, e.g., by Insurance Europe [2], it is safe to assume that misaligned expectations persist. Indeed, this is to be expected, as new products and new players are entering the market, and as companies with little experience of cyber insurance and perhaps cyber risk management as such look for appropriate insurance coverage. Furthermore, the ambiguity in the cyber insurance policy language should not be underestimated [3], even though there are efforts to standardize terminology [4].

The purpose of our research has been to examine the expectations that early and prospective customers have towards cyber insurance, and see if these are in line with contemporary incidents and claims. If there are discrepancies between what the customer needs, what the product offers and what kind

U. Franke was partially supported by the Swedish Civil Contingencies Agency, MSB (agreement no. 2015-6986). P.H. Meland was supported by the Norwegian Research Council (agreement no. 259869).

of incidents take place in the real world, cyber insurance will struggle reaching adulthood.

Our research approach has been to perform qualitative interviews with companies in Sweden and Norway. These two countries are at the top of the Digital Evolution Index [5], which means that they have an economy that relies strongly on the digital infrastructure, hence constituting a region where cyber insurance should be able to get a foothold. The results from these interviews have been compared with reports describing recent incident claims, claims statistics from 2018, as well as data breach statistics for different industry domains and a few cyber insurance loss scenarios. The use of these disparate data sources reflects a fundamental and well-known problem with cyber insurance: lack of data on cyber incidents [1]. Nevertheless, we believe that the sources used give interesting perspectives on our research, even if definite answers cannot be found.

We have tried to address the following research questions:

- 1) Are there different expectations in different business domains?
- 2) Are there discrepancies between coverage expectations and the costs of prevalent incidents as seen in incident data?
- 3) Are there discrepancies between coverage expectations and the costs of prevalent incidents as seen in scenarios?

The remainder of the paper is structured as follows. Section II gives a brief overview of related work. Section III details our employed method, before Section IV summarizes the results from customer interviews and incident data. A discussion and analysis of these results are presented in Section V. Section VI concludes the paper and also gives some directions for future work.

## II. RELATED WORK

Marotta et al. [6] provide a comprehensive review of the available cyber insurance literature up until 2017. They point out that there has been a “*slow start and many problematic issues*”, and enlist a number of research gaps where many are related to *lack of experience*. The main body of this literature has been focused on topics related to the provider side, and this has been the trend among the most popular papers in recent years as well, e.g. the cyber insurance assessment process [7], insurance policies analysis [8], insurance claim disputes [9], [10], [11], contract design and pricing [12], [13], [14] and

G

characterisation of markets and trends [15], [16], [17], [18], [19], [20], [21], [22] to mention a few.

Our work is mostly concerned with the demand side of cyber insurance. Here, the recent literature is rich on various investment strategies. For instance, Wang [23] proposes analytical models to quantify effects of security spending, including cyber insurance. He also suggests how innovative cyber insurance products should look like. Hoang et al. [24] propose an algorithm that owners of electric vehicles can use to determine whether to use cyber insurance as a risk transfer option. Bodin et al. [25] have created a model for selecting the optimal set of insurance policies. Here, it is pointed out that *“cybersecurity insurance premiums are commonly viewed as being poorly aligned with the risks and coverage needs of private sector firms”*. Tosh et al. [26] use game theoretic models to study self-defence investment for organisations, optimal attack rate for adversary and optimal coverage level for insurers. Similarly, Massacci et al. [27] use game theory to model firms, attackers, insurers and a public policy coordinator, where their findings show that the aggregated security level of the targets may be eroded. Mukhopadhyay et al. [28] propose a framework that allows organisations to select among cyber insurance, self-insurance or self-protection as a strategy to minimize losses. Meland and Seehusen [29] have suggested a data-driven decision support model for companies considering buying cyber insurance. Shetty et al. [30] propose a tool that allows organisations reduce insurance premiums by optimally chosen mitigation policies. Vakilinia and Sengupta [31] present three models where organisations collaboratively insure a common platform instead of themselves.

Unlike these contributions, our approach has been to empirically investigate the expectations towards cyber insurance. The most closely related paper that we are aware of is by De Smidt and Botzen [32], who have analysed professional decision makers’ perceptions of cyber risks. They found that the *“overall awareness of the cyber risks is high, the perceived probability is high, but expected impacts of a cyber-attack may be underestimated”*. Another notable finding was that *“the low uptake of cyber insurance may be explained by the low expected damage of a cyber-attack”*.

### III. METHOD

#### A. Company interviews

In the autumn of 2016, interviews on cyber insurance were performed with 10 Norwegian companies (4 of which had not considered cyber insurance previously, 3 of which had actively decided not to procure, 2 of which were considering, and 1 of which had acquired insurance). These results were first partially reported in [33], but oriented towards what kind of uncertainties existed as seen from the demand side. To complement the initial Norwegian results, a follow-up study was conducted in Sweden in the autumn of 2018. For maximum relevance, only companies who had actively considered cyber insurance were approached. 3 companies were interviewed (2 of which were considering and 1 of which had acquired insurance).

All interviews, in 2016 and 2018, were conducted in a semi-structured form for the duration of an hour at the office of the informant or using teleconferencing. The Norwegian interviews were carried out by two researchers, one of whom asked the questions and one of whom took notes. These interviews were digitally recorded, transcribed and coded, before a draft summary report was presented to the informants so they could give additional comments. The Swedish interviews were carried out by a single researcher. The raw transcription notes of the Swedish interviews were distributed to informants shortly after the interview, to give the opportunity to correct and complement the findings, and also for the informants to confirm that their level of anonymity was sufficient.

All interviews were conducted using the same semi-structured template given in Appendix A. The Swedish translation was created based on the Norwegian original in early 2017 in preparation for the second, Swedish, data-collection phase. All transcriptions were independently analysed by two researchers (the authors) in order to identify opinions about expectations among the informants.

#### B. Incident data analysis

A set of anonymized incident claims reports were obtained from the global insurance intermediary Willis Towers Watson, along with their aggregated cyber claims data [34] and data breach event statistics for 2018 [35]. The claims reports covered the following business domains and included short narratives about the event, consequences, coverage categories and claims costs:

- *Finance*, with three breach incidents related to loss of sensitive data, malware and insider data theft.
- *Healthcare*, with two incidents related to insider data theft and theft of servers containing sensitive data.
- *Education*, with three incidents related to ransomware/extortion, loss of personal data and malware information theft.
- *Manufacturing*, with two incidents related to unintentional business interruption and information theft.
- *Retail*, with two incidents related to intentional business interruption and information theft.

The claims data showed what kind of claim coverages were implicated and the data breach event statistics were based on 330 breaches and their incident costs related to a broader set of industry/business domains than the claims reports. As a supplement to this, historical data about proportion of breaches per domain had previously been obtained from Advisen, which is a commercial data provider for the insurance market.

### IV. RESULTS

We have summarized the interview results in Table I. The first column specifies the identifier of the informant. The second column defines what kind of domain they belong to, using the industries defined in [35]. In the third column we have tried to briefly summarize what kind of expectations the informants have on current products and what they would like to see more of.

TABLE I  
SUMMARY OF EXPECTATIONS FROM THE INFORMANTS

ID	Domain	Expectations
Comp1	IT	Emphasizes business interruption, and in particular related to catastrophic events.
Comp2	Finance	Expect that the insurance they have bought will cover incident response and costs related to reconstruction of lost data. Motivated to buy insurance if this gives them access to highly skilled expertise.
Comp3	Process industry	Stresses the width of insurance triggers. Coverage for ransomware attacks and similar would be nice to have, but not the only thing that would make them buy an insurance. They expect that e.g. CEO-phishing attacks would be covered by their extended crime insurance.
Comp4	Food	Expect the insurance will cover incidents with low probability but high consequence. Stresses the costs of business interruption and motivates not procuring cyber insurance with limits being too small, so that it is more rational to self-insure. Would like transfer short time business interruption risks to an insurance company.
Comp5	Transport	Stresses business interruption as the most important coverage factor.
Comp6	Media	Business interruption should be the most important coverage, not so concerned about data breach. Notes that current limits are too small. Insurance should cover the truly catastrophic events, and then some 20 MEUR is not so much.
Comp7	Finance	Stresses the non-monetary impact to reputation, questioning if insurance can help. Seemingly more interested in SLAs with guarantees than insurance to manage business interruption.
Comp8	IT	Has not considered cyber insurance, currently has self-insurance for many types of incidents. Data breach might become relevant after 2018 (GDPR implementation).
Comp9	Energy	Worries mostly about business interruption, but not relevant to insure this. Reasons that in principle, (any) insurance is only relevant if it covers what you cannot cover yourself. For instance, handling reputation loss in case of an incident is something they should take care of themselves.
Comp10	Retail	Emphasizes business interruption, for instance a virus attack that could take down the whole business.
Comp11	Retail	Places equal importance on the three components (i) data breach, (ii) business interruption, and (iii) incident response.
Comp12	Manufacturing	Deems business interruption more important than data breach.
Comp13	Manufacturing	Notes that the up until now, it has been rational for them to self-insure, as the limits have been too small and the consequences not so severe. Do not possess sensitive data and not too worried about data breach coverage today, but foresee that it may become more relevant in the future.

Fig. 1 shows actual coverages implicated based on the claims data from [34]. The coverage categories were defined by the insurance intermediary. The most interesting finding here is that incident response constitutes the bulk of the coverages (61%), while business interruption has only a mere 4%.

Table II shows informants grouped according to domains, and relates these to incidents described in claims report and breach claims data. Here, we have grouped Comp12 and Comp13 (manufacturing), Comp3 (Process industry) and Comp9 (Energy) within the same group (manufacturing). Similarly, Comp11 (Retail), Comp4 (Food) and Comp10 (Retail) are grouped as retail. We only have relevant claims reports (from [34]) for finance, manufacturing and retail.

The breach claims data, by contrast, is richer and consists of several statistics: The first is the relative frequency of breaches per domain. This data stem from Advisen and is described in [29]. The proportion of claims, average cost and total cost values stems from [35] and encompasses 281 claims worldwide from 2018. Note that *retailers* have the greatest *total* breach costs associated with them (\$3 473 550), while *hospitality and leisure* has the highest *average cost* per claim (\$173 908). *Healthcare* is the domain that has greatest *number* of claims (27% of the data set). These two latter along with some other minor categories are not displayed in Table II, as we did not interview any informants from those domains.

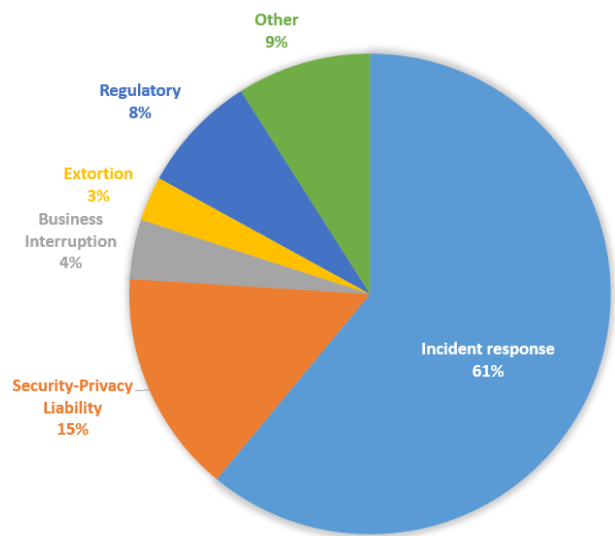


Fig. 1. Coverages implicated

## V. DISCUSSION

### A. Expectations from domains

There is no clear pattern in what companies from different domains expect from their cyber insurance policies. In the



TABLE II  
DATA RELATED TO DOMAINS

Domain	Informants	Incident examples from claims reports	Breach claims data
Finance	Comp2, Comp7	<ul style="list-style-type: none"> <li>Lost laptop containing data about individuals. Lawsuit defence costs exceeds \$700 000.</li> <li>Malware infection, expensive forensic investigations and customer credit monitoring.</li> <li>Insider stole customer information, insurance carrier helped with coordinate legal, forensics, notification, call center &amp; credit monitoring.</li> </ul>	21% of the breaches. 19.6% of the claims in 2018, average cost \$83 242, total cost \$2 422 106.
Manufacturing	Comp3, Comp9, Comp12, Comp13	<ul style="list-style-type: none"> <li>Failed software upgrade lead to business interruption, \$2M insurance pay out.</li> <li>Malicious software scraped customer credit cards, insurance carrier covered legal counselling and forensic assistance. Additional expenses related to customer services.</li> </ul>	5.6% of the breaches. 1.8% of the claims in 2018, average cost \$152 900, total cost \$764 500.
Retail	Comp4, Comp10, Comp11	<ul style="list-style-type: none"> <li>Stolen login credentials, 50 000 customer credit card numbers stolen. \$1M income loss.</li> <li>DDoS attack lead to service disruption and a \$300 000 income loss.</li> </ul>	9.9% of the breaches. 8.5% of the claims in 2018, average cost \$144 731, total cost \$3 473 550.
Transport	Comp5	-	4.6% of the breaches. 1.1% of the claims in 2018, average cost \$10 331, total cost \$30 994.
IT	Comp1, Comp8	-	Part of the wider Advisen category <i>services</i> , which has 42.6% of the breaches. 4.6% of the claims in 2018, average cost \$6 968, total cost \$90 586.
Media	Comp6	-	Part of the wider Advisen category <i>services</i> , which has 42.6% of the breaches. 1.4% of the claims in 2018, average cost \$14 879, total cost \$59 516.

*retail* domain, there is typically an emphasis on business interruption (Comp4, Comp10) or it is at least deemed as important as other aspects (Comp11). This is somewhat contrary to our expectations of a greater emphasis on data breach, based on retail handling large amounts of personal data and credit card data from customers. One of our incident examples in Table II showed that this had been the case, and that the income loss was much higher than the average cost covered by insurance. As already mentioned, the claims statistics show that retailers have the largest amount of total breach costs.

In the *manufacturing* domain, business interruption is deemed more important than data breach (Comp9 and Comp12). Comp12 motivates this by noting that being in the business-to-business rather than business-to-consumer segment gives less exposure to sensitive personal data and the potential consequences of a breach. This is more in line with expectations. One of the incident examples showed that the business interruption pay-out was much higher (13x) than the average cost covered by insurance, and also larger than the total cost. We assume that this incident must be prior to 2018, but it is a good illustration how one incident can dominate a market when the number of claims is small with variable

costs. Another noteworthy finding is that this domain has a much lower proportion of claims compared to the proportion of breaches.

Considering the *finance* domain, Comp2 were concerned about incident response and recovery, while Comp7 stressed that good SLAs were more important than insurance. Expectations on response and recovery are aligned with the two latter incident examples for this domain in Table II. On the other hand, SLAs would not have made a difference for any of the three incident examples. It is also interesting to note that finance is where the largest proportion of breaches and claims occur among the companies we interviewed. This is in accordance with Forbes, claiming that US financial services firms are attacked more than 300 times more frequently than businesses in other industries [36].

For the remaining domains, *transport*, *IT* and *media*, coverage of business interruption is the common expectation. These domains constitute a significant portion of the breaches; however, the claims data show that there are only a few claims, and these are all very low in terms of costs. This is an indication of a discrepancy between coverage expectations, claims and incidents.

Independent of domain, several informants (Comp4, Comp9, Comp13) reason about self-insurance, arguing that this is a more rational option when limits are too small or important consequences such as reputational damage are not covered anyway. Comp13 (manufacturing), however, goes on to say that this may change as their line of business is expected to undergo digital disruption in the coming years, where their exposure both to personal data (covered by cyber insurance) and to cyber risks that might entail physical damage or bodily injuries (not covered by cyber insurance) will increase. Comp13 also stated that even though cyber is a great risk, it is currently not their greatest. If we compare this with the Allianz Risk Barometer from 2019 [37], business interruption and cyber incidents are on the top for Europe considering all domains. However, for manufacturing, *natural catastrophes* trump cyber incidents, so this is in accordance with the view of Comp13.

### B. Expectations and claims paid out

Comparing customer expectations with statistics on the claims paid out, some discrepancies can be identified.

First, it is evident that incident response expenses are by far the most common of the coverages implicated in the claims data [34], as seen in Fig. 1. The coverage for incident response expenses is implicated some 4 times more often than security/privacy liability (roughly the same as data breach) and some 15 times more often than business interruption. This is clearly out of proportion compared to customer expectations. As we saw in the previous section, there are informants who value incident response highly (Comp2) or at least places it on a par with data breach and business interruption (Comp11), but they are a minority.

However, it should also be noted that the incident response coverage is a bit different from the other categories in the sense that the latter count different kinds of incidents (e.g., data breach is one kind of incident and business interruption is another), but incident response counts them all. Hypothetically, for any number of incidents that each implicate (i) a particular category and (ii) the general incident response category, incident response would account for half of the coverages implicated. Furthermore, if some incidents do not reach the thresholds for activation in the particular categories, e.g., the waiting periods always included in business interruption coverage, incident response would account for strictly more than half of the coverages implicated. From this perspective, the apparent over-representation of incident response fully disappears.

Conversely, business interruption coverage was highly valued by informants from retail (Comp4, Comp10) and manufacturing (Comp9 and Comp12) alike, yet accounts for only 4% of the coverages implicated. Here, it is clear that customer expectations are not in line with claims actually paid out. However, business interruption policies have waiting periods before they are activated. These are rarely shorter than some 6 or 8 hours and often significantly longer, e.g., 24, 36, 48, or 72 hours [15]. From this perspective, it is not surprising that

business interruption claims are limited. It is also worth observing that many informants emphasize that insurance should cover incidents with low probability but high consequence (Comp1, Comp4, Comp6). Thus, their expectations may not be that business interruption should represent a large proportion of claims in any given year, but rather that *if* there are interruptions with very long durations, these will be covered. This leads us to the question of how expectations align with business interruption scenarios.

### C. Expectations and scenarios

Since it is well-known that statistics on cyber insurance are rare [1], [15], [6], it is reasonable to expect that expectations – both of customers and insurers – are also formed by hypothetical scenarios. Some such scenarios are private, e.g., internal risk analyses carried out before procuring an insurance, or as part of the underwriting. However, other scenarios are made public. It is instructive to compare some recent such published scenarios with the customer expectations expressed in the interviews.

In early 2018, Lloyd's released a report mapping the impact on US companies of a major cloud service provider outage, i.e. an outage in the order of several days [38]. While the big enterprise public cloud providers – Amazon Web Services, Microsoft Azure, Google Cloud and IBM – are all remarkably reliable, outages do happen and the impact of such downtime is substantial. In the scenarios investigated in the report, where a major cloud service provider is down for 3-6 days, the US manufacturing industry would experience ground up losses of some \$4.2-\$8.6 billion, and the US wholesale and retail trade industry would experience ground up losses of some \$1.4-\$3.6 billion. It is noteworthy that these are the industries carrying the greatest losses.

In early 2019, Lloyd's together with the Cambridge Centre for Risk Studies, and Nanyang Technological University, released a report exploring the consequences of a global infection by contagious malware [39] – not unlike the real cases of WannaCry and NotPetya in 2017. In a less severe version of the scenario, retail suffers the most (\$15 billion), followed by healthcare (\$10 billion) and manufacturing (\$9 billion). In a more severe version, retail and healthcare are on a par (\$25 billion each), followed by manufacturing (\$24 billion). This report also explores the insurance coverages implicated, showing that business interruption is the main driver of the insured losses (with some 71% of total losses in the less severe scenario and 59% in the more severe one). The second and third largest claims arise in incident response costs and liability, respectively.

In brief, it can be concluded that scenarios such as these are well aligned with the concerns of the many informants, who emphasize business interruption coverage (Comp1, Comp4, Comp5, Comp6, Comp9, Comp10, Comp11 and Comp12).

### D. Who is to blame?

An interesting remark by Comp11, which at the time of the interview had just signed their cyber insurance policy, is

that its existence will not be announced to the employees. The reason given is moral hazard: the risk of more reckless behaviour as a consequence of insurance protection (this also applies to other insurance policies at Comp11). Comp12, which at the time of the interview expected to soon request quotes through an insurance intermediary, reasoned in a similar way.

Such management of human error is prudent, since 66% of the incidents in the claims statistics from Willis Towers Watson were blamed on “employee negligence or malfeasance”. For example, phishing and ransomware that is introduced by an employee opening a malicious e-mail attachment is typically considered “human error” by insurers.

#### E. Validity and reliability

The 13 companies interviewed were selected with a kind of *purposive sampling*, actively looking for companies considering cyber insurance. Thus, the sample cannot be claimed to be representative of all Norwegian or Swedish companies. However, such broad representativity was never the goal of the research. It is known from previous work that the number of companies with cyber insurance is still low in Norway and Sweden [15], so a random sample of all companies would mostly generate informants oblivious to cyber insurance. Instead, the aim was to investigate attitudes of companies who had considered cyber insurance. From this perspective, the sample is more representative. Among the 13 companies, there is broad representation from different industries, such as finance, media, retail, manufacturing, critical infrastructure and IT.

Most of the informant companies are relatively large and many of them are international companies active on many markets. This means that they represent the large company market segment of cyber insurance. While most insurance companies offering cyber insurance have relatively large customers [15], there is also an SME cyber insurance market segment. Both in Denmark and in Sweden, thousands of small cyber insurance packages have been sold, typically with comparably small indemnity limits and mostly focusing on incident response [15, Table 1, Insurance company 1]. Thus, it is important to bear in mind that the results do not represent this SME segment, but the large company segment of cyber insurance.

The claims data set represents events from all over the world (though mostly North America and Europe). Thus, while the claims data set has a much broader scope than the interviews, it allows to assess the expectations of (would-be) cyber insurance policy holders with the kinds of events for which insurance claims are actually paid out. While it might be argued that it would be more accurate to compare the expectations of Scandinavian policy holders with Scandinavian claims, this is not feasible, as it is known that the number of such claims is still very low [15]. Instead, to have a reasonable frame of comparison, it is necessary to look at claims from a larger area, and the data set used can thus be deemed suitable, though results need to be interpreted with some caution.

## VI. CONCLUSIONS AND FUTURE WORK

Revisiting the research questions posed in Section I, first, we can discern no obvious pattern of discrepancies between different domains. What differences there are between of informants do not correspond to their domains, and the coverage that seems to be the most valued among all informants is business interruption.

This naturally leads us to our second research question, because this is *not* aligned with the incident data. Here, incident response constitutes 61% of coverages implicated, whereas business interruption represents a mere 4%. However, this dominance of incident response is not surprising when accounting for the fact that this is a generic category that applies to all incidents. Similarly, the small fraction of business interruption can to some extent be explained by the fact that this coverage has waiting periods before it is activated. Many informants also reason in a mature way about this: the important thing is not coverage of many small incidents that happen every year, but rather coverage of rare but substantial incidents. From this perspective, the waiting periods are not misaligned with customer expectations.

This leads to the third question, about scenarios. Indeed, some recently published scenarios on possible major business interruptions, due to cloud service outages or rapidly spreading ransomware, are more aligned with informants’ emphasis on business interruption coverage. At the very least, these scenarios show that major business interruption events are not at all implausible, and are in this respect aligned with customer expectations on business interruption coverage as expressed by the informants.

A few avenues for further research suggest themselves. Cyber insurance is still in its development phase and the number of claims paid out in the Nordic region is still very low [15]. Thus, it would be interesting to investigate how expectations on coverage change over time.

Second, the interview guide given in Appendix A can be used to conduct comparative studies in other regions. Are customer expectations uniform all over the world, or do they differ? Based on previous research [15], we hypothesize that customers in Europe still focus more on 1st party costs such as business interruption, whereas customers in the US still focus more on 3rd party liabilities connected to data breaches. However, the advent of the General Data Protection Regulation (GDPR) in Europe might change this, so longitudinal studies, tracking the expectations over time, are interesting.

A third research direction is to look at how emerging threats may change the expectations of what policies should cover. For instance, there is a generally high awareness about ransomware, which has very noticeable consequences and is present in the majority of the policies [6], [33]. However, a more recent trend is that the ransomware threat is being dethroned by cryptojacking or cryptomining malware, which also provides direct revenue to the criminals but with a lower risk of penalty. According to the latest Internet Organised Crime Threat Assessment from Europol [40], this is a type of

threat that only has a small impact on the victim's system, and it is hard to quantify the damages and difficult to investigate due to the lack of reporting. It is comparable to theft of electricity, which may go unnoticed over a long period of time though the accumulated costs can become significant. How the insurance market will position themselves towards large-scale, low-impact threats is still an open question.

#### ACKNOWLEDGMENT

The authors would like to thank all the interviewed informants. We would also like to express our gratitude towards Willis Towers Watson and Advisen for sharing their data with us.

#### APPENDIX A: INTERVIEW GUIDE

In the following, the interview guide used is outlined, translated from Norwegian/Swedish and somewhat abbreviated.

##### A. About the research

- 1) Short introduction of the scope of the research project and the interview.
- 2) Do you give your informed consent to the use of the material gathered for scientific publication?

##### B. Background on the informant and the enterprise

- 1) What is your role, how long have you had it, and what is your background?
- 2) Is there a CISO in the enterprise?
- 3) Has the enterprise procured cyber insurance?

##### C. Evaluation of cyber insurance

*For enterprises that **have procured** cyber insurance.*

- 1) What made you consider cyber insurance?
- 2) Can you describe the process (roles, intermediary, understanding market offerings)?
- 3) Can you describe the process of obtaining an insurance quote (relevance of insurer questions, proposal forms, etc.)?
- 4) What should insurance quotes and insurance policies look like to be attractive (e.g., price, incident response service, claims payment, simplicity, flexibility, coverage of many small incidents, coverage of catastrophic incidents)?
- 5) Can you describe how the decision was reached on which insurance policy to choose (easy or difficult decision, comfortable with it, the right competence to decide, based on risk-analysis or quantification of security, insurance vs. other measures)?
- 6) Does the existence of a cyber insurance affect how you work with security?
- 7) Have you experienced incidents covered by the insurance?

*For enterprises that have **not procured, but considered** cyber insurance.*

- 1) What made you consider cyber insurance?

- 2) Can you describe the process (roles, intermediary, understanding market offerings)?
- 3) Was there a process of obtaining an insurance quote (relevance of insurer questions, proposal forms, etc.)?
- 4) What should insurance quotes and insurance policies look like to be attractive (e.g., price, incident response service, claims payment, simplicity, flexibility, coverage of many small incidents, coverage of catastrophic incidents)?
- 5) Can you describe how the decision was reached on not taking out insurance (easy or difficult decision, comfortable with it, the right competence to decide, based on risk-analysis or quantification of security, insurance vs. other measures)?
- 6) If you had procured a cyber insurance, do you think it would have affected how you work with security?

*For enterprises that have **not procured and not considered** cyber insurance.*

- 1) Why have you not considered cyber insurance?
- 2) To what extent are you familiar with cyber insurance products and their meaning?
- 3) Do you have other kinds of insurance that also cover cyber crime related incidents?
- 4) What would be important to make cyber insurance relevant for you?
- 5) How do you make decisions on the kinds of cyber security measures you need to implement (risk-analysis, evaluation of measures against each other, quantification of security, roles involved, easy or difficult decisions, comfortable with them)?

##### D. Evaluation of enterprise cyber risk

- 1) How exposed are your enterprise to cyber risk (why, what are the potential consequences)?
- 2) Have you experienced cyber incidents (and did they affect your future risk management)?

##### E. Conclusion

- 1) Thanks and a brief description of the road from interview to scientific publication.
- 2) Is there anything you would like to add about cyber insurance?
- 3) Are there any questions related to the interview that future research should address?
- 4) Do you want information about future research project activities?

#### REFERENCES

- [1] OECD, "Enhancing the Role of Insurance in Cyber Risk Management," 2017.
- [2] Insurance Europe, FERMA, and BIPAR, "Preparing for cyber insurance," 2018, accessed February 27, 2019. [Online]. Available: <https://www.insuranceeurope.eu/preparing-cyber-insurance>
- [3] "Cyber-insurance: Black swans and fat tails," *The Economist*, pp. 61–62, 2019.
- [4] ENISA, "Commonality of risk assessment language in cyber insurance," European Union Agency for Network and Information Security, Tech. Rep., 2017.



- [5] R. Wallace, "Digital evolution index maps competitiveness of 60 countries," 2017. [Online]. Available: <http://www.thenextsiliconvalley.com/2017/07/21/4784-digital-evolution-index-maps-competitiveness-of-60-countries/>
- [6] A. Marotta, F. Martinelli, S. Nanni, A. Orlando, and A. Yautsiukhin, "Cyber-insurance survey," *Computer Science Review*, vol. 24, pp. 35–61, 2017.
- [7] D. Woods, I. Agrafiotis, J. R. Nurse, and S. Creese, "Mapping the coverage of security controls in cyber insurance proposal forms," *Journal of Internet Services and Applications*, vol. 8, no. 1, p. 8, 2017.
- [8] S. Romanosky, L. Ablon, A. Kuehn, and T. Jones, "Content analysis of cyber insurance policies: How do carriers write policies and price cyber risk?" 2017. [Online]. Available: <http://dx.doi.org/10.2139/ssrn.2929137>
- [9] J. P. Kesan and C. M. Hayes, "Strengthening cybersecurity with cyberinsurance markets and better risk assessment," *Minn. L. Rev.*, vol. 102, p. 191, 2017.
- [10] B. R. Ostrager and T. R. Newman, *Handbook on Insurance Coverage Disputes*. Aspen Publishers, 2018.
- [11] E. S. Knutsen and J. W. Stempel, "The techno-neutrality solution to navigating insurance coverage for cyber losses," *Penn St. L. Rev.*, vol. 122, p. 645, 2017.
- [12] M. M. Khalili, P. Naghizadeh, and M. Liu, "Designing cyber insurance policies: The role of pre-screening and security interdependence," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 9, pp. 2226–2239, 2018.
- [13] R. Pal, L. Golubchik, K. Psounis, and P. Hui, "Security pricing as enabler of cyber-insurance a first look at differentiated pricing markets," *IEEE Transactions on Dependable and Secure Computing*, pp. 1–1, 2018.
- [14] M. Xu and L. Hua, "Cybersecurity insurance: Modeling and pricing," 2017.
- [15] U. Franke, "The cyber insurance market in Sweden," *Computers & Security*, vol. 68, pp. 130–144, 2017.
- [16] G. Strupczewski, "The cyber-insurance market in poland and determinants of its development from the insurance brokers perspective," *Economics and Business Review*, vol. 3, no. 2, pp. 33–50, 2017.
- [17] M. Camillo, "Cyber risk and the changing role of insurance," *Journal of Cyber Policy*, vol. 2, no. 1, pp. 53–63, 2017.
- [18] C. Biener, M. Eling, and J. H. Wirfs, "Insurability of cyber risk," *Methodology*, p. 9, 2018.
- [19] D. Woods and A. Simpson, "Policy measures and cyber insurance: a framework," *Journal of Cyber Policy*, vol. 2, no. 2, pp. 209–226, 2017.
- [20] P. Low, "Insuring against cyber-attacks," *Computer Fraud & Security*, vol. 2017, no. 4, pp. 18 – 20, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1361372317300349>
- [21] E. Kopp, L. Kaffenberger, and N. Jenkinson, *Cyber Risk, Market Failures, and Financial Stability*. International Monetary Fund, 2017.
- [22] G. Peters, P. V. Shevchenko, and R. Cohen, "Understanding cyber-risk and cyber-insurance," 2018. [Online]. Available: <http://dx.doi.org/10.2139/ssrn.3200166>
- [23] S. Wang, "Integrated framework for information security investment and cyber insurance," 2017. [Online]. Available: <http://dx.doi.org/10.2139/ssrn.2918674>
- [24] D. T. Hoang, P. Wang, D. Niyato, and E. Hossain, "Charging and discharging of plug-in electric vehicles (pevs) in vehicle-to-grid (v2g) systems: A cyber insurance-based model," *IEEE Access*, vol. 5, pp. 732–754, 2017.
- [25] L. D. Bodin, L. A. Gordon, M. P. Loeb, and A. Wang, "Cybersecurity insurance and risk-sharing," *Journal of Accounting and Public Policy*, vol. 37, no. 6, pp. 527–544, 2018.
- [26] D. K. Tosh, I. Vakilinia, S. Shetty, S. Sengupta, C. A. Kamhoua, L. Njilla, and K. Kwiat, "Three layer game theoretic decision framework for cyber-investment and cyber-insurance," in *International Conference on Decision and Game Theory for Security*. Springer, 2017, pp. 519–532.
- [27] F. Massacci, J. Swierzbinski, and J. Williams, "Cyberinsurance and public policy: Self-protection and insurance with endogenous adversaries," *Paragraph*, vol. 1, no. 2, p. 2, 2017.
- [28] A. Mukhopadhyay, S. Chatterjee, K. K. Bagchi, P. J. Kirs, and G. K. Shukla, "Cyber risk assessment and mitigation (cram) framework using logit and probit models for cyber insurance," *Information Systems Frontiers*, Nov 2017. [Online]. Available: <https://doi.org/10.1007/s10796-017-9808-5>
- [29] P. H. Meland and F. Seehusen, "When to treat security risks with cyber insurance," in *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*. IEEE, 2018, pp. 1–8.
- [30] S. Shetty, M. McShane, L. Zhang, J. P. Kesan, C. A. Kamhoua, K. Kwiat, and L. Njilla, "Reducing informational disadvantages to improve cyber risk management," *The Geneva Papers on Risk and Insurance-Issues and Practice*, vol. 43, no. 2, pp. 224–238, 2018.
- [31] I. Vakilinia and S. Sengupta, "A coalitional cyber-insurance framework for a common platform," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 6, pp. 1526–1538, June 2019.
- [32] G. de Smidt and W. Botzen, "Perceptions of corporate cyber risks and insurance decision-making," *The Geneva Papers on Risk and Insurance-Issues and Practice*, vol. 43, no. 2, pp. 239–274, 2018.
- [33] P. H. Meland, I. A. Tøndel, M. Moe, and F. Seehusen, "Facing uncertainty in cyber insurance policies," in *International Workshop on Security and Trust Management*. Springer, 2017, pp. 89–100.
- [34] "Cyber risk exposure – what are the business risks?" Willis Towers Watson, Tech. Rep., 2018.
- [35] "Intelligence & risk insight report: Data breach event statistics," Willis Towers Watson, Tech. Rep., 2018.
- [36] B. Mirchandani, "Laughing All The Way To The Bank: Cybercriminals Targeting U.S. Financial Institutions," 2018, accessed March 3, 2019. [Online]. Available: <https://www.forbes.com/sites/bhaktimirchandani/2018/08/28/laughing-all-the-way-to-the-bank-cybercriminals-targeting-us-financial-institutions/>
- [37] "Allianz risk barometer results appendix 2019," Allianz, Tech. Rep., 2019. [Online]. Available: [https://www.agcs.allianz.com/assets/PDFs/Reports/Allianz\\_Risk\\_Barometer\\_2019\\_APPENDIX.pdf](https://www.agcs.allianz.com/assets/PDFs/Reports/Allianz_Risk_Barometer_2019_APPENDIX.pdf)
- [38] "Cloud Down: Impacts on the US economy," Lloyd's of London, Tech. Rep., 2018, accessed March 19, 2018. [Online]. Available: <https://www.lloyds.com/news-and-risk-insight/risk-reports/library/technology/cloud-down>
- [39] J. Daffron, S. Ruffe, C. Andrew, J. Copic, K. Quantrell, S. A., and E. Leverett, "Bashe attack: Global infection by contagious malware," Cambridge Centre for Risk Studies, Lloyds of London and Nanyang Technological University, Tech. Rep., 2019, accessed February 4, 2019. [Online]. Available: <https://www.lloyds.com/news-and-risk-insight/risk-reports/library/technology/bashe-attack>
- [40] "Internet organised crime threat assessment 2018," Europol, Tech. Rep., 2018. [Online]. Available: <https://www.europol.europa.eu/internet-organised-crime-threat-assessment-2018>

## **H: ‘An Experimental Analysis of Cryptojacking Attacks’**

©2019 IEEE. Reprinted, with permission, from Per Håkon Meland, Bent Heier Johansen and Guttorm Sindre, 2019 Nordic Conference on Secure IT Systems (NordSec), November 2019.

H

# An Experimental Analysis of Cryptojacking Attacks

Per Håkon Meland<sup>1,2</sup>[0000-0002-5509-0184], Bent Heier  
Johansen<sup>2</sup>[0000-0001-7454-6557], and Guttorm Sindre<sup>2</sup>[0000-0001-5739-8265]

<sup>1</sup> SINTEF Digital, Trondheim, Norway

`per.h.meland@sintef.no`

`https://www.sintef.com`

<sup>2</sup> Norwegian University of Science and Technology, Trondheim, Norway

`{per.hakon.meland,guttorm.sindre}@ntnu.no`

`bent.heier@gmail.com`

`https://www.ntnu.edu`

**Abstract.** Cryptojacking is the illicit exploitation of Internet users' bandwidth and processing power to mine cryptocurrencies. This paper presents an experimental analysis of how different types of cryptojacking attacks impact a selection of consumer-grade devices, and the perceived annoyance by the user. This is seen in relation to the expected cost and revenue the attacker would expect. The results show that a well-configured cryptojacking attack does not significantly harm its victims, hence can be very difficult to detect, and even aware users might not bother getting rid of the infection. The costs and risk associated with performing cryptojacking are low, but the attacker would rely on a pool of infected devices over a prolonged period of time in order to make any significant revenue. The main cost is therefore the opportunity cost, as there are more profitable ways to abuse compromised systems due to the general decline in cryptocurrency values. Though the heyday of cryptojacking has gone by, several adversaries are likely to have made quite a profit from it. It can therefore emerge as a serious threat again due to market externalities.

**Keywords:** Cryptojacking · Cryptomining · Drive-by mining · Monero · Blockchain · Malware · Experiment · Economy.

## 1 Introduction

Cryptojacking is one of the youngest members in the family of cryptocurrency related crimes, including blatant theft, illegal trading, money laundering, extortion and ransomware among others. Since the investment costs of hardware and electricity in most cases exceed the expected profit from mining cryptocurrencies, the goal with cryptojacking is to illicitly exploit Internet users' bandwidth and processing power to mine on behalf of the attacker. In contrast to many other types of attacks, cryptojacking is not about stealing or altering data, nor does it want to interrupt the victims workflow or operations. Instead it wants



to stay hidden and extract as many CPU cycles as possible. In 2018, Europol [17] proclaimed that the industry reported an explosion in the volume of illicit cryptomining and in the latter part of 2017, it overshadowed almost all other malware threats. However, in a more recent report from 2019 [40], Symantec finds that cryptojacking dropped by 52% between January and December 2018, and that the declining trend is continuing.

In order to explain some of the reasons why cryptojacking promptly declined as a cyber threat, this paper presents an experimental analysis of cryptojacking impacts using a selection of six types of consumer-grade devices. This is seen in relation to the expected cost and profit the attacker would expect. The goal of the experiment has been to answer the following research questions:

1. How is performance on different types of devices affected by cryptojacking measured objectively and perceived subjectively?
2. What are the expected revenues and costs for the attacker based on the targeted devices?

The outline of this paper is as follows: Section 2 explains the basics of cryptojacking attacks and cryptocurrencies typically associated with them. In Section 3 we present the experiment setup and measurement types. Section 4 presents the results from the experiment in terms of mining efficiency and how the devices are degraded seen objectively and subjectively. Section 5 discusses mining times and investments compared to expected profit, as well as limitations and related work. Finally, Section 6 revisits the research questions and concludes the paper.

## 2 Background

### 2.1 Types of cryptojacking attacks

There are two main types of cryptojacking attacks; one which require a malicious payload to be installed on the user's computer and the other which runs inside the user's browser upon visiting dubious web sites. In the former case, the simplest attacks typically fool users to download and launch an executable file or open an email attachment. More advanced methods exploit unpatched vulnerabilities, often zero-days to bypass the user entirely and install the payload.

The second type is an even more subtle way of attacking. About 95% of all web sites use *JavaScript* [43], and due to its popularity JavaScript is supported by all major web browsers. JavaScript is a quite powerful programming language running inside the web browser and uses the computing power of the client, not the web server. This allows for a lot of processing power, including the power to mine cryptocurrency. Such a *drive-by download* attack [11] terminates as soon as the web page is closed, leaving no trace on the victim's computer.

The most well-known script for cryptojacking was offered by *Coinhive* [7]. It allowed web site owners to deliberately put a cryptominer on their web site, letting visitors choose to allow the use of their CPUs for mining. However the same script was also frequently injected into compromised sites [12,32]. The business

model of Coinhive was to take 500 EUR for an account creation, then a 30% share of the mining itself. The services offered by Coinhive were not nefarious or illegal, in fact, they presented themselves as an alternative to advertisement, which is one of the main sources of revenue on the Internet to day. However, Coinhive was quite controversial and received their share of criticism since their initial script did not ask web site visitors for consent, and the users did not have to upgrade to the one that required this. On February 26th 2019, the Coinhive Team announced they were shutting down their service as of March 8th 2019. They proclaimed that it was no longer profitable to keep the service operating anymore, citing that Monero had depreciated more than 85% over the last year and that the hashrate dropped over 50% after the last hard fork[8].

Coinhive accounted for 70-75% of the cryptojacking JavaScripts on the web in 2018 [31,35], and while the default setting was to use 100% of the victims CPU, researchers have found that most sites throttled themselves to use between 25% and 70% of available CPU power [16,27]. This was likely done to make the mining unnoticeable to the user. A report by R uth et al. [35] also found that merely 10 user accounts were responsible for 80% all short links, meaning that only a handful of people were reaping the vast majority of the profits.

An alternative approach for cryptojacking is to mine using *plugins* that are used by web sites, such as *Wordpress* plugins. This will require a compromise of the browser extension itself, which is easier to detect. In the past, Wordpress had cryptomining plugins on its official plugin page, including several miners using the Coinhive script [44]. These could be included by legitimate web site owners, but they could also be deployed on compromised sites. *Browser extensions* are yet another vector for attackers.

Cryptojacking can also target smart phones and IoT devices. For the Android operating system, it is possible to download applications as *APK*-files from the Internet and install them directly without going through Google's Play Store. If the *side loading* setting is not set to off, cryptomining apps like *HiddenMiner* [45] can take advantage of the device. The auto update feature can also be exploited to install a cryptominer. It should be noted that Google have recently removed all cryptomining apps from the Google Play Store. Apple's iOS has been less susceptible to these kinds of attack due to a stricter lock-down policy. However, there have been incidents where apps suddenly begun mining cryptocurrency, such as the *Calendar 2* app [24]. Apple has also proclaimed that they do not allow cryptominers in their App Store [4], but mining can still be done using developer accounts or a jailbroken device.

Luckily the security industry has developed many techniques to prevent cryptojacking [38]. For native miners, all the same procedures that prevent other kinds of malware will be effective. For instance, anti-virus programs have caught up and can detect the well-known cryptominers [14,20,18]. To protect against web miners there exist a lot of options as well, such as *browser extensions*, specialized *addons* and general purpose *ad-blockers*.

## 2.2 Coins suited for cryptojacking

The first mainstream and most popular cryptocurrency, *Bitcoin*, was created to establish a decentralized global currency [28]. While Bitcoin in theory is anonymous, linking an account to a person is considered manageable when the coins at some point are exchanged or used to buy items. To preserve the integrity of the blockchain, all Bitcoin transactions and associated wallets are public. This means that if a person is linked to a wallet, all previous transactions can be traced back as well.

*Monero* [26] is a cryptocurrency based on Bytecoin. Bytecoin was abandoned when the community found out that its creators had mined about 80% of the supply for themselves, but the technology was sound, so Monero rose from the ashes. Monero uses an algorithm called *CryptoNote*, which is virtually untraceable and unlinkable [37]. This is a desirable feature when you are exploiting somebody else's hardware. Monero is currently only on the 10th place among cryptocurrencies when it comes to market capitalization [9], however it is a very popular payment option among Dark Net marketplaces trading illegal goods and services.

In cryptomining everyone that mines is competing to solve the next block and get the next payout. Bitcoin and similar technologies use primarily raw computing power and can be effectively done in parallel. This makes expensive *High Performance Computers* (HPC) desirable targets for native Bitcoin mining. However, these machines tend to be well protected and not easy to infect with native cryptojacking attacks.

CryptoNote is less CPU intensive, but requires a relatively large amount of memory (CPU-cache or RAM) instead. Compared to Bitcoin, the benefits of using large computing clusters, GPUs and ASICs over regular CPUs are severely diminished. This means that average consumer-grade hardware has a decent chance of solving the puzzle and get the payout. This in turn makes Monero an attractive currency to mine when someone has access to a large number of regular and cheap devices, such as laptops, IoT-devices and smart phones. These devices exist in enormous quantities around the world with limited protection, hence very suitable targets for cryptojackers seeking Monero.

## 3 Method

For our cryptojacking experiment we decided to focus on Monero mining and a selection of consumer-grade devices typically found in homes and work places. The goals were to understand how Monero performs under different configurations, how efficient web mining is in comparison to native mining, how much power is consumed and how noticeable this kind of mining would be on an infected device. The devices we included were the following:

- **NUC** (Intel NUC7i5BNK) was a tiny computer running Linux Ubuntu 18.04. It had a two core, four thread, i5 2.3 GHz CPU with a turbo mode at 3.4 GHz, 4 MB CPU cache memory and 8 GB RAM. It was released in Q1 2017 and represents low-to-medium powered computers.

- **Mac** was a mid-2014 laptop running MacOS High Sierra. It had a dual core, four thread CPU running at about 2.6 GHz with a turbo mode at 3.1 GHz. With 3 MB CPU cache memory, 8 GB RAM and no discrete graphics, it was chosen to represent laptops.
- **Chromebook** was a low powered ASUS laptop running Chrome OS with developer access. It had a 2.16 GHz dual core CPU without hyper threading, 1 MB of L2 cache and 2 GB RAM. It represents devices that do not have true access to the hardware.
- **Stationary** was a custom made desktop PC with Microsoft Windows 10, build in early 2014 with a four core, eight thread, i7 CPU running at 3.40 GHz, with turbo up to 3.90 GHz, 8 MB of L3 cache, 16 GB RAM and a discrete Nvidia GTX 760 graphics card, making it the most powerful device in the experiment.
- **Phone** was a Sony H4113 Android smart phone from 2018 with root access. It had two ARM CPUs, both dual core, four thread, one running at 2.2 GHz and one running at 1.8 GHz with 3 GB RAM. It was included to analyze web mining on phones.
- **Rpi** was a Raspberry Pi 2 Model B with a 900 MHz ARM Cortex-A7 CPU with 256 KB of L2 cache and 1 GB RAM, This was the least powerful device used in this experiment and dates back to 2015. It represents fairly advanced IoT devices and was only used for native mining.

In order to determine Monero performance, the hashrate was the main parameter. The peak and average hashrates were recorded by the mining software. The peak tells us what the device is capable of when the miner has most of the device's resources for itself, while the average hashrate tells us how much is likely to be mined when the device is used in a regular manner. The tests were performed with a varying amount of threads mining simultaneously, which allowed us to see the overhead effects as well.

For the native tests a miner called *XMR-stak* [46] was used. It runs natively on x86 versions of Linux, Windows and MacOS for both the CPU and GPU. Unfortunately, XMR-stak does not run on ARM devices, and as thus it could not be used on the Rpi. Instead another program, *cpuminer-multi* [13] was used in this case. We also employed a mining pool named *supportxmr.com* [39] that allowed us to extrapolate the number of required hashes for one coin of Monero without actually having to mine a whole coin. A mining pool works by connecting the resources of many miners together. When a block is solved, every member of the pool gets a share of the coinage based on the amount of work they contributed. In this way a mining pool can provide a steady and predictable income as opposed to the random nature of solo mining. The effectiveness of web mining was measured by employing several different web sites, including *coinhive.com* [7] (before its shutdown), *coinwebmining.com* [10] and *minero.cc* [25].

Each device ran for at least 1 hour for each configuration of native mining and for at least 10 minutes of web mining. Though the time intervals are somewhat short, initial testing showed that the hashrate was quite stable, so it was deemed

unnecessary to prolong the experiment. The web mining gave a real time update and had far less variance than the native mining.

The power consumption was measured by the spot Watt usage when running idle and when mining under different configurations. The consumption was measured over a few minutes, this was enough time to get an estimate that could be extrapolated. For the devices that did not have batteries (NUC, Stationary and Rpi), a simple hardware power recorder was installed in the power outlet and read directly. For the Chromebook, a build-in utility was used (`chrome://power` in the URL-bar). With the Macbook we used a utility called *iStats Menu* [21], and with the Phone *Android Studio* and *Battery Historian* [3] were used.

To measure the actual impact of cryptojacking, we had one objective and one subjective approach. Objective measurements were collected with *Sysbench* [41] on the supported systems. On Android no comparable benchmarking tool to Sysbench was found and thus no benchmark data have been collected for the Phone.

For the subjective testing, a scale of annoyance was recorded for the different configurations. It ranged from *0 - not annoyed at all* to *4 - the device is practically unusable*. The devices were tested doing some common tasks such as surfing the web, streaming HD-video, using office applications and gaming.

## 4 Results

### 4.1 Relationship between hashrate and power

Figure 1 shows the highest recorded hashrate and power consumption for the NUC in both native and web mining mode when varying the number of threads used.

When mining natively it peaked at 2 threads, and decreased somewhat when adding more threads, probably due to increased overhead. During web mining adding more threads seemed to work well to increase the hashrate, although the 4th threads did not add much. The power consumption was very similar between native and web mining, at about three times the power consumption when idling. Interestingly, adding more threads to mine did not increase the power consumption by a whole lot.

We saw similar trends with the other devices as well, native mining outperforms web mining by a factor between 3-8, and in most cases uses less power. With the Mac, adding a second thread does not affect the hashrate beyond the margin of error, and the third thread adds less than a 10% increase. The fourth thread does not add anything at all. The Chromebook did not mine very efficiently, but the power consumption was also quite low. Both the Rpi and Phone scaled almost linearly when adding threads for mining. With the Stationary, we noticed that the power drain during mining was almost exactly the same whether mining natively or web, but when using the GPU the power consumption went up significantly. The next noticeable thing was that the hashrate peaked at four threads during native mining, GPU or no GPU. Adding even more threads

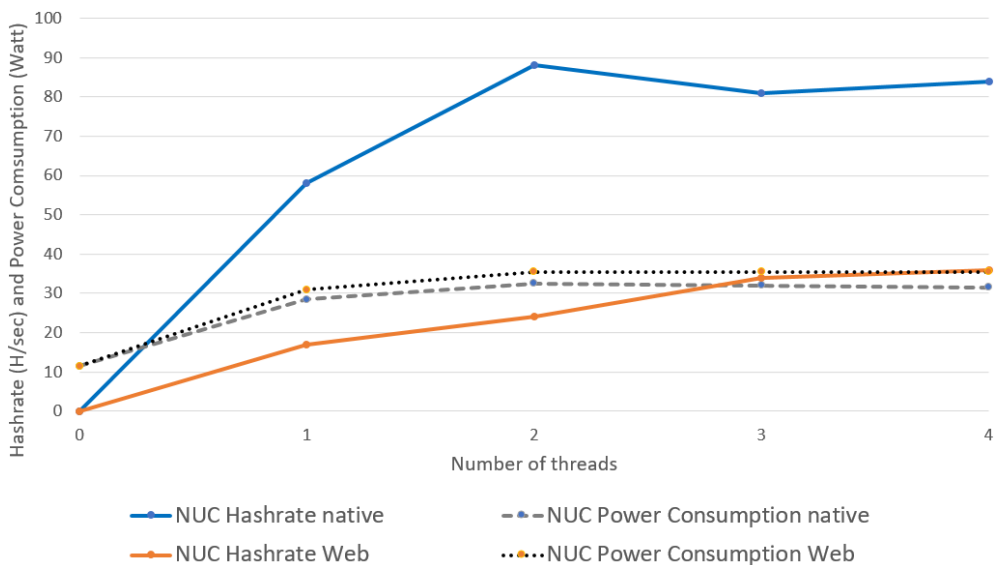
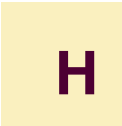


Fig. 1. Hashrate and power consumption of the NUC in native and web mining modes.

made the hashrate drop significantly. This was not true for web mining where the hashrate flattened out.

Figure 2 shows mining performance on all devices where the hashrate has been divided by the power consumption. The device that truly stands out is the Rpi at 3 threads, but this was not very efficient mining. The highest hashrate of the Rpi was 12.1, while the Stationary had 218.7 in native mode with the same number of threads.

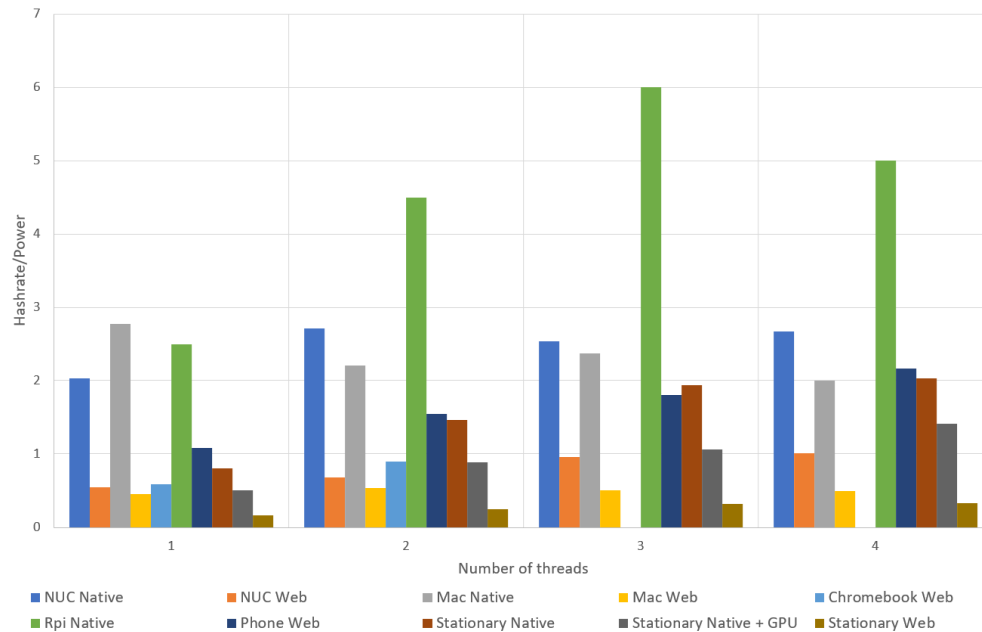


#### 4.2 Objective impact on performance and latency

To get measurements on the impact cryptomining had on device performance, Sysbench was used both while the devices were idling and mining. Sysbench works by running a large amount of math problems by the CPU to test how many events it can process in a given time. The number of threads for both mining and performing events were varied to see how this competition for resources turned out. Figures 3 and 4 show how mining affected performance and latency for the NUC. For the Mac and Stationary the trends were the same, when mining at full speed the performance of all devices drop to about half and the latency increases dramatically.

#### 4.3 Subjective impact on casual use

While a benchmark is very useful for getting an objective measurement on how a stressed device is affected by cryptojacking, this does not necessarily tell the whole story. If the users are not bothered by cryptominers using their CPUs,



**Fig. 2.** Relationship between hash efficiency and power consumption (higher is better).

they are less likely to do anything about it. In Figure 5 we have recorded the annoyance level for five of the devices when they were exposed to an increasing number of threads used for mining. The Phone was omitted since multitasking (running several apps in parallel) could not be done in practice while mining. The scores were all given by a single individual (one of the authors), the results are thus highly subjective and not very reliable, but they still give an indication of how mining might impact the perceived performance of cryptojacked devices.

The stationary was tested while performing several different tasks including steaming HD video, working with office documents and playing some game (real-time and turn-based strategy games, turn-based card game, real-time fighting game). When mining using the GPU in XMR-stak and no CPU threads the graphical I/O were severely impacted, to the point of making the whole computer unusable for anything else. However, when running as many as 7 out of 8 CPU threads the impact was negligible when simultaneously streaming HD-video and playing games. When running all 8 threads the impact was noticeable, but the computer was still fully usable. Even so, the increased latency was only significantly noticeable when performing context switches, such as loading new maps in a game, starting a new video, open new documents for editing and switching between different web sites rapidly. When staying within a single application, document or map for a long time the perceived performance hit was much less noticeable. The NUC and Mac were tested in much the same way, but with fewer games. The results were similar, both devices were slower than the Stationary even with no mining, but the reduced performance was only noticeable when us-

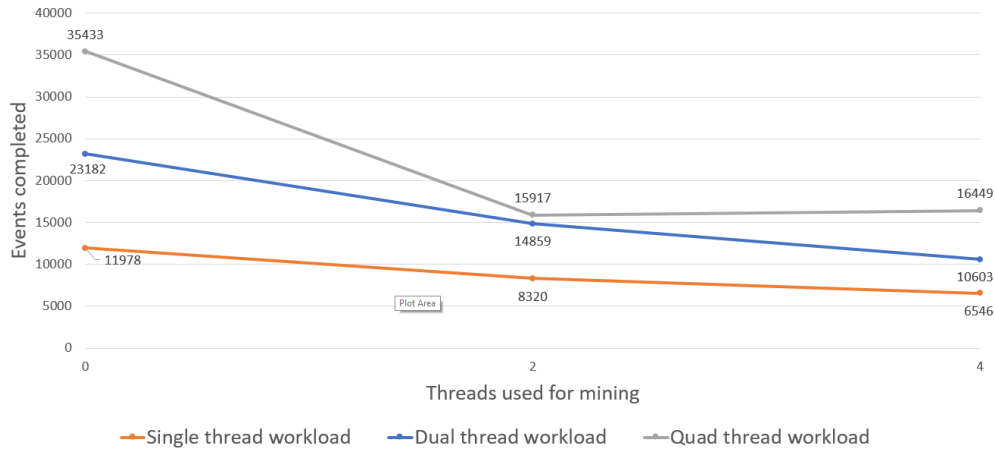


Fig. 3. NUC performance during mining (higher is better).

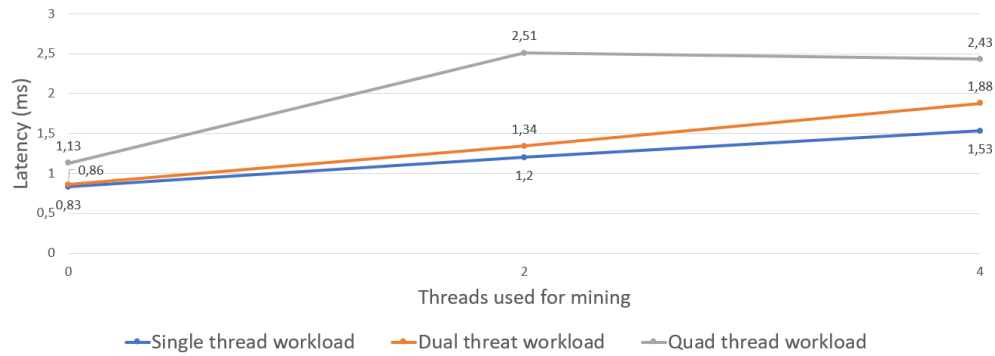


Fig. 4. NUC latency during mining (lower is better).

ing all available threads for mining. Latencies in load times and context switching were somewhat more noticeable on these devices. On the Chromebook, a web miner was set up running at 100% using both available threads. At the same time full HD videos were streamed and documents were opened in the browser. While the Chromebook was slower in comparison to the other machines, there was little performance impact from the web miner. The Rpi was only tested with a web browser running in the GUI. It was very slow to begin with and the mining made it virtually unusable.

## 5 Discussion

### 5.1 Best buck for the bang

An important aspect of cryptomining is how long it takes to accumulate the currency. In our case we worked with Monero, and during our experiment we ran





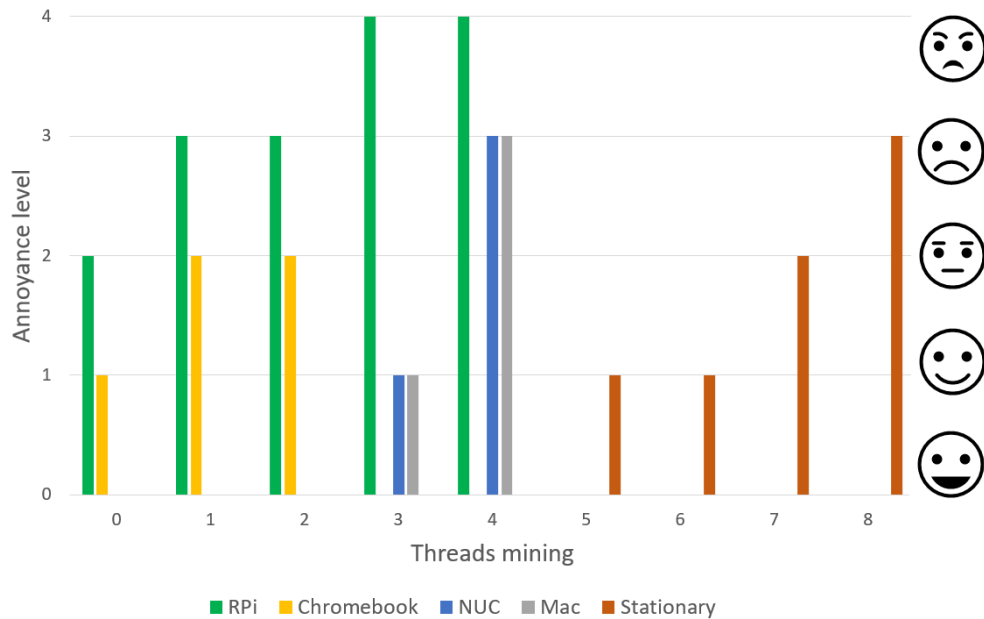
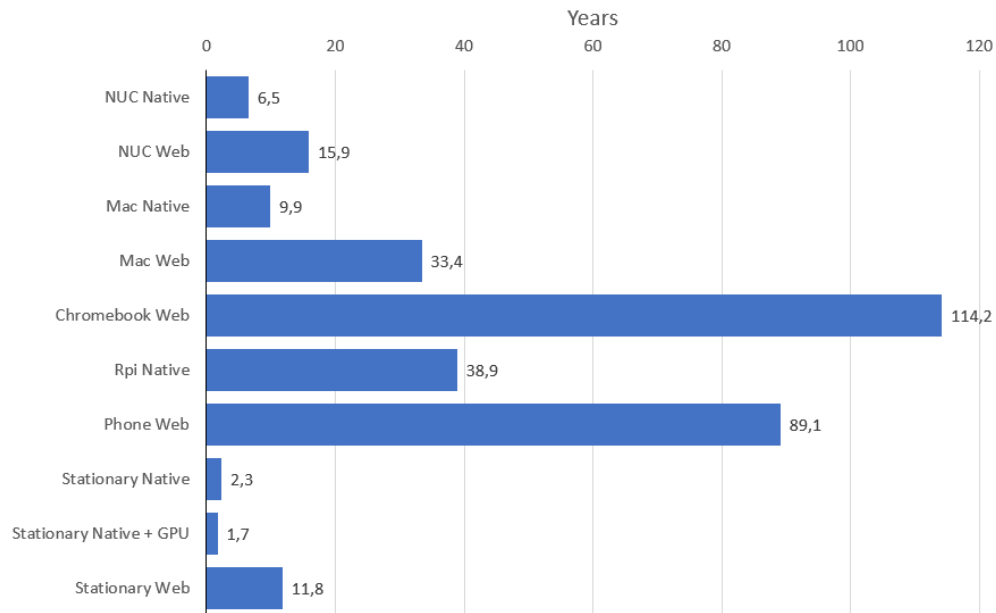


Fig. 5. Subjective experience of performance during mining.

about 223,680,786 hashes that generated about 0.0125 XMR. This indicates that it takes about 18 billion ( $1.789\,446\,288 \times 10^{10}$ ) hashes to produce a single coin. Note that this estimate is subject to the ever-changing nature of Monero mining. The Monero network limits the payouts to only once every second minute, thus the more participants in the network the more hashes are required to acquire one coin. Additionally, the payouts decrease over time, effectively increasing the amount of hashes necessary to acquire one coin. Since our work was done prior to the hard fork on the 9th of March 2019, we have looked at the potential value at that time. On the 1st of March 2019, one coin had a value of about 50 USD [9]. As with most cryptocurrencies, this number fluctuates a lot. On the 7th of January 2018, Monero peaked at about 500 USD. Even so, we can use our estimates to make a comparison between the devices and give an indication of how long a miner must run to yield valuable results on a single device. This is shown in Figure 6, where we have used the maximum recorded hashrate from our devices.

As can be seen from the figure, even greedy configurations of the script need years to mine a single coin even when running on high-end devices. In order to have any reasonable chance of making a revenue from this kind of mining, a cryptojacker would need to infect a large number of devices, preferably in native mode. For cyber criminals this means that there might be more profitable ways to use compromised devices, such as encrypting the data and demanding ransom, have the device participating in denial of service attacks or just leave it dormant until some use for it can be found.



**Fig. 6.** Years to mine a single Monero coin (lower is better).

Rational attackers will not only consider the potential revenue of cryptojacking, but also their own investments in order to perform the attack. Since cryptomining by itself is perfectly legal, a lot of the necessary code can already be found in the public domain, thus the job of the cryptojacking developers is to weaponize the code. This includes making it run stealthily and undetected by the user and perhaps include an auto update feature. As an example, we can look at *DeepMiner's* source code [15], which is freely available and constitutes about 1000 lines of code when excluding the cryptography itself. Assuming an investment of 18 USD per line [34], a rough estimate would be 18 000 USD for the weaponization. Alternatively, the attacker could buy off-the-shelf cryptojacking software from e.g. Dark Net markets. We have observed such items being sold for about 150 USD, although the prices vary between vendors and markets. However, there is a significant risk of being scammed when purchasing items on the Dark Net, which must also be considered in this equation. Once created, the cryptojacking software must be maintained and updated, which can be even more challenging than for legitimate software. Malware usually takes advantage of some vulnerability to infect other software, but such vulnerabilities are patched regularly, so there is a limited window of opportunity. Additionally, anti-virus programs and ad-blockers will quickly be on the lookout for cryptojacking signatures and behavior. There is also a significant cost of distribution, which is difficult to put a price tag on. Often the people distributing cryptojacking attacks are not the same people that wrote the software. This was the whole business idea behind Coinhive and its affiliates. Native miners have an even higher distribution cost in order to be installed and executed, whether through

social engineering, as a trojan, using an exploit or through physical access to the devices.

Some of these costs might be better measured in terms of time rather than money spent, which we often refer to as opportunity cost. Opportunity cost refers to the cost of doing one thing rather than another. Every hour, every buck and every bit of effort put into cryptojacking attacks could be used to do something else. As we have shown in Figure 6, an attacker would need long-term infection periods on a vast number of devices just to make some small revenue, and it is therefore understandable why the heyday of cryptojacking has gone by. Having said that, the cryptocurrency market might skyrocket again, making cryptojacking a very relevant threat again.

## 5.2 Limitations

One obvious limitation in our experiment is the relatively small sample size of six devices, running different operating systems and hardware configurations. Also, the subjective annoyance recording could have involved more people, but it is doubtful that the results would have been very different. Additionally, it was difficult to account for other running processes even when comparing idle states with mining activities. Prolonged mining would also create a temperature increase making the different devices behave differently, but this was not something we recorded.

Monero's Cryptonote mining algorithm was using memory blocks of 2 MB at the time of our experiment, meaning that in theory each CPU or GPU thread running CryptoNote would be most efficient if they could get 2 MB of cached memory for themselves. There are now plans for Monero to switch to another proof-of-work algorithm that requires miners to dedicate over 2 GB of RAM to the process, making cryptojacking attempts harder to hide [48] and probably useless on low-end devices. It would therefore be useful to repeat the experiment as the algorithm changes to see how this affects the impact on different device types.

## 5.3 Related work

Cryptojacking is a relatively new phenomenon, hence there has been limited research on this kind of threat prior to 2018. Musch et al. [27] wrote an extensive report on web-based cryptojacking this year, describing how to identify mining scripts among the Alexa [2] Top 1M web sites and expected mining revenues. They found that about 1 out of 500 web sites contained miners and that there was moderate profit to be made at that time. Tahir et al. [42] have done a later study on Alexa Top 50K web sites looking for cryptojacking, and also discovered that mining-prevention plugins often fail to detect such scripts.

Eskandari et al. [16] have analyzed the profitability of cryptojacking web sites using a real-world data set, showing that over a period of three months little revenue can be earned. They also discuss whether the web site visitors giving consent to mine have a clear mental model of what they are paying.

This is supported by Carlin et al. [6], who discuss the legality of cryptomining, referring to UK legislation.

Similarly to our work, Saad et al. [36] have analyzed the impact cryptojacking has on system resources on various devices, in their case three different laptops and one smart phone, but only for web-based mining. They also examined the economic basis for cryptomining as an alternative to advertisement on web sites, and concluded that cryptomining was not a feasible alternative. In parallel to this work, Papadopoulos et al. [31] studied the profitability of in-browser mining and developed a testbench that ran on a Linux desktop. They found that advertisements were 5.5 more profitable than web-cryptomining, but that hybrid solutions would allow for maximum profits. However, on the user side the device temperature and power consumption would increase 52.8% and 2X respectively on a desktop computer.

Hong et al. [19] have done a systematic study on cryptojacking and present a detector that automatically tracks mining scripts. This detector has been applied to the Alexa Top 100K list, and they estimated a danger to more than 10M web users and extra spending of electricity that is similar to powering a city. Kharraz et al. [22] present another detector that has been applied to Alexa Top 1M and conclude that cryptojacking operations can be detected with minimal human interventions. Konoth et al. [23] did another crawl of Alexa Top 1M, but in contrast to related studies, they analyzed more than just the landing pages. They found that only 3.86% of cryptomining web sites informed their users of this activity, and that the most profitable web site was earning 17K USD a month from 29M visitors. However, the vast majority of web sites were making very little revenue from cryptomining. Pastrana and Guillermo [33] have conducted a longitudinal study where they analyze about 1M malicious miners to see where the profit goes in the underground economy. They found that at least 56M USD have gone to criminals. A broader paper on how to monetize from web attacks has been published by Nguyen et al. [29], who also suggest countermeasures to this. A paper by Norman [30] also focus on many of the same countermeasures. In a review paper by Al Hajri et al. [1] a particular warning goes to enterprises due to their broad attack surface.

Sigler [38] show the trend where web/script-based cryptojacking attacks became more favorable than the native counterpart due to their easiness. Zimba et al. [47] have proposed how digital autopsies on both native and web-based miners, as well as extortion malware, can be performed. They found that most of the scripts they analyzed were very simple and relied on communication to Command and Control servers to receive further directives.

Bijmans et al. [5] have performed a recent large study on organized cryptojacking. They discovered that cryptojacking campaigns have been heavily underestimated in previous studies, and that third-party software such as Wordpress is the new preferred method of spreading infections. After having crawled about 20% of the Internet, they estimate cryptomining without user consent in 0.011% of all domains, mostly prevalent in adult content sites. They also describe nu-

merous hiding techniques present in scripts making them more difficult to detect by blocking application.

## 6 Conclusion

Related to our first research question, the experiment measurements show that native mining clearly outperforms web mining. Though relatively simple devices such as the Raspberry Pi had the highest hashrate per Watt, mining simply takes too much time on these that they are desirable targets. When we measured performance and latency during mining using an objective benchmarking tool, these values went down as expected as we added more mining threads, making cryptojacking easily detectable on an already stressed device. However, we got a somewhat different impression when the devices had more casual usage patterns involving video streaming, office apps, surfing and games. On devices with many available threads, the mining was hardly noticeable as long as the algorithm did not take all available resources. Since most regular users are accustomed to natural performance variations, it can therefore be very difficult to naturally recognize cryptomining running in the background.

By addressing our second research question we saw that it was difficult to justify a sound attacker business model. There was a relatively large market for it up until 2018, but as the cryptocurrencies fell in value, the cyber criminals started to revert back to other, more profitable ways of making a revenue. It is important to remember that if the cryptocurrency markets should resurge, it is likely that cryptojacking will follow suit. The attacks are relatively easy to carry out, and since they seldom cripple the infected devices, users might not detect the mining or bother to do something about it. Luckily, the security industry is now more aware of this threat and there are many tools that can protect the users.

## References

1. Al Hajri, H.H., Al Mughairi, B.M., Hossain, M.I., Karim, A.M.: Crypto jacking a technique to leverage technology to mine crypto currency. *International journal of academic research in business and social sciences* **9**(3) (2019)
2. Alexa. <https://www.alexa.com/>, last accessed: 2019-08-23
3. Analyze power use with battery historian. <https://developer.android.com/topic/performance/power/battery-historian>, last accessed: 2019-05-24
4. App store review guidelines. <https://developer.apple.com/app-store/review/guidelines/>, last accessed 2019-08-22
5. Bijmans, H.L., Booiij, T.M., Doerr, C.: Inadvertently making cyber criminals rich: A comprehensive study of cryptojacking campaigns at internet scale. In: 28th USENIX Security Symposium. pp. 1627–1644 (2019)
6. Carlin, D., Burgess, J., O’Kane, P., Sezer, S.: You could be mine (d): The rise of cryptojacking. *IEEE Security & Privacy* (2019)
7. Coinhive.com. <https://coinhive.com/>, last accessed: 2019-04-08

8. Discontinuation of coinhive. <https://coinhive.com/blog/en/discontinuation-of-coinhive> (2019), last accessed: 2019-05-24
9. Top 100 cryptocurrencies by market capitalization. <https://coinmarketcap.com> (2019), last accessed: 2019-08-22
10. Mine monero from your browser. <https://coinwebmining.com/browser-miner/monero>, last accessed: 2019-05-24
11. Cova, M., Kruegel, C., Vigna, G.: Detection and analysis of drive-by-download attacks and malicious javascript code. In: Proceedings of the 19th international conference on World wide web. pp. 281–290. ACM (2010)
12. Cox, J.: Creators of in-browser cryptocurrency miner 'coinhive' say their reputation couldn't be much worse. [https://motherboard.vice.com/en\\_us/article/vbpbz4/creators-of-in-browser-cryptocurrency-miner-coinhive-say-their-reputation-couldnt-be-much-worse](https://motherboard.vice.com/en_us/article/vbpbz4/creators-of-in-browser-cryptocurrency-miner-coinhive-say-their-reputation-couldnt-be-much-worse) (2018), last accessed: 2019-05-24
13. Cpuminer-multi. <https://github.com/tpruvot/cpuminer-multi>, last accessed: 2019-05-24
14. Dean, M.: 5 best cryptojacking blockers to use on your windows pc. <https://windowsreport.com/cryptojacking-blockers/>, last accessed: 2019-05-24
15. Deep miner. <https://github.com/deepwn/deepMiner>, last accessed: 2019-05-24
16. Eskandari, S., Leoutsarakos, A., Mursch, T., Clark, J.: A first look at browser-based cryptojacking. 2018 IEEE European Symposium on Security and Privacy Workshops (2018)
17. Internet Organised Crime Threat Assessment (IOCTA) 2018. <https://www.europol.europa.eu/activities-services/main-reports/internet-organised-crime-threat-assessment-iocta-2018> (2018)
18. Frigioiu, A.: Crypto miners: the rise of a malware empire. <https://blog.avira.com/crypto-miners-coinhive-malware-empire/> (2018), last accessed: 2019-05-24
19. Hong, G., Yang, Z., Yang, S., Zhang, L., Nan, Y., Zhang, Z., Yang, M., Zhang, Y., Qian, Z., Duan, H.: How you get shot in the back: A systematical study about cryptojacking in the real world. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security. pp. 1701–1713. ACM (2018)
20. Hron, M.: Protect yourself from cryptojacking. <https://blog.avast.com/protect-yourself-from-cryptojacking> (2018), last accessed: 2019-08-23
21. istat menus. <https://itunes.apple.com/us/app/istat-menus/id1319778037?mt=12>, last accessed: 2019-05-24
22. Kharraz, A., Ma, Z., Murley, P., Lever, C., Mason, J., Miller, A., Borisov, N., Antonakakis, M., Bailey, M.: Outguard: Detecting in-browser covert cryptocurrency mining in the wild. In: The World Wide Web Conference. pp. 840–852. WWW '19, ACM, New York, NY, USA (2019). <https://doi.org/10.1145/3308558.3313665>, <http://doi.acm.org/10.1145/3308558.3313665>
23. Konoth, R.K., Vineti, E., Moonsamy, V., Lindorfer, M., Kruegel, C., Bos, H., Vigna, G.: Minesweeper: An in-depth look into drive-by cryptocurrency mining and its defense. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security. pp. 1714–1730. ACM (2018)
24. Liao, S.: Calendar app in mac app store mines cryptocurrency in the background. <https://www.theverge.com/2018/3/12/17110810/apple-app-store-mac-cryptocurrency-monero-calendar-2-qbix> (2018), last accessed: 2019-08-22
25. Monero miner for web browsers. <https://minero.cc/>, last accessed: 2019-05-24
26. Monero.com. <https://monero.org/>, last accessed: 2019-05-24
27. Musch, M., Wressnegger, C., Johns, M., Rieck, K.: Web-based cryptojacking in the wild. arXiv preprint arXiv:1808.09474 (2018)

28. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system. <https://bitcoin.org/bitcoin.pdf> (2008), last accessed: 2019-08-23
29. Nguyen, V.L., Lin, P.C., Hwang, R.H.: Web attacks: defeating monetisation attempts. *Network Security* **2019**(5), 11–19 (2019)
30. Norman, J.: How not to become a crypto-jacking statistic. *Computer Fraud & Security* **2019**(4), 18–19 (2019)
31. Papadopoulos, P., Ilija, P., Markatos, E.P.: Truth in web mining: Measuring the profitability and cost of cryptominers as a web monetization model. arXiv:1806.01994v1 (2018), last accessed: 2019-05-24
32. Partz, H.: Coinhive code found on 300+ websites worldwide in recent cryptojacking campaign (2018), last accessed: 2019-05-24
33. Pastrana, S., Suarez-Tangil, G.: A first look at the crypto-mining malware ecosystem: A decade of unrestricted wealth. arXiv preprint arXiv:1901.00846 (2019)
34. Pedersen, P.: The open source community as a top 100 country. <http://www.inside-open-source.com/2007/11/open-source-community-as-top-100.html>, last accessed: 2019-08-23
35. R uth, J., Zimmermann, T., Wolsing, K., Hohlfeld, O.: Digging into browser-based crypto mining. Chair of Communication and Distributed Systems, RWTH Aachen University (2018), last accessed: 2019-05-24
36. Saad, M., Khormali, A., Mohaisen, A.: End-to-end analysis of in-browser crypto-jacking. arXiv:1809.02152v1 (2018), last accessed: 2019-05-24
37. Saberhagen, N.v.: Cryptonote v 2.0 (2013), last accessed: 2019-05-24
38. Sigler, K.: Crypto-jacking: how cyber-criminals are exploiting the crypto-currency boom. *Computer Fraud & Security* **2018**(9), 12 – 14 (2018). [https://doi.org/https://doi.org/10.1016/S1361-3723\(18\)30086-1](https://doi.org/https://doi.org/10.1016/S1361-3723(18)30086-1), <http://www.sciencedirect.com/science/article/pii/S1361372318300861>
39. Supportxmr.com. <https://www.supportxmr.com/>, last accessed: 2019-05-24
40. Internet security threat report. <https://www.symantec.com/en/sg/security-center/threat-report> (2019), last accessed: 2019-05-24
41. Sysbench. <https://github.com/akopytov/sysbench>, last accessed: 2019-05-24
42. Tahir, R., Durrani, S., Ahmed, F., Saeed, H., Zaffar, F., Ilyas, S.: The browsers strike back: Countering cryptojacking and parasitic miners on the web. In: IEEE INFOCOM 2019 - IEEE Conference on Computer Communications. pp. 703–711 (April 2019). <https://doi.org/10.1109/INFOCOM.2019.8737360>
43. Historical trends in the usage of client-side programming languages for websites. [https://w3techs.com/technologies/history\\_overview/client\\_side\\_language/all](https://w3techs.com/technologies/history_overview/client_side_language/all), last accessed: 2019-05-24
44. Plugin tag: mining. <https://wordpress.org/plugins/tags/mining/>, last accessed: 2019-05-24
45. Wu, L.: Monero-mining hiddenminer android malware can potentially cause device failure. TrendMicro Last accessed: 2019-05-24
46. Xmr-stak: Cryptonight all-in-one mining software. <https://github.com/fireice-uk/xmr-stak>, last accessed: 2019-05-24
47. Zimba, A., Wang, Z., Chen, H., Mulenga, M.: Recent advances in cryptovirology: State-of-the-art crypto mining and crypto ransomware attacks. *KSII Transactions on Internet and Information Systems (TIIS)* **13**(6), 3258–3279 (2019)
48. Zmudzinski, A.: Monero developers consider adopting new proof-of-work algorithm in october. <https://cointelegraph.com/news/monero-developers-consider-adopting-new-proof-of-work-algorithm-in-october> (2019), last accessed: 2019-08-23

**I: ‘Cyber Attacks for Sale’**

©2019 IEEE. Reprinted, with permission, from Per Håkon Meland and Guttorm Sindre, 2019 International Conference on Computational Science and Computational Intelligence (CSCI), December 2019.





# Cyber Attacks for Sale

Per Håkon Meland

Norwegian University of Science and Technology  
and SINTEF Digital  
Trondheim, Norway  
per.hakon.meland@ntnu.no

Guttorm Sindre

Norwegian University of Science and Technology  
Trondheim, Norway  
guttorm.sindre@ntnu.no

**Abstract**—The infamous darknet hosts an underground economy for illegal goods and services, some of which can be purchased and used for cyber attacks. By analyzing the properties and popularity of such items, we can get indications about the type and capabilities of potential attackers, what assets they are targeting and which vulnerabilities they are likely to exploit. We have conducted an online study of eleven marketplaces residing within the darknet, addressing what kind of cyber attack items are available and where the profit lies. The results have been used to create a detailed categorization of items, showing a distribution based on item type and availability. This has been compared to the number of sold items and revenue from four of the marketplaces, and we discuss these different views. Aided by related studies, we have identified trending cyber threats such as phone hacking, information theft and bitcoin stealing.

**Type of submission:** Full/Regular Research Paper

**Symposium:** CSCI-ISCW

**Index Terms**—cyber threat, darknet, underground economy

## I. INTRODUCTION

The infamous darknet hosts an underground economy for illegal goods and services, where the identities of vendors and buyers stay hidden through cryptographic mechanisms. Within popular marketplaces residing here, there are numerous types of software and services that are sold for the purpose of performing cyber attacks, and which allow actors with limited technical expertise and resources to obtain malicious capabilities. Knowledge of mechanisms and trends in this market can improve our situational awareness about threats towards our systems [1], i.e. the popularity of malicious digital goods may indicate the type and capability of potential attackers, what assets they target and which vulnerabilities they are likely to exploit. This is comparable to the military arms market; high demand for aggressive weapons indicates a potential threat. If the buyer of these weapons happens to be a group or country with a grudge against you, then it is wise to install defense mechanisms that can counter such weapons. In the cyber world, these dynamics works at a much higher pace, giving the defenders a preparation time of maybe a few days only.

The purpose of this paper is to provide an overview of contemporary marketplaces and items related to cyber attacks. We do this by addressing the following research questions:

- RQ1: What kind of cyber attack items are available on the darknet marketplaces?
- RQ2: What are the most profitable items for the vendors?

Answering these might give us *forward-looking indicators* [2] of the cyber threat landscape, and according to Broadhurst et al. [3], a way for tracking trends in potential victimization.

Section II describes how we have conducted our study and the research space. Section III presents the categories, exclusions and different views on the market, which are further discussed in Section IV. Section V concludes the paper.

## II. METHOD

We have conducted an online study of the virtual community residing on darknet marketplaces, with a specific focus on tools and services that can be purchased and used for cyber attacks. Kozinets [4] uses the term *netnography* for such online studies, and we have followed his guidelines for planning, ethical considerations, data collection and interpretation. It was important to us that the research would not cause harm to individuals or groups. Users on the darknet are anonymous, and we would not collect any data that could be used to reveal their identities. We have also been conscious not to put ourselves or others at risk. In practice this means a passive data collection of archival data already available in the public space. To avoid supporting illegal activities we have not purchased anything. Finally, we have not tried to deceive, intimidate or confuse people within this research space, e.g., pretending to be a vendor, customer (though we had to create user accounts), malware software writer or marketplace administrator.

DarknetLive [5], found to have the most up-to-date index of TOR market links and mirrors, was used to identify marketplaces for our study, supplemented with a few extra links from TheDarkWebLinks [6] and DarknetStats [7]. Screened out dead and seized markets, as well as irrelevant ones (e.g., only dealing drugs, no malware), yielded the sample shown in Table I. Data from this sample were collected during the month of September 2019. For each market we identified the relevant inventory categories, and did a manual inspection of the items enlisted in each of these. Due to variance in functionality between the marketplace platforms, the data recorded from each market differed somewhat. We could record item name and price for almost all, while for instance number of successful sales and views were only visible for some (detailed in Table I). Where possible, we filtered out items with zero sales to

TABLE I  
MARKETPLACES INCLUDED IN OUR STUDY.

Name	Description and data recorded	Selected categories (available items)
Apollon Market	Established in March 2018, selling a large variety of items (12 836 in total) in all kinds of categories, but mostly <i>drugs</i> , <i>digital goods</i> and <i>fraud</i> . We recorded relevant items, their price and number of sales, but filtered out items with zero sales.	Software and malware (72) Services - Social engineering (16) Services - Hacking (38) Services - Cracking (6)
Berlusconi Market	Established in July 2017 and had the largest inventory (150 034 items) in our sample until it died right after our observation period. Clearly dominated by <i>drugs</i> and <i>counterfeit</i> items, but contained digital goods as well. We recorded items, sales and price. Filtering: At least one sale per item, vendor activity within the last 30 days.	Software and malware (1 459) Digital products (8 555) - Fraud software Services (2 759)
Canadian HeadQuarters	Established early in 2018. The market has a particular focus on <i>fraud related items</i> (2 117 items, such as bank logs, personal information profiles, utility bills, passports and bar code generators) and one of the few markets we saw that was not dominated by <i>drugs</i> (184 items). We recorded all relevant items and price.	Fraud - Scampages (84) Services - Other (87)
Cave Tor	A small marketplace of unknown origin with 464 items in total, whereas <i>financial services</i> (cloned credit cards, fake identity cards, etc) and <i>drugs</i> were the main categories. We recorded 31 <i>hackers-for-hire</i> services and 1 <i>phishing kit</i> , but these were not enlisted with price.	Service (85)
DarkBay	A market named DarkBay was originally shut down in 2014, and it is unclear whether the current operating is related. It had 4 213 items where <i>guides &amp; tutorials</i> (44%) was the most comprehensive category, followed by <i>digital goods</i> (99,8% e-books) and <i>drugs</i> . We recorded relevant items and price.	Fraud software (2) Services (12) Software and malware (2)
Dream Alt	Established early in 2019 and should not be confused with the original Dream Market that was shut down in March 2019. Out of 21 646 items in total, 40% were found under <i>digital goods</i> (32% e-books) and 34% under <i>drugs &amp; chemicals</i> . We recorded relevant items and price.	Digital goods - Software (220) Digital goods - Security (110) Services - Hacking (374)
Empire Market	Established around April 2018 and regarded as the successor of the seized Alphabay market. Out of 49 501 items in total, 68% were related to <i>drugs &amp; chemicals</i> . We recorded relevant items, number of views and successful sales per item. Filtering: At least one sale per item.	Software and malware (364) Services - Social engineering (108) Services - Other (237) Digital Products - Other (1 443) Fraud - Other (569) Guides & tutorials - Hacking (363)
Grey Market	Officially launched July 2019, enlisting 3 360 items in total. Out of these, 62% were related to <i>digital</i> and 33% related to <i>drugs</i> . We recorded relevant items, number of views and successful sales per item.	Digital - Information - Other (1 160) Digital - Fraud - Other (12) Digital - Fraud - Software (140) Service - Hacking (32) Service - Other (68)
Samsara	Samsara opened in July 2019 and is an updated and rebranded version of Dream Market. Out of 28 859 items in total, 54% were related to <i>drugs</i> and 43% to <i>digital goods</i> . We recorded relevant items and price.	Digital goods - Hacking (209) Digital goods - Fraud (340) Digital goods - Software (627) Services - Hacking (23)
Tochka	A.k.a. <i>Point</i> , has been operating since 2015. We found 6 669 items in total, divided into categories <i>drugs</i> (70%), <i>prescriptions</i> (21%) and <i>steroids</i> (5%) (the remaining 4% was unaccounted for). Under <i>drugs</i> , there was a subcategory <i>other</i> that contained relevant digital goods. We recorded relevant items and price.	Drugs - Other (389)
Undermarket 2.0	Marketplace of unknown origin where vendors are enlisted under each category, and items under each vendor. The total number of vendors was 70, where <i>carding</i> (17%) and <i>drugs</i> (17%) were the most prominent categories. We recorded relevant vendors, their items, prices, successful sales and number of reviews.	Services (9)

let the buyers help us rule out untrustworthy or undesirable items. Observations were listed in a spreadsheet, all currencies converted to USD, and we took screenshots of interesting items and wrote descriptive and reflective field notes during the study.

### III. RESULTS

#### A. An Overall Inventory of Cyber Attacks

We found the granularities of the categories used in the marketplaces to be rather low. In order to get a more detailed view on what kind of malicious cyber items were available on the marketplaces, we defined a more specific categorization of software and services that all recorded items were mapped against. The following bullet list describes this categorization,

and shows the percentage of items put in each from the total of 885 we considered relevant. Where suitable, we have adopted definitions from the *Structured Threat Information Expression* (STIX) framework [8].

- Ransomware (4.1%): Encrypts files on a victim's system, demanding payment in return for the access codes required to unlock files [8]. Products offered were typically source code or customized binaries.
- Remote Access Trojans (RAT) (3.8%): A trojan horse capable of controlling a machine through commands issued by a remote attacker [8]. We observed RATs that could activate webcams, take screenshots, monitor user behavior or access sensitive information.
- Keyloggers (4.1%): Malware that monitors keystrokes

and either records them for later retrieval or sends them back to a central collection point [8].

- Scanners and sniffers (1.4%): Network analysis tools typically used during attack reconnaissance. Scanners find IP addresses and look for vulnerable ports, sniffers intercept and analyze network packages.
- Stealers and grabbers (8.1%): Exploit clipboard data. A stealer will look for bitcoin addresses, and replace these with the attacker's account when pasting. Grabbers look for usernames, passwords, bank accounts, etc. that can be stolen or manipulated.
- Hardware stealers (0.5%): Physical attack devices such as custom-made USB-sticks used to copy/steal data or inject malware.
- Account/password crackers (12.4%): Software used to brute force into specific operating systems or user accounts of popular web sites.
- Phone hacking (6.6%): Toolsets used to hack into phones or other devices running an Android/iOS operating system. This category also includes RATs especially made for phones/tablets.
- Cryptominers (2.7%): Malware that steals a system's resources [8], such as code and binaries that illicitly make use of CPU/GPU cycles, RAM and power to mine cryptocurrencies on behalf of the attacker.
- Exploit kits (0.9%): Tools used to automate attacks on popular applications with specific vulnerabilities. These were either sold as collections or single-system attack software.
- Hack packs (9.7%): Large collections of the various hacking tools mentioned here, along with guides. These are often several GBs in size and can contain hundreds of applications.
- Wifi hacking (2.7%): Software for setting up fake wireless access point software or hacking directly into wireless networks.
- Phishing kits (11.6%): Ready-made scam-pages of popular web sites, sold either as collections or individual sites.
- Botnet software (3.4%): Malware for forming and administration tools for botnets, which are mostly used to execute DDoS attacks.
- Injection tools (1.8%): Tools to generate and send malicious input into web pages that gets executed by an interpreter. We saw mostly SQL injection tools.
- Spamming kits (2.4%): Software for sending out large amounts of emails or SMSs to specific addresses. Letter templates in various languages were also registered in this category.
- Spamming/bombing services (3.2%): Services that will send out a specific number of emails or SMSs. Usually in the range of tens of thousands.
- Hackers-for-hire (19.9%): Diverse hacking services, such as breaking into specific social media accounts, changing school grades or site takedowns.
- DDoS services (0.2%): Specific services for taking down sites through DDoS attacks, often advertised with down-

time guarantees.

- Botnet services (0.5%): Rent control over a botnet for a specific amount of time.
- RAT services (0.1%): High-level remote access to number of already compromised computers.

Figure 1 shows how the items are distributed by type and among the eleven different marketplaces. In terms of availability, the top three categories were *hackers-for-hire*, *account/password crackers* and *phishing kits*. Items from Sam-sara dominated the two former (36% and 49%), and Canadian HQ offered 82% of the items from the latter. If we remove these two marketplaces from the sample, the top three becomes *hackers-for-hire*, *stealers and grabbers* and *account/password crackers*.

A general observation is that the type of items and number of items are unevenly distributed among the marketplaces.

### B. Exclusions

Among the inspected items, there were several types that can be deemed malicious, but not used directly for a cyber attack and therefore excluded from our study. Examples include *credit card data*, *zero-day exploits and vulnerabilities for sale*, *anonymity tools* (private SOCKS, cleaners, antidection), *software licenses*, *hacked user accounts and digital identities* (studied in detail by Wehinger [9]), *money laundering services*, *tutorials and guides*, *contact details of experienced hackers*, *physical skimming devices*, *automatic account creators*, *fake social media followers and web-site visitors* (or popularity-as-service), *search engine optimizers* (SEOs) and *gift card generators*. Also, we excluded *binders*, used to combine a malicious payload with an executable file, and *crypters*, which can obfuscate malicious code, though both of these types were commonly found within *hack packs*.

### C. A better view on the market

The availability and distribution of items is one view on the market, but other studies [9], [10] have indicated that fake items and scams thrive on the darknet. Therefore, we made use of the marketplaces that reported number of successful sales and mapped these to the same categories. In Figure 2, we show the number of sales per category from the Apollon, Berlusconi, Empire and Grey market. Out 371 items with 6257 sales in total, we can see here that the top three cyber attack items are *phone hacking* (26%), *hack packs* (20%) and *stealers and grabbers* (17%).

Another way of looking at the market is where the revenue lies. Multiplying the number of successful sales with the latest listing price per item, we estimated what vendors have earned from sales. In Figure 3, the topmost (blue) bars in each category show the accumulated revenue, and the lower (red) bars show the average revenue per item. The standard deviations are shown as extensions to the red bars, indicating how much the average revenue vary between individual items within the same category. The main takeaway from this view is that *hackers-for-hire* are now back on top due to a high average price. There was one item in particular that had a lot

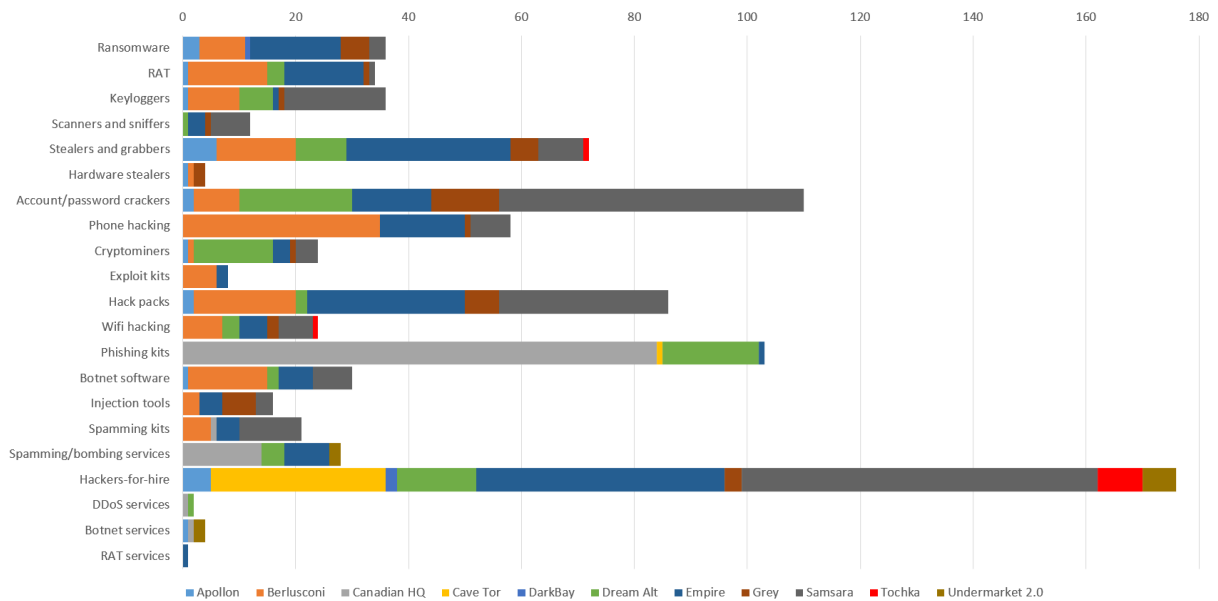


Fig. 1. Categorical distribution of items from eleven marketplaces.

of sales (311 successful sales, 39% of the total revenue). In the following three places we find the same top three cyber attack items as with the number of sales, no surprise since these items have a similar average price (97-113 USD). For all of these, the standard deviation is quite large, as the number of sales is unevenly distributed among the items. The most sold items also tend to be the most pricy ones, benefiting from buyers that will use the high number of sales as a sign of legitimacy and therefore are willing to pay more. A similar trend could be seen from the ratio between number of sales and views, where the most successful items stayed between 0.05 and 0.10, while unpopular items were several factors of ten lower.

As seen from Figures 1 and 2, the Apollon, Berlusconi and Grey markets are weak when it comes to availability and sales of services. Undermarket 2.0 reports number of successful sales per vendor, with two vendors that specialized in cyber attack services such as *DDoS*, *spamming*, *information theft* and *account hacking* at the time of our observations. The sales figures of these were 32 540 and 72 259, exceeding the combined sales of all relevant items in the four marketplaces stating those figures. Either, these are among of the most successful cyber attack service providers on the darknet, or the numbers are fabricated and the marketplace a scam. Some darknet forum posts claim the latter, and the number of reviews (mostly positive) for each of these two vendors are exactly 85% of the number of sales, possibly indicating that reviews are automatically generated.

#### IV. DISCUSSION

We have addressed our first research question by categorizing and looking at the distribution of cyber attack items

found in our largest sample of eleven marketplaces. The second research question is addressed by looking at number of successful sales and prices from a smaller sample of four marketplaces. Except for *hackers-for-hire*, the top items differ between the views, and an obvious limitation is the difference between the samples. Therefore, it is debatable which view, if any, gives us the best indication of what kind of cyberthreats we should worry about based on darknet trade. In our opinion, there is more confidence in the view based on sales. This is based on a more qualitative assessments of the items offered in the marketplaces that do not state sales figures, where we noted the following:

- Many of the offered items have descriptions which are short, vague or written in poor English, hence difficult for potential buyers to assess.
- Only a few vendors have many reviews, and these seem to be obtained more from drugs and carding items, less from cyber attack items.
- Many vendors put out the same or similar items multiple times, seeking visibility by flooding the market.
- Many of the items sold seem to have little value. E.g. the tools are old or can be found for free on the surface web (e.g. *Oracle VirtualBox*, the *Mirai* source code, various password crackers).

In contrast, items with a significant amount of sales have clearer descriptions, prices seem more appropriate and duplicate entries are more sparse.

Our dataset consists of a snapshot from September 2019, lacking trends over time. In previous work [11] we studied availability and price fluctuations for ransomware over a longer period aided by archival datasets. Such studies are

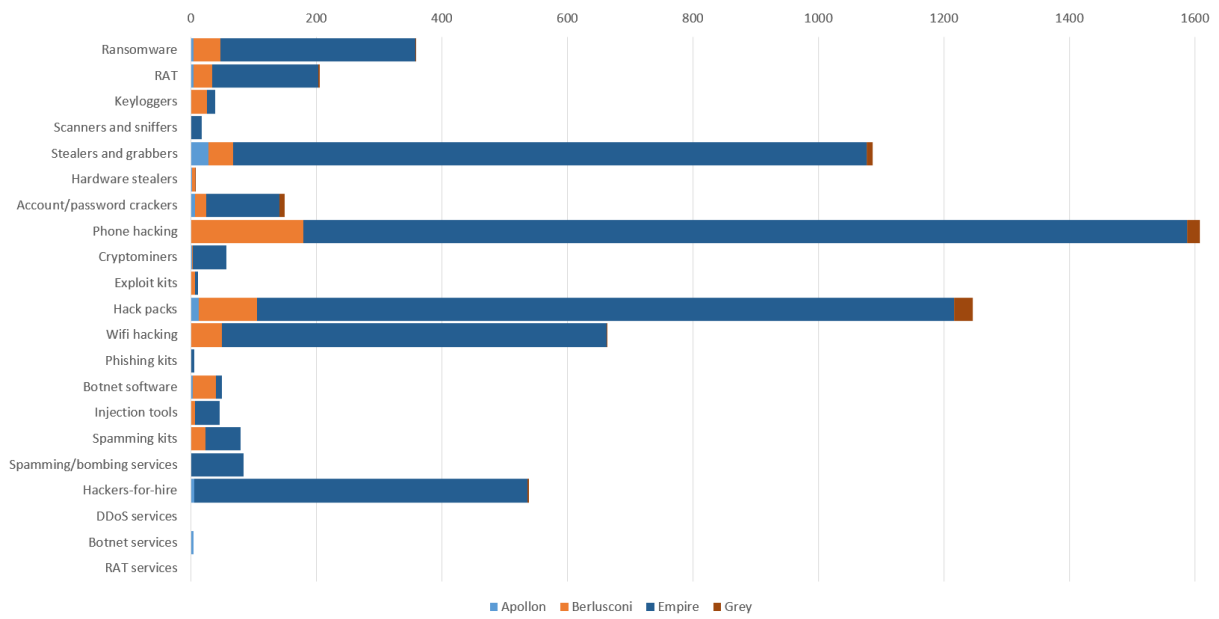


Fig. 2. Number of successful sales per category in four marketplaces.

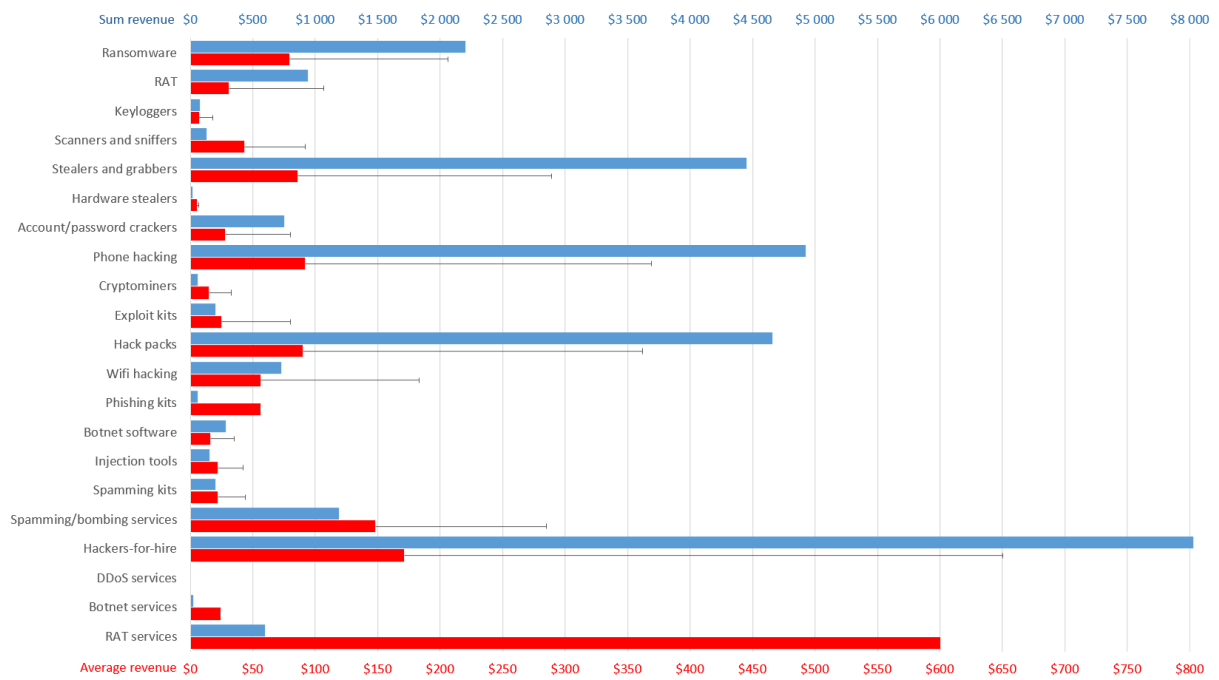


Fig. 3. Accumulated revenue per category and average revenue per item from four marketplaces.

interesting for projections, but also increasingly difficult to perform since law enforcement agencies are more effectively taking down marketplaces. The majority of marketplaces in our largest sample have been established quite recently, while infamous ones such as Silk Road, AlphaBay, Hansa, Dream and WallStreet are now gone. As future work it would be interesting to repopulate the categories with new observations, and analyze how vendors transition themselves in this volatile environment of marketplaces.

#### A. Related work

Our results can be more informative in the light of related work. In 2014, Ablon et al. [12] classified and exemplified hacking tools and services on black and gray markets. Their approach was to interview subject-matter experts and conduct a literature review. Their classification is more abstract than our categories and lacks elements such as *stealers and grabbers*. For exploit kits and zero-day vulnerabilities, they were able to show price developments over time. The year after, Thomas et al. [13] surveyed existing research in order to develop a taxonomy for reasoning about the flow of capital within the underground economy, making estimations about price and revenue from underground studies and their own investigations. This taxonomy has a broad cybercrime spectrum, but not our level of detail. They also showed that a lot of published studies have an unknown collection methodology. Broadhurst et al. [3] reviewed malware trends on darknet markets and categorized digital products found on Dream Market between September 2017 and April 2018. Again, these categories are fewer and more abstract than ours, but we can for instance see a comparative increase in the presence of keyloggers and a general increase in average prices. Van Wegberg et al. [14] have carried out a six-year longitudinal study tracking the evolution of commoditization on eight marketplaces up until 2017 (all now defunct). Their categorization was based on earlier work by Soska and Christin [15], which is less detailed than ours as well. The way they estimated sales figures was based on customer feedback, which is less accurate than the exact sales from our smallest marketplace sample. They found that ransomware was dominating the malware category, which is different from our data where *stealers and grabbers* prevail. McGuire [16] analyzed fifteen darknet platforms between November 2018 to March 2019. Only Empire and Berlusconi were common with our sample, and their top three were *malware* (25%), *DDoS* (20%) and *RATs* (17%). By comparing their findings with archival data from 2016, they found that there has been a 20% rise in the number of darknet listings that have the potential to harm the enterprise. By responding to ads and actively pretending to be buyers they were also able to get prices for targeted attacks (enterprises around 4 500 USD, individuals 2 000 USD) and espionage (1 000-15 000 USD). They never went through with any of the purchases, but prices are probably more realistic than the ones published within marketplaces.

#### V. CONCLUSION

There are different ways of looking at the underground market for cyber attacks, and we deem threat indicators based on sales to be more reliable than availability of items. This comes at a cost of a smaller sample size of markets, so we recommend considering both views in combination. The demand for *phone hacking* tools is prevalent, which is a natural consequence of our societies increasing use of phones for everyday digital activities. When comparing our result with past related studies, especially *stealers and grabbers* seem to be trending items. Such items were clearly present in most marketplaces and had a high number of sales. They are typically used for digital fraud and information theft, which indicates threat agents with a rational behavior and economic motivation. Bitcoin stealers are the most popular, and even though the price of individual items tends to be low (around 4 USD), the volume of sales suggests a decent revenue to the vendors.

#### REFERENCES

- [1] S. L. Pfleeger and D. D. Caputo, "Leveraging behavioral science to mitigate cyber security risk," *Computers & security*, vol. 31, no. 4, pp. 597–611, 2012.
- [2] R. Anderson, R. Böhme, R. Clayton, and T. Moore, "Security economics and the internal market," *Study commissioned by ENISA*, 2008.
- [3] R. Broadhurst, D. Lord, D. Maxim, H. Woodford-Smith, C. Johnston, H. W. Chung, S. Carroll, H. Trivedi, and B. Sabol, "Malware trends on 'darknet' crypto-markets: Research review," Australian National University Cybercrime Observatory and the Korean Institute of Criminology, Tech. Rep., 2018.
- [4] R. V. Kozinets, *Netnography*. Wiley Online Library, 2015.
- [5] Darknetlive. <https://darknetlive.com/>. Last accessed: 2019-10-03.
- [6] Thedarkweblinks. <https://www.thedarkweblinks.com/>. Last accessed: 2019-10-03.
- [7] Darknetstats. <https://www.darknetstats.com/>. Last accessed: 2019-10-03.
- [8] OASIS, "STIX Version 2.0. Part 1: STIX Core Concepts," OASIS Cyber Threat Intelligence (CTI) TC, Tech. Rep., July 2019.
- [9] F. Wehinger, "The dark net: Self-regulation dynamics of illegal online markets for identities and related services," in *2011 European Intelligence and Security Informatics Conference*, Sep. 2011, pp. 209–213.
- [10] Ken. (2018, June) Dream market: A hotbed of scammers. <https://darkwebnews.com/darkwebmarkets/dream-market/dream-market-a-hotbed-of-scammers/>. Last accessed: 2019-06-27.
- [11] Y. F. F. Bayoumy, P. H. Meland, and G. Sindre, "A netnographic study on the dark net ecosystem for ransomware," in *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*. IEEE, 2018, pp. 1–8.
- [12] L. Ablon, M. C. Libicki, and A. A. Golay, "Markets for cybercrime tools and stolen data: Hackers' bazaar," Rand Corporation, Tech. Rep., 2014.
- [13] K. Thomas, D. Huang, D. Wang, E. Bursztein, C. Grier, T. J. Holt, C. Kruegel, D. McCoy, S. Savage, and G. Vigna, "Framing dependencies introduced by underground commoditization," in *14th Workshop on the Economics of Information Security*, 2015.
- [14] R. Van Wegberg, S. Tajalizadehkhoob, K. Soska, U. Akyazi, C. H. Ganan, B. Klievink, N. Christin, and M. Van Eeten, "Plug and prey? measuring the commoditization of cybercrime via online anonymous markets," in *27th USENIX Security Symposium*, 2018, pp. 1009–1026.
- [15] K. Soska and N. Christin, "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem," in *24th USENIX Security Symposium*, 2015, pp. 33–48.
- [16] M. McGuire, "Behind the dark net black mirror: Threats against the enterprise," Bromium and University of Surrey, UK, Tech. Rep., June 2019.

**J: ‘The Ransomware-as-a-Service economy within the darknet’**

Included is the published material [26], following the Creative Commons Attribution 4.0 International (CC BY 4.0) licensing arrangement used by Elsevier.



**J**



Contents lists available at ScienceDirect

Computers &amp; Security

journal homepage: [www.elsevier.com/locate/cose](http://www.elsevier.com/locate/cose)

# The Ransomware-as-a-Service economy within the darknet

Per Håkon Meland<sup>a,b,\*</sup>, Yara Fared Fahmy Bayoumy<sup>a</sup>, Guttorm Sindre<sup>a</sup>

<sup>a</sup> Department of Computer Science, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway

<sup>b</sup> Department of Software Engineering, Safety and Security, SINTEF Digital, NO-7034 Trondheim, Norway

## ARTICLE INFO

### Article history:

Received 28 November 2019

Revised 30 January 2020

Accepted 18 February 2020

Available online 29 February 2020

### Keywords:

Ransomware

RaaS

Malware

Darknet

Marketplace

Netnography

## ABSTRACT

Ransomware is an epidemic that adversely affects the lives of both individuals and large companies, where criminals demand payments to release infected digital assets. In the wake of the ransomware success, Ransomware-as-a-Service (RaaS) has become a franchise offered through darknet marketplaces, allowing aspiring cybercriminals to take part in this dubious economy. We have studied contemporary darknet markets and forums over a period of two years using a netnographic research approach. Our findings show that RaaS currently seems like a modest threat relative to popular opinion. Compared to other types of illegal digital goods, there are rather few RaaS items offered for sale in darknet marketplaces, often with questionable authenticity. From our data we have created a value chain and descriptions of the actors involved in this economy.

© 2020 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY license. (<http://creativecommons.org/licenses/by/4.0/>)

## 1. Introduction

The darknet is an unregulated *Wild West* of the Internet, cyber crime's safe haven for communication and exchange of illegal goods and services. It is easily accessible, and with the help of anonymisation technology and modern-day digital currencies, a full-fledged economy takes place on a global scale right under the nose of impaired law enforcement agencies. An estimated USD 1 billion has been spent here during the first nine months of 2019 (Europol, 2019).

We have been especially interested in ransomware, which enables extortion of victims by taking control of their digital assets. On the darknet markets, *Ransomware-as-a-Service* (RaaS) is being offered as a franchise model that allows people without programming skills to become active attackers and take part in the ransomware economy. This is a way of democratising crime, giving ordinary people and smaller players an easier way into the criminal market (Jaishankar, 2008; Naylor, 2000), while reducing the risk of exposure for the ones on top of the value chain. For instance, a dissatisfied employee might decide to partner up with a RaaS developer to effectively infect an organisation from the inside and then splitting the profit.

### 1.1. Objective

In order to devise effective countermeasures against RaaS, it is helpful to understand the intricate relationships of people operating within the opaque darknet markets (Thomas et al., 2015). Currently, the relationships between organised crime and the Internet is under-investigated (Lavorgna, 2015). This research gap can be narrowed down by looking at the motivations and incentives of the people involved, and Waldrop (2016) suggests that this can be accomplished by embracing behavioural science and economics as part of the research. The research objective of our work has been to obtain a better understanding of the darknet market for RaaS as we have tried to address the following research questions:

1. How severe is the RaaS threat?
2. What are the value chains related to this market?

The answers to these questions are of significance when estimating the current impact of RaaS and the participating actors, and to guide further research both for academic and commercial purposes.

### 1.2. Scope

We have studied RaaS within popular contemporary darknet markets and forums over a period of two years (fall of 2017 to fall of 2019) using a netnographic research approach. Our observations have been complemented with historical data found in archives and published interviews with stakeholders involved in darknet operations. Our study has been limited to English-speaking spaces

\* Corresponding author at: Department of Software Engineering, Safety and Security, SINTEF Digital, Strindvegen 4, NO-7034 Trondheim, Norway.

E-mail address: [per.h.meland@sintef.no](mailto:per.h.meland@sintef.no) (P.H. Meland).

<https://doi.org/10.1016/j.cose.2020.101762>

0167-4048/© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license. (<http://creativecommons.org/licenses/by/4.0/>)

not residing behind walls requiring pay-for-access or other unethical contributions.

### 1.3. Outline

In [Section 2](#) we present background information about the environment in which we have conducted the study. [Section 3](#) gives an overview of related research that we have built our knowledge on. [Section 4](#) details our methodological approach and data sources, including the ethical issues we had to consider. [Section 5](#) summarises our most important results, which are discussed in the following [Section 6](#). Finally, [Section 7](#) concludes the paper.

## 2. Background information

### 2.1. The darknet and dark web

The term *darknet* is commonly associated with hidden networks on the Internet, and most prominently, *The Onion Router* (TOR), originally developed by the US Naval Research Laboratory to protect communication with agents stationed abroad but later made open to anyone who wants to anonymously interact with others. Another darknet example is the *Invisible Internet Project* (I2P), but it currently has fewer users and is thus considered less anonymous than TOR.

The collection of websites that reside on the secret space of the darknet is commonly referred to as the *dark web*. The dark web can also be thought of as a subset of the *deep web* (aka *invisible* or *hidden web*). What distinguishes any website in the deep web from what we refer to as the *surface web*, *lightnet*, or *clearnet*, is that it is not indexed, and therefore, cannot be found by the everyday search engines most people use. Though most of the deep web content is perfectly legitimate, the story is quite different when it comes to dark web. A study by [Moore and Rid \(2016\)](#) gave a conservative estimate that 57% of the TOR websites facilitated criminal activities related to drugs, arms, murder and child pornography.

### 2.2. Marketplaces and forums

Both within the surface web and darknets there are websites similar in structure to online shopping sites that facilitate the illegal transactions. These websites go by the name of *darknet markets*/*marketplaces*, *underground markets* or *cryptomarkets*. For the sake of simplicity, the rest of this paper will refer to them as darknet markets.

The pioneering Adamflowers/Farmer's Market started out as a surface web market in 2006 but transitioned to TOR in 2010. It had been selling illegal drugs to more than 34 countries before it was eventually shut down by law enforcement agencies in 2012 ([Vaas, 2012](#)). Learning from the mistakes of the Farmer's Market, Silk Road became the first darknet market that used cryptocurrency for payment in 2011. The business model of Silk Road was very successful, and its administrators were making a living off vendor fees and commissions. It was shut down by the FBI in 2013, but a multifold of markets emerged in its wake using similar models.

Sometimes darknet marketplaces are shut down for other reasons than law enforcement. Money stored in escrow has on several occasions been stolen from or by the administrators, so-called *exit scams*. The Sheep Marketplace is a well-known example, where one of the vendors exploited a site vulnerability and took off with 54 000 bitcoins in 2013, while the administrator shut down the site and stole 40 000 bitcoins for himself in 2015 ([DIVIDEDBY0, 2017](#)).

Most darknet markets are accompanied with a discussion forum. Such forums help the users tackle uncertainties related to the quality of the offered goods and services ([Yip et al., 2013](#)). For

instance, vendor review is a common discussion topic. This helps identify potential *scammers*, i.e., vendors that actively manipulate their own product reviews.

### 2.3. Ransomware and Ransomware-as-a-Service

[Gallo and Liska \(2016\)](#) define *ransomware* as "a blanket term used to describe a class of malware that is used to digitally extort victims into payment of a specific fee". Typically, malicious code makes specific files or a whole system unavailable to the victim through encryption or change of usage rights. After a limited time, the ransom fee must be payed, or the damage becomes permanent. In most cases (65%) ([Hernandez-Castro et al., 2017](#)), the system is recovered after the ransom has been payed.

The first ransomware, known as AIDS, was observed in the wild already in 1989, spreading through the exchange of floppy disks ([O'Kane et al., 2018](#)). In the years to follow, ransomware was not a serious threat. Studies by [O'Gorman and McDonald \(2012\)](#) and [Kharraz et al. \(2015\)](#) have shown that the number of ransomware families was quite low for more than two decades, especially the ones with sophisticated destructive capabilities. However, this all changed with the introduction of stronger encryption schemes in the ransomware code and especially the availability of cryptocurrency as a payment method difficult to track by law enforcement ([Young and Yung, 2017](#)). Ransomware has been recognized as one of the fastest growing cybercrimes in recent history ([Grobman and Cerra, 2016](#)), and even though the overall number of infections started to decline in 2018, the current trend is that businesses are becoming the primary targets, whereas regular citizens are to a lesser extent being hit ([Symantec, 2019](#)).

In the wake of the ransomware success, *ransomware-as-a-service* (RaaS) has become an entry point for criminals with little programming skills to participate and earn money from ransomware ([O'Kane et al., 2018](#)). Contacting ransomware service providers using darknet markets, the criminals can cheaply obtain tailor-made ransomware ready to be used on their prospective victims. In addition to the creation fee, the service providers may take a 20–30% cut of the ransom as well. RaaS can have different formats, such as source code that the buyer compiles himself, pre-compiled binaries or an interface where the buyer inputs information about the victims. This collaborative strategy is a way of achieving a faster rate of infections with a lower risk of getting caught.

## 3. Related research

### 3.1. Marketplace and forum research

The vast body of research on darknet markets is related to illegal drugs, while there is limited literature focusing solely on ransomware markets. However, if we glance towards the broader category of digital goods and services, we find many studies that are of relevance to ransomware. [Ablon et al. \(2014\)](#) published a book describing structures, types of participants, products of open and closed black markets. Though their focus was mostly on botnets and zero-day vulnerabilities, they also show the price development for exploit kits and the evolution of markets over time. The year after, [Thomas et al. \(2015\)](#) surveyed existing research in order to systematize the community's understanding of the underground economy and develop a taxonomy of profit and support centres for reasoning about the flow of capital. [Broadhurst et al. \(2018\)](#) wrote a research review of malware trends on darknet markets. In their own six-month study (Sep 17 - Feb 18), they were able to observe increasing interaction between cybercriminals and state or quasi-state cybersecurity actors. Their analysis of the Dream market product listing in this period showed that ransomware only

constituted 0.73% of the offered goods, while compromised accounts and credit cards represented 72% of the listed products.

Van Wegberg et al. (2018) carried out a six-year longitudinal study tracking the evolution of commoditization on eight marketplaces, spanning from Silk Road to Alphabay. Within the malware category, the ransomware clusters around the Stampado and Philadelphia stood out as the most prominent. However, they also claim that there has been limited growth due to bottlenecks in outsourcing critical parts of the criminal value chain. This can be seen in relation to the exploratory darknet study by Cusack and Ward (2018). Based on observations from the business processes and technologies associated with ransomware, their opinion is that over time, erosion of trust will render the ransomware crime model economically infeasible.

### 3.2. Stakeholders, roles and value chains

The stakeholders involved in the underground economy have different responsibilities and expose themselves to different types of risks. Several research papers have modelled value chains that illustrate the roles involved and the direction of communication and responsibility. Zhuge et al. (2009) have modelled the underground economy in China, with an emphasis on online games. They defined several roles, including *virus writers*, *website masters/crackers*, *envelope (account) stealers*, *virtual asset stealers and sellers and players* (buyers). Yip (2010) compared the Chinese cybercrime underground with the West and added other types of roles for faux website design. In another stakeholder classification, Cárdenas et al. (2009) identified the *malware distributors* role. O'Kane et al. (2018) have described *mixers* and *tumblers* involved in the money laundering services. A report by the security company Carbon Black (2017) defined three core economic tiers for the ransomware supply chain; *author*, *RaaS* and *distributor*.

Yip et al. (2013) examined the structure of organised cybercrime and sources of uncertainty given the masked identities of the traders and presence of undercover agents. Rossy and Décarv-Héty (2017) further examined trust issues as vendors often face the threat of identity theft by people who want to take advantage of their established reputation. Holt et al. (2012) identified network structures for information sharing amongst *malware writers* and other members of the community. della Torre (2018) analysed the strategic dynamics of vendors in the darknet markets, discovering that the fittest and richest vendors focus on a limited subset of products (3–5) with little updates. Kwon and Shakarian (2018) studied information sharing between actors during take-downs, finding examples of both collaboration for alternative economic routes and distrustful communication during such events.

For a thorough overview of the contemporary cybercrime ecosystem and its developments, we refer to Broadhead (2018).

### 3.3. Economics of ransomware

There have been many papers that analyse the economics of ransomware as seen from the offender's and victim's point of view. Economic incentives from developing and distributing ransomware are high, simply because the revenue is high, whereas the costs of resources and probability of apprehension are low. Hernandez-Castro et al. (2017) put forth an economic model based on the victim's willingness to pay. Here, the amount for a single ransomware variant can either be a fixed price for all victims, or fluctuating based on a set of factors (*price discrimination*). Laszka et al. (2017) proposed a game-theoretic model of the ransomware ecosystem, including backup and recovery investments, and incentives to pay the ransom. Lee and Lee (2017) observed that the cost of acquiring ransomware was determined by complexity of the vulnerability the malware is exploiting.

Aurangzeb et al. (2017) have done a literature survey on ransomware families including their payment methods.

Another category of studies has tried to *follow the money*, analysing the cryptocurrency transaction logs associated with ransomware. For instance, Huang et al. (2018) do this from the time victims acquire bitcoins to pay the ransom and through to the time ransomware operators cash them out. Paquet-Clouston et al. (2019) have a similar approach. They found that this market is highly skewed with a low number of players and that the total amount of ransom is relatively low compared to the hype surrounding the issue. The analysis by Anderson et al. (2018) of ransom payments on the blockchain indicated that substantial ransom sums may have been mixed in and obfuscated with drug transactions. Conti et al. (2018) have conducted a longitudinal study on twenty ransoms and how they have impacted the economy of bitcoin payments.

Within academic publications, there has been less research focusing on the economy of ransomware-as-a-service. However, a few security companies have published reports on this franchise model. For instance, Check Point and IntSight (2016) disclose the business operation of the Cerber RaaS from end-to-end, and Carbon Black (2017) describe how novice criminals are included to minimize the risk of the ransomware authors.

## 4. Methodological approach

*Netnography* is a research approach centred on the study of *online traces*, which are various types of data people make available online to anonymous or networked others (Kozinets, 2019). In this sense, they also represent social information on which research can be done. We answer to Kozinets' four defining elements by having a *cultural focus* on ransomware trade, *social media data* that primarily stem from darknet marketplaces and forums, an *immersive engagement* through actively learning and reflecting on the focal phenomenon by members of the research team, and finally a *praxis* that follows particular netnographic research procedures.

As an initial *movement*, we decided upon the ethical concerns related to this research. Online traces such as archived data are publicly available and should technically be regarded as published open content. However, the personal identities of the people involved are secret, and they operate behind pseudonyms. Connecting data and giving them unwanted exposure could lead to retributive actions, e.g., towards the researchers or affiliated organisations. To reduce such risks, we decided to avoid direct interaction with subjects creating or selling ransomware. This is stressed by Martin and Christin (2016) for two main reasons. Firstly, the research after publication will not be pertinent to any proof for prosecution against any individual. Secondly, there will be no need to ask for permissions or consent. The pseudonyms we recorded in our field notes are either altered or not included in this paper, hence no data linked to the user's identity or personal background are exposed. To avoid supporting illegal activities, we have not purchased anything. Finally, we have not tried to deceive, intimidate or confuse people within this research space.

Our study spanned over two years with four phases of data collection further described below.

### 4.1. Phase 1: pre-study

This initial phase was a pre-study of contemporary darknet markets and forums performed during the fall of 2017. Following the recommendations of Kozinets et al. (2014) we found it best to start the investigation with a small number of sites to gain a cultural sense of "what is going on" in that particular social space. Our sample was selected using DNStats (2019), which at that time

offered links to the most popular darknet websites along with uptime and availability. We chose the *Dream* and *WallStreet* markets, being the two most prominent markets dealing with ransomware, and the discussion forum *Intel Exchange*, which was the only open market that allowed members to promote ransomware services (aka *vending*). By searching for “ransom” and manual inspection we collected RaaS item price listings and descriptions in our field notes, as well as vendor profiles and ratings/reviews/comments from buyers. Within forums we also used the search keyword “ransom” and recorded relevant discussions, e.g., related to the process of buying and partner search for development or distribution.

#### 4.2. Phase 2: expansion

We expanded our research sample in the spring of 2018, covering additional contemporary sites, historical data and published interviews with stakeholders. These were selected using *DNStats*, (2019), *Reddit* (2019), *DeepDotWeb* (2017) and *Darknet Markets* (DNetX, 2019). Prior to 22nd of March 2018, Reddit offered several subreddits with posts concerning darknet markets and activity, but these were all banned to shut out illegal activities. DeepDotWeb provided news and an overview of the top darknet markets and forums based on ratings and uptime status. Darknet Markets provided news and a directory listing of active and dead sites.

In addition to the previous marketplaces from phase 1, we chose to include the *Berlusconi* market, which was growing quickly at that time. We identified historical archives by *Branwen et al.* (2015), containing scraped data of 89 different marketplaces and 37 forums between 2013–2015, *McKenna and Goode's archive* (2017) of *Alphabay* between 2016–2017, and *Lewis' (2017)* item listings and buyer feedback from the *Hansa* and *Valhalla* markets from October and December 2016. Additional forums were the top ranked *OnionLand*, *HUB*, and *HiddenAnswers*. Both *Onionland* and *HUB* were taken down in the beginning of 2018.

We gained insight into the thoughts and opinions of darknet community stakeholders by studying interviews published on DeepDotWeb, covering marketplace administrators (*TheRealDeal*, *Alphabay* and *German Plaza*), a marketplace platform developer, a forum moderator, a forum vendor, a money launderer, and a ransomware developer.

#### 4.3. Phase 3: iteration

During the Winter of 2018/2019 we revisited the contemporary marketplaces and forums to capture the latest trends and developments with respect to RaaS. We included the *Tochka* (aka *Point*) market due to its then high ranking at Darknet Markets and DeepDotWeb, the *Empire* market, which had emerged in February 2018 to become one of the fastest growing markets, and the *Dread* forum, which had become a popular discussion site on the darknet after the subreddit crackdowns. For our stakeholder analysis, we included additional published interviews with the administrators of *Valhalla*, *Outlaw*, *Minerva*, *Oasis* and *Tochka* found on DeepDotWeb, as well as one with the *Empire* market administrator found in a *Dark Web News* article by *C.M.* (2018).

#### 4.4. Phase 4: a new line-up

By Fall 2019, several of our previous data sources were debunked or shut down (*Dream*, *WallStreet*, *IntelExchange*). As *DeepDotWeb* had also been seized by law enforcement, the identification of marketplaces relied on *DarknetLive* (2019), which we found to have the most up-to-date index of marketplace links, supplemented by *TheDarkWebLinks* (2019) and *DNStats*. From the living marketplaces we found RaaS in the following sample: *Apollon*, *Berlusconi*, *Darkbay*, *Empire*, *Grey* and *Samsara* (successor of *Dream*).

*Berlusconi* went offline around September 22nd, right after we had completed our observations, possibly due to an exit scam or takedown. We excluded *Tochka* since there were no RaaS items there anymore.

## 5. Results

We have integrated the collected data from each phase and made an incarnation showing phenomena related to vendor resilience despite of marketplace takedowns, that there is a strong decline in the availability of RaaS items, that there is a high risk of buying fraudulent items, what kind of buyers/distributors the vendors are targeting, and finally, a larger picture of the RaaS economy and its actors.

### 5.1. Vendor resilience

Our first study phase started right after the takedown of the dominant darknet markets *Alphabay* and *Hansa* as a part of Operation Bayonet (Europol, 2017). This led to a rapid growth of the *Dream* userbase also observed by *Van Wegberg et al.* (2017), both when it came to vendors and buyers. We believe that one of the reasons that *Dream* succeeded in taking this business was its relatively high uptime and performance compared to its competitors at the time. Another reason could be related to a rapid establishment of trust between the actors. We observed that *Dream* had a specific feature that allowed vendors to present their previous rating from *Alphabay* and *Hansa* on their profile page. This let them maintain their existing reputation and buyers could base their trust on trade ratings from dead markets. This phenomenon reappeared in phase 4 after the death of *Dream*, as *Empire* allowed vendors to display their sales stats from *Dream*. This is an example of resilience in a volatile environment where people are anonymous, and trust is a great market advantage.

### 5.2. Market size perspectives

The most popular goods sold on open darknet markets are drugs. Where available, RaaS items are usually found under the *Digital Goods* or *Services* categories, but RaaS is rare in these inventories. The most popular digital goods or service is *carding* or *credit card fraud*. Fig. 1 shows an overview of items offered on *Dream*, the largest market in phase 2 and 3 of our study, comparing ransomware to *digital goods* and to *carding*.

Though the number of total items had increased about 38% between 2018 and 2019, the number of RaaS items declined 22%. In phase 3, RaaS items constituted about 0.15% of the total items available at *Dream*. We could not extend this trend analysis to phase 4 as *Dream* died before that, however the successor *Samsara* contained merely 3 RaaS items. In fact, the total number of ransomware items across the six remaining markets were now only 69. 65 of these items were sold from the markets *Apollon*, *Berlusconi*, *Empire* and *Grey*, which were the only ones that also stated the number of successful sales per item. Only 28 items had any sales at all, and the total number of successful sales from these were 359, constituting a total sales profit of approximately USD 2 202 based on the listed price per item.

### 5.3. No honour among thieves

The authenticity of RaaS items sold on the darknet markets was questioned throughout our research and we found several indications of scam. Firstly, most of the renowned RaaS vendors had gained their high rating from credit card gift cards or drug related sales in the past, and not because of RaaS. Secondly, the descriptive RaaS information tended to be copied from other RaaS items.

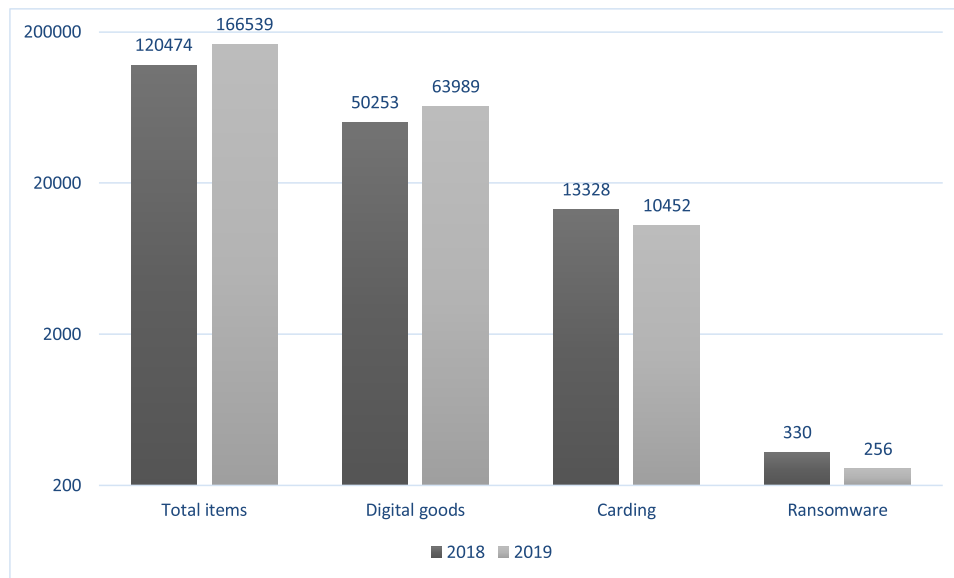


Fig. 1. Inventory excerpt from the Dream market shown with a logarithmic Y-axis scale.

Thirdly, a lot of the data in the feedback fields, which include obscured aliases and star ratings, seemed to be artificially created since they were identical and registered at the same time. Buyers using the free text fields tended to give negative feedback. Such observations lead us to believe that most of the RaaS items sold on the darknet markets are frauds, where the buyers either get rubbish or ransomware that redirects the whole payment somewhere else than the buyer's wallet. For instance, one of the most trusted RaaS vendors we found on *WallStreet* received this feedback comment:

"...these files are all open source files found for free at github, and are old"

The fraud assumption was further supported by a question posted on the *OnionLand* forum, where a user questioned the validity of services offered by software dealers on the marketplaces:

"Is there anyone or any vendor/market out there that isn't a scam)? I mean, seriously!!! I'm beginning to think this whole Darknet is just an urban legend!!"

The moderator of the forum responded as follows:

"The public space is supposed to be filled with scams and stupid products, because you don't have to prove your worth to get into the public sphere. The only way to experience the inner workings is to be able to convince others that you should be allowed into invite-only spheres as mentioned."

Gaining access to such walled spaces can be a challenge if you do not already know someone on the inside. For instance, one of the most popular walled forums, named *Hell*, requested a payment of 0.01 Bitcoin or a trusted referral in order to get access. Additionally, users would need to prove their worth for the community. Upon an inspection of the *Hell* bitcoin wallet we could not see a substantial amount of transactions, which either means that there few members or they are invited by acquaintances.

#### 5.4. RaaS target market

In order to gain an understanding about the type of customers the vendors were targeting, we looked more closely at our gathered RaaS item descriptions. A common piece of information is the recommended level of technical expertise a buyer should have. During phase 1, we analysed the 20 items that provided such descriptions and found out that most of them (65%) targeted experts, while novice users should be able to use the other portion (35%). Moreover, popular items tended to include links to detailed guides and tutorial videos with step-by-step instructions on how to distribute and activate the ransomware, claim the ransom (or even give mercy to the victim).

We also analysed the anonymous social interactions that took place on the forums. During phase 2, we categorized the frequency of the RaaS topics that we found on *HiddenAnswers*, which was the oldest forum and had the highest number of posts concerning RaaS compared to *OnionLand* and *HUB*. Based on 79 posts in English, we created 8 different groups of Q&A as shown in [Fig. 2](#).

The majority of these posts were about ransomware acquisition or development, indicating that this forum was dominated by non-experts. This was to be expected since experienced developers would rather stick to walled forums or IRC-channels.

#### 5.5. Value chain

Based on marketplace observations, forums posts, available interviews and literature we have created a simplified map of the value chain related to RaaS as depicted with blue arrows in [Fig. 3](#). RaaS items follow the red arrows until they become ransomware infections at victims. The green arrow indicates a close coupling between marketplaces and forums. The stakeholders are briefly described in [Table 1](#), where we have also tried to classify them according to the risk categories *high*, *medium*, *low* based on how likely it is that they will be exposed and possible consequences. Note that there are law enforcement agencies, security companies, researchers and neutral darknet bystanders entangled in this anonymised online community as well, but they are not directly involved in the economy.

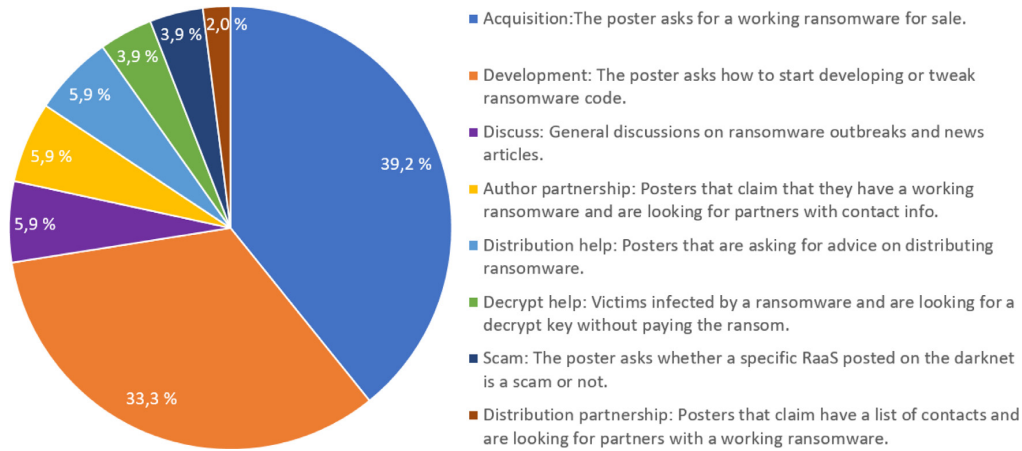


Fig. 2. Question categories related to ransomware in the Hidden Answers forum.

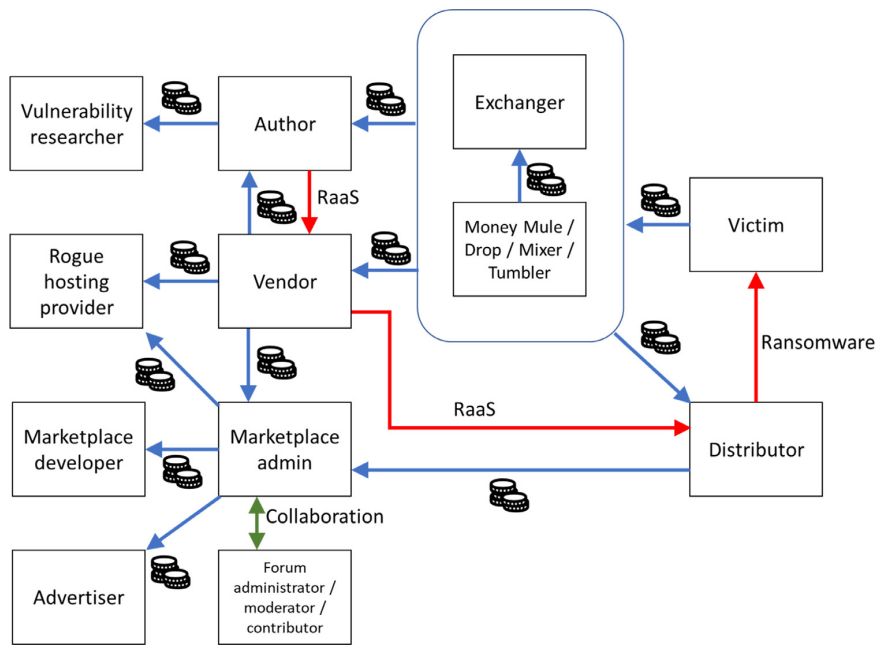


Fig. 3. Value chain for the RaaS economy.

### 6. Discussion

Netnography studies are useful for getting a better understanding of the activities taking place on the darknet. In our case, we narrowed the scope down to phenomena related to RaaS, but RaaS is often tied to other types of activities as well, such as a plethora of different infection methods and money laundering schemes. The size, unstructured nature and instability of our data sources have been a challenge in the data collection and analysis. However, this instability is a reality that the darknet community must deal with as well. On the surface web, we are all used to search engine functionality when looking for information, but on the darknet, links to markets, forums and websites are a commodity listed in market inventories. In addition, access to walled sites is seldom granted for free. When facing such research barriers, it is important to acknowledge that we will never get a complete picture of the social interactions and economy within this somewhat obfuscated world.

However, we argue that through our research approach, we have been able to find clear indications, trends and examples of phenomena that contribute to the general knowledge of RaaS activities on the darknet.

Darknet markets, though constantly hunted by law enforcement agencies, have proven themselves to be quite resilient. In spite of numerous takedowns of high-profiled markets, vendors persist and quickly move on to other markets, using their PGP key to preserve their reputation. This is in accordance with [Everton's \(2008\)](#) general finding that "covert and illegal (i.e., dark) networks are quick to adapt to changing environmental pressures".

A trend that [Europol \(2018\)](#) has documented, is an increasing number of smaller vendor shops and secondary markets catering to specific languages or nationalities. Smaller vendor shops are more difficult to come by, and were not within the scope of our study, so we cannot say if this is also the case for RaaS, but we observed that some of the most well-known RaaS items are provided

**Table 1**  
Actor descriptions and risk categories.

Actor	Description
Vulnerability researcher	Vulnerability researchers (Cárdenas et al., 2009) discover and sell information about zero-day vulnerabilities to others who can write the exploit code. They have high expertise in hardware and software, and a forum member mentioned that many of them were <i>sysadmins</i> in respected companies. Risk category: Low, little exposure and minor consequences of getting caught.
Author	Authors are professional developers that create the malware that takes advantage of vulnerabilities, some of which are purchased from vulnerability researchers. There are authors offering services for signing ransomware with stolen code certificates to make the payload look legit (Abrams, 2016). As pointed out by Yip (2010), there can be fierce competition between malware authors. Risk category: Low, authors seldom expose themselves on the darknet and rather outsource the risk taking to others while harvesting a significant portion of the ransom amount.
Vendor	Vendors do marketing and sales on marketplaces or on their own private website. Vendors can be authors, but the majority of darknet vendors have little programming knowledge and sell a wide range of products that are not necessarily digital goods. Some vendors offer technical support. Risk category: Medium, can be compared to weapons dealers that facilitate crime, but do not directly take part in the offensive action. Highly exposed on the darknet.
Distributor	The distributors buy or get hold of RaaS and infect the devices of victims. Distributors can be observed on the darknet. They share experiences and feedback on ransomware purchases. Some distributors search for partnerships involving malware developers on forums and offer vulnerability information of their target system. As shown in earlier studies (Bayoumy et al., 2018), two levels of malware distributors can be defined; <i>novice</i> and <i>experienced</i> . Risk category: High, severe consequences if they get caught (depending on different legal jurisdictions).
Victim	Victims suffer from ransomware infections and may lose their data or pay the ransom (or both). They may need the help of an exchanger to obtain the ransom amount in cryptocurrency. Risk category: High, the main source of income for all other parties.
Marketplace admin	Provides a market platform that vendors and distributors can use for trade. Should be a trusted third party that governs the money transaction. There have been several examples of administrators running off with the money (exit scams). Risk category: High, law enforcement agencies put a lot of effort in shutting down these services. High penalty when caught. Also, other marketplaces may try to get rid of competition.
Marketplace developer	Person with technical expertise that develops the marketplace platforms for the administrators. Requires a high security competence. Risk category: Low, creating marketplace infrastructure is probably not a crime in itself.
Advertiser	Marketplace affiliate that posts darknet links on the surface web and receives kickback money when there are successful transactions originating from these. Example DeepDotWeb. Risk category: Medium, high penalty when getting caught, but this does not happen often.
Forum admin/moderator / contributor	People responsible for managing the forum contents and membership access. Usually have a close relationship with the administrator of one or more marketplaces. Risk category: Medium, forums are targeted by law enforcement agencies just as marketplaces, but probably a lesser penalty if they get caught.
Rogue hosting provider	Provide website hosting services on the darknet that reduces the risk of getting caught (Cárdenas et al., 2009). Risk category: Low, difficult to prove that they are responsible for the website contents.
Money Mule / Drop / Mixer / Tumbler	Transactions received from victims are transferred through an intermediary, either a professional money launderer or someone who unknowingly forwards the money. Modern ransomware actors tend to immediately launder their gains through well-known bitcoin laundering operations, who take a fee (around 2.5%) for their services (Hernandez-Castro et al., 2017). A marketplace administrator (Empire) operating with Monero has said that tumblers are not needed due to the anonymity features of that cryptocurrency. Risk category: High, unknowing mules can be traced and prosecuted even though they are innocent. New investigation techniques can better track cryptocurrency transactions.
Exchanger	Exchangers own verified accounts and use their immunity to offer currency exchange services to cybercriminals. Risk category: Medium, as their actions can be investigated by authorities or financial institutions.

from dedicated sites, and that several vendors were unhappy with the commission and vendor fees of the larger markets. Contrary to the findings by della Torre (2018), showing that the “best” vendors focused on few products, we have observed in the case of RaaS that the vendors deal with a large variety of products in several different categories.

What we can say with a large degree of certainty, is that RaaS constitutes a relatively small portion of the inventory for the major darknet markets. There have been reports from security companies that seem to be inaccurate or biased. For instance, one report from 2017 (CarbonBlack, 2017) claimed that there were 45,000 current listings, and that the sales of ransomware in the darknet increased by 2500% from 2016 to 2017. These estimates were based on measurements from a small sample that were extrapolated based on the assumed size of the darknet. Our latest observations showed that there were merely 69 ransomware related items for sale in the dominating markets after a strongly decreasing trend from 2018 to 2019. In addition, we saw indications many of these items were duplicates and frauds, leading us to believe that the real availability of RaaS seems exaggerated. Indeed, our assumptions regarding RaaS fraud support the findings of Wehinger (2011) and Cusack and Ward (2018) related to lack of trust and amount of fraud on the darknet. Compared to RaaS, carding services are more

prevalent on the darknet, arguably since they require less technical skills and a different economic model where the buyers ask the vendor to deduct the price of the service from the total amount of money in the card instead of buying it in cryptocurrency. Unlike RaaS, the reviews on carding services are considered more authentic since they are more expressive and greater in number.

Open darknet forums allow members to share knowledge and eventually improve their skills and create partnerships with others. Getting into an invite-only forum requires a history with darknet activity, and this can be achieved through prolonged discussions on the open forums. This is in line with the *apprentice work ethics* phenomenon as reported by Mann and Sutton (1998). Holt et al. (2012) have presented a sociograph for connectivity and centrality of darknet members showing that low-skilled hackers have a lot less connections than the highly skilled, who are very much aware of their peers. This was evident in the forum activities we were able to observe as well. Those who openly want to acquire information are indeed low-skilled and publicly post on forums, putting them at the edge of the sociography, whereas the highly skilled are usually active in invite-only forums or have been assigned to be the moderator of the forum. This is in accordance with the two-tier model of Herley and Florêncio (2010); an open



tier for inexperienced users and a more closed tier for experienced criminals.

Our study has been limited to English-speaking markets and forums. These are known to be more concerned with drug related items and carding services compared to, e.g., Russian sites. Leah (2019) has given an historical overview of Russian-specific darknet markets and forums that complements our study. According to her, Russian criminals are notorious for selling malicious software, while Russian authorities have "historically turned a blind eye to online crimes". The most well-known darknet marketplace and forum, RAMP, was reportedly taken down in July 2017, but the vendors successfully moved to other key marketplaces. She also reports that digital goods markets such as MEGA and Hydra require direct communication between buyer and vendor before the transaction takes place. This mechanism is a way of increasing trust between the actors, and it will be interesting to see if the Western markets will implement the same strategy.

## 7. Conclusion and further work

Based on our own field notes from studying the darknet over two years and additional archival data going further back, the answer to our first research question is that the RaaS threat currently seems more modest than indicated in the media and reports from security companies. There are now relatively few RaaS items offered for sale in the most popular darknet marketplaces, and the number of successful sales does not indicate a large economy. Moreover, the authenticity of many items was questionable. In a virtual economy where people are anonymous and real trust is hard to come by, there are plenty of opportunists trying to make money of naïve cybercriminals. Retribution is difficult, and reporting RaaS fraud to the police is not viable for several reasons. There are professional RaaS vendors that ask for a share of the ransom revenue instead of an investment up front. They tend to host their merchandise in privately-owned websites, but these are difficult to find due to the limited search capabilities on the darknet, and the fact that advertisements are banned from most of the forums. However, it is important to remember that ransomware prevails as a serious threat when committed by experienced cybercriminals, and the forums may be considered a recruitment ground for their organisations. The value chain we have outlined to address our second research question can be useful when trying to break the underground economy behind ransomware and subsequently mitigate this cyber threat.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## CRediT authorship contribution statement

**Per Håkon Meland:** Conceptualization, Investigation, Writing - original draft, Visualization, Resources, Data curation. **Yara Fareed Fahmy Bayoumy:** Conceptualization, Methodology, Investigation, Data curation. **Guttorm Sindre:** Supervision, Writing - review & editing.

## References

Ablon, L., Libicki, M.C., Golay, A.A., 2014. *Markets for Cybercrime Tools and Stolen Data: Hackers' Bazaar*. Rand Corporation.  
 Abrams, L. (2016, February 23). CTB-Locker for websites: reinventing an old ransomware. Retrieved from <https://www.bleepingcomputer.com/news/security/ctb-locker-for-websites-reinventing-an-old-ransomware/>.  
 Anderson, R.J., Shumailov, I., Ahmed, M., Rietmann, A., 2018. Bitcoin redux. Paper presented at the Workshop on the Economics of Information Security (WEIS).

Aurangzeb, S., Aleem, M., Iqbal, M.A., Islam, M.A. Security, 2017. Ransomware: a survey and trends. *J. Inf. Assur.* 6 (2).  
 Bayoumy, Y., Meland, P.H., Sindre, G., 2018. A netnographic study on the dark net ecosystem for ransomware. Paper presented at the International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA).  
 Branwen, G., Christin, N., Décary-Héту, D., Andersen, R.M., StExo, Presidente, E., ..., Goode, S. (2015). Dark net market archives, 2013-2015. Retrieved from <https://www.gwern.net/DNM-archives>.  
 Broadhead, S., 2018. The contemporary cybercrime ecosystem: a multi-disciplinary overview of the state of affairs and developments. *Comput. Law Secur. Rev.* 34 (6), 1180-1196.  
 Broadhurst, R., Lord, D., Maxim, D., Woodford-Smith, H., Johnston, C., Chung, H.W., ..., Sabol, B. (2018). Malware trends on 'Darknet'Crypto-Markets: research review. Available at SSRN 3226758.  
 C.M. (2018). Interview: empire market admin talks DNM community, security, Crypto & Plans for Future. Retrieved from <https://darkwebnews.com/darkwebmarkets/empire-market/empire-market-admin-interview/>.  
 CarbonBlack. (2017). The ransomware economy: how and why the dark web marketplace for ransomware is growing at a rate of more than 2,500% per year. Retrieved from <https://www.carbonblack.com/company/news/press-releases/dark-web-ransomware-economy-growing-annual-rate-2500-carbon-black-research-finds/>.  
 Cárdenas, A., Radosavac, S., Grossklags, J., Chuang, J., Hoofnagle, C., 2009. An economic map of cybercrime. Paper presented at the Telecommunications Policy Research Conference (TPRC).  
 CheckPoint. (2016). CerberRing: an in-depth exposé on cerber Ransomware-as-a-Service. Retrieved from <https://blog.checkpoint.com/2016/08/16/cerberring/>.  
 Conti, M., Gangwal, A., Ruj, S., 2018. On the economic significance of ransomware campaigns: a Bitcoin transactions perspective. *Computers & Security*, 79 162-189.  
 Cusack, B., Ward, G., 2018. Points of failure in the ransomware electronic business model. Paper presented at the Twenty-fourth Americas Conference on Information Systems.  
 DarknetLive. (2019). Darknet markets. Retrieved from <https://darknetlive.com/>.  
 DeepDotWeb. (2017). DeepDotWeb. Retrieved from <https://www.deepdotweb.com/>.  
 della Torre, G.G., 2018. Business Strategies in Darknet Marketplaces: An attempt to Model Competition in the Framework of Economic Complexity. Politecnico di Torino. Retrieved from <https://webthesis.biblio.polito.it/9063/>.  
 DIVIDEDBYO. (2017). Sheep marketplace owner indicted and face years in prison. Retrieved from <https://www.deepdotweb.com/2017/04/21/sheep-marketplace-owner-indicted-face-years-prison/>.  
 DNetX. (2019). Darknet Markets. Retrieved from <https://www.darknetmarkets.com/>.  
 DNStats. (2019). Dark Market Tracker. Retrieved from <https://dnstats.net/>.  
 Europol. (2017, 20 July). Massive blow to criminal dark web activities after globally coordinated operation. Retrieved from <https://www.europol.europa.eu/newsroom/news/massive-blow-to-criminal-dark-web-activities-after-globally-coordinated-operation>.  
 Europol. (2018). Internet Organised Crime Threat Assessment (IOCTA) 2018. Retrieved from <https://www.europol.europa.eu/activities-services/main-reports/internet-organised-crime-threat-assessment-iocta-2018>.  
 Europol. (2019). Internet Organised Crime Threat Assessment (IOCTA) 2019. Retrieved from <https://www.europol.europa.eu/activities-services/main-reports/internet-organised-crime-threat-assessment-iocta-2019>.  
 Everton, S.S., 2008. Tracking, Destabilizing and Disrupting Dark Networks with Social Networks Analysis. The NPS Institutional Archive DSpace Repository Retrieved from Calhoun <https://calhoun.nps.edu/handle/10945/34415>.  
 Grobman, S., Cerra, A., 2016. *The Second Economy: The Race for Trust, Treasure and Time in the Cybersecurity War*. Apress.  
 Herley, C., Florêncio, D., 2010. Nobody sells gold for the price of silver: dishonesty, uncertainty and the underground economy. In: *Economics of Information Security and Privacy*. Springer, pp. 33-53.  
 Hernandez-Castro, J., Cartwright, E., & Stepanova, A. (2017). Economic analysis of ransomware. arXiv:1703.06660v1.  
 Holt, T.J., Strumsky, D., Smirnova, O., Kilger, M., 2012. Examining the social networks of malware writers and hackers. *Int. J. Cyber Criminol.* 6, 891.  
 Huang, D.Y., Aliapoulos, M.M., Li, V.G., Invernizzi, L., Bursztein, E., McRoberts, K., ... McCoy, D., 2018. Tracking ransomware end-to-end. Paper presented at the 2018 IEEE Symposium on Security and Privacy (SP).  
 Jaishankar, K., 2008. Space transition theory of cyber crimes. In: Schmallegger, F., Pittaro, M. (Eds.), *Crimes of the Internet*. Pearson, pp. 283-301.  
 Kharraz, A., Robertson, W., Balzarotti, D., Bilge, L., Kirda, E., 2015. Cutting the gordian knot: a look under the hood of ransomware attacks. Paper presented at the International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment.  
 Kozinets, R.V., 2019. *Netnography: The Essential Guide to Qualitative Social Media Research*. SAGE Publications Limited.  
 Kozinets, R.V., Dolbec, P.-Y., Earley, A., 2014. Netnographic analysis: understanding culture through social media data. In: *The SAGE Handbook of Qualitative Data Analysis*. SAGE, pp. 262-276.  
 Kwon, K.H., Shakarian, J., 2018. Black-Hat Hackers' crisis information processing in the Darknet: a case study of cyber underground market shutdowns. In: *Networks, Hacking, and Media-CITA MS@ 30: Now and Then and Tomorrow*. Emerald Publishing Limited, pp. 113-135.  
 Laszka, A., Farhang, S., Grossklags, J., 2017. On the economics of ransomware. Paper presented at the International Conference on Decision and Game Theory for Security.

- Lavorgna, A., 2015. Organised crime goes online: realities and challenges. *J. Money Launder. Control* 18 (2), 153–168.
- Leah, M. (2019). Russians on the darknet part II: marketplaces & forums. Retrieved from <https://www.darkowl.com/blog/2019/russians-on-the-darknet-marketplaces-amp-forums>.
- Lee, J., Lee, K., 2017. Spillover effect of ransomware: economic analysis of web vulnerability market. *Res. Brief. Inf. Commun. Technol. Evolut. (ReBICTE)* 3.
- Lewis, S.J. (2017). Dark web data dumps. Github Repository. Retrieved from <https://polecats.mascherari.press/onionscan/dark-web-data-dumps>.
- Liska, A., Gallo, T., 2016. Ransomware: Defending Against Digital Extortion. O'Reilly Media, Inc..
- Mann, D., Sutton, M., 1998. »NETCRIME: more change in the organization of thieving. *Br J Criminol* 38 (2), 201–229.
- Martin, J., Christin, N., 2016. Ethics in cryptomarket research. *Int. J. Drug Policy* 35, 84–91.
- McKenna, M., & Goode, S. (2017). Alphabay crawl 20170128. Retrieved from <https://www.dropbox.com/s/0w74dz4c83tzhar/20170128-alphabay.tar.xz>.
- Moore, D., Rid, T., 2016. Cryptopolitik and the darknet. *Survival (Lond)* 58 (1), 7–38.
- Naylor, R.T., 2000. Expert Panel on Emerging Crimes: Hosted by the Department of Justice, Canada Retrieved from [https://www.justice.gc.ca/eng/rp-pr/csj-sjc/crime/rr03\\_20/rr03\\_20.pdf](https://www.justice.gc.ca/eng/rp-pr/csj-sjc/crime/rr03_20/rr03_20.pdf).
- O'Gorman, G., & McDonald, G. (2012). *Ransomware: a growing menace*. Retrieved from [https://www.symantec.com/content/en/us/enterprise/media/security\\_response/whitepapers/ransomware-a-growing-menace.pdf](https://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/ransomware-a-growing-menace.pdf).
- O'Kane, P., Sezer, S., Carlin, D., 2018. Evolution of ransomware. *IET Networks* 7 (5), 321–327.
- Paquet-Clouston, M., Haslhofer, B., Dupont, B., 2019. Ransomware payments in the bitcoin ecosystem. *J. Cybersecur.* 5 (1) tyz003.
- Reddit. (2019). The front page of the internet. Retrieved from <https://www.reddit.com/>.
- Rossey, Q., Décarry-Héту, D., 2017. Internet traces and the analysis of online illicit markets. In: *The Routledge International Handbook of Forensic Intelligence and Criminology*. Routledge, pp. 249–263.
- Symantec. (2019). Internet security threat report. Retrieved from <https://www.symantec.com/content/dam/symantec/docs/reports/istr-24-executive-summary-en.pdf>.
- TheDarkWebLinks. (2019). Darkweb markets. Retrieved from <https://www.thedarkweblinks.com>.
- Thomas, K., Huang, D., Wang, D., Bursztein, E., Grier, C., Holt, T.J., ... Vigna, G., 2015. Framing dependencies introduced by underground commoditization. In: Paper presented at the Workshop on Economics of Information Security (WEIS).
- Vaas, L. (2012, April 23). Tor-hidden online narcotics store, 'The farmer's market', brought down in multinational sting. Retrieved from <https://nakedsecurity.sophos.com/2012/04/23/farmers-market-tor-narcotics/>.
- Van Wegberg, R., Tajalizadehkhoob, S., Soska, K., Akyazi, U., Gañán, C., Klievink, B., ... Van Eeten, M., 2018. Plug and prey? measuring the commoditization of cybercrime via online anonymous markets. In: Paper presented at the Proceedings of the 27th USENIX Conference on Security Symposium. Baltimore, MD, USA.
- Van Wegberg, R., Verburgh, T., Van den Berg, J., & Van Staalduinen, M. (2017). Alphabay exit, hansa-down: dream on? Retrieved from <https://dws.pm/download/PUB/17-9099-factsheetbrochure-dws-05.pdf>.
- Waldrop, M.M., 2016. How to hack the hackers: the human side of cybercrime. *Nature* 533 (7602).
- Wehinger, F., 2011. The dark net: self-regulation dynamics of illegal online markets for identities and related services. Paper presented at the European Intelligence and Security Informatics Conference (EISIC).
- Yip, M., 2010. An investigation into Chinese cybercrime and the underground economy in comparison with the West (Master Thesis). University of Southampton.
- Yip, M., Webber, C., Shadbolt, N., 2013. Trust among cybercriminals? Carding forums, uncertainty and implications for policing. *Polic. Soc.* 23, 516–539.
- Young, A.L., Yung, M., 2017. On ransomware and envisioning the enemy of tomorrow. *Computer* 50 (11), 82–85.
- Zhuge, J., Holz, T., Song, C., Guo, J., Han, X., Zou, W., 2009. Studying malicious web-sites and the underground economy on the Chinese web. In: *Managing Information Risk and the Economics of Security*. Springer, pp. 225–244.

**Per Håkon Meland** is a senior research scientist at the independent research institute SINTEF in Norway. He obtained his M.Sc. degree in Computer Science at the Norwegian University of Science and Technology in 2002, where he is also a Ph.D. fellow in the intertwined fields of threat modelling and security economics.

**Yara Fareed Fahmy Bayoumy** completed her Information Systems Master degree at the Norwegian University of Science and Technology in 2018. She earned a Bachelors Degree in Computer and Communication Engineering from Alexandria University in 2015.

**Guttorm Sindre** is a professor at the Department of Computer Science, Norwegian University of Science and Technology, and is also leader for the Excited Centre for Excellence in IT Education. He obtained his Ph.D. from the Norwegian Institute of Technology in 1990. His research interests are in requirements engineering, security requirements, and IT education and didactics.

**K: ‘Breaking the cyber kill chain by modelling resource costs’**

A license to reproduce the published material [27] for inclusion in this thesis has been obtained from Springer Nature.



# Breaking the Cyber Kill Chain by Modelling Resource Costs

Kristian Haga<sup>1</sup> , Per Håkon Meland<sup>1,2</sup> , and Guttorm Sindre<sup>1</sup> 

<sup>1</sup> Norwegian University of Science and Technology, Trondheim, Norway  
{kristian.haga,per.hakon.meland,guttorm.sindre}@ntnu.no

<sup>2</sup> SINTEF Digital, Trondheim, Norway

per.h.meland@sintef.no

<https://www.ntnu.no/>

<https://www.sintef.no/>

**Abstract.** To combat cybercrime, a clearer understanding of the attacks and the offenders is necessary. When there is little available data about attack incidents, which is usually the case for new technology, one can make estimations about the necessary investments an offender would need to compromise the system. The next step would be to implement measures that increase these costs to a level that makes the attack unattractive. Our research method follows the principles of *design science*, where cycles of research activities are used to create artefacts intended to solve real-world problems. Our artefacts are an approach for creating a *resource costs model* (RCM) and an accompanying modelling tool implemented as a web application. These are used to find the required attacker resources at each stage of the cyber kill chain. End user feedback show that structured visualisation of the required resources raises the awareness of the cyberthreat. This approach has its strength and provides best accuracy with specific attacks, but is more limited when there are many possible attack vectors of different types.

**Keywords:** Cyber kill chain · Costs · Resources · Profiling · Attack tree

## 1 Introduction

As our use of technology in almost every aspect of life steadily increases, so does our exposure to cybercrime. To combat this growing form of criminality, a clearer understanding of the costs, benefits and attractiveness of cyberattacks is necessary [18]. This is in accordance with *Routine Active Theory* [5], extended to include cybercrime [6,8], which states that crime will occur when all of the following four conditions are met: There exist an 1) *accessible and attractive target*, 2) *the absence of a capable guardian* and the presence of 3) *a motivated offender* with 4) *the resources required to commit the crime*. For the latter case, it is not just a question of technical skills, but also a requirement that the offender

is able to invest in software development and hardware acquisition, as well as the time it takes to plan, prepare and perform the attack. Alternatively, the offender could bribe an insider or hire someone else to do it through cybercrime-as-a-service [21] being offered by third parties.

We hypothesize that during threat analysis, it is possible to reduce the complexity of the resource requirement to a monetary concern, complemented by a limited set of attacker characteristics. This will allow us to identify the potential offenders and come up with technical and non-technical mitigations that will significantly increase the attacker costs.

The contribution of this paper is a modelling approach that maps resource costs to each stage of a cyberattack, and derives the total cost of the attack. We have utilized principles from Schneier's *attack trees* [32] and the Lockheed Martin's *cyber kill chain* [13], both already widely known in the security community, to structure this approach. A dedicated prototype tool has been developed to simplify and visualise this process, and we have completed the first rounds of iterative evaluation among experts. This tool is able to show calculations interactively and extract potential offenders based on a built-in library from available cybercriminal profile literature. Our goal is to improve the accuracy of threat analysis, and especially increase the understanding and awareness of cyberthreats among sectorial domain stakeholders.

This paper is structured as follows. Section 2 gives an overview of background knowledge and literature, and Sect. 3 explains our method. Results are given in Sect. 4, which are discussed in the light of evaluations in Sect. 5. Finally, Sect. 6 concludes the paper.







## 2 Background

### 2.1 The Cyber Kill Chain

Already in 1998, Meadows [23] presented a way of dividing attacks into different stages or phases to make visual representation easier. The next stage would not commence before the previous one had completed, and she used different colours to represent the assumed difficulty of each stage. The stages were not predetermined, but varied according to the nature of the attack. Later on, McQueen et al. [22] defined a set of five fixed stages, *reconnaissance*, *breach*, *penetrate*, *escalation* and *damage*, which were then modelled as a compromise graph in order to find the weakest link(s) in the attack path based on expected time-to-compromise. Hutchins et al. [14] describe different phase-based models from military usage (countering terrorist attacks) and the information security field (between 2008–2010), and present their own version nicked the *intrusion kill chain*. This model was later on renamed and branded as the *cyber kill chain* [13] by Lockheed Martin, and has proven to be widely popular among defenders of IT and enterprise networks [1]. The seven stages of the cyber kill chain are:



**Reconnaissance:** Research, identification and selection of target.

-  **Weaponization:** Coupling a malware (e.g. remote access trojan) with an exploit into a deliverable payload, e.g. a media file.
-  **Delivery:** Transmission of the weapon to the targeted environment, e.g. an email attachment or USB-drive.
-  **Exploitation:** Triggers malicious code. Ranges from auto-executing within the host's operating system to users triggering execution.
-  **Installation:** Installation of the malware on the victim system, allowing the adversary to maintain presence inside the environment.
-  **Command and Control (C2):** Establishes a channel for the adversary to access the target environment.
-  **Actions on Objectives:** Complete attack objectives, such as data extraction, establish hop point, break integrity or make system unavailable.

According to Hahn et al. [10], a developed cyber kill chain provides the basis for a “systematic study of how the various cyberattack steps and phases can perturb the system layers and eventually impact physical operations”. This is subsequently used in their analysis framework to develop security properties and design systems resilient to cyberattacks. As shown by Pols [27], there are many variants of the kill chain found in the literature. Some with different stage types and others with up to eighteen different stages. We chose to focus our work on the original seven stage cyber kill chain due to its popularity.

## 2.2 Attack Tree Cost Modelling

Attack trees are acyclic graphs used to model threats from the viewpoint of the perpetrator. Schneier's original attack tree paper [32] showed how different costs could be assigned to alternative leaf nodes and how these propagated to define the cheapest way of attack. A fundamental paradigm for this kind of modelling is the assumption of a *rational attacker* [3], meaning that 1) *there will be no attack if the attack is unprofitable* and 2) *the attacker chooses the most profitable way of attacking*.

There have also been several approaches where costs are used in combination with other attributes. For instance, Buldas et al. [3] include costs, gains, penalties and associated probability values. Further examples of different attributes and references to papers that utilize costs in attack trees is given by Bagnato et al. [2]. Having more attributes enables additional ways of analysing attack trees, for instance Kumar et al. [19] show how to find the minimum time to complete an attack given a specific budget. Jensen et al. [15] present an approach where cost is a function of time instead of a constant cost per atomic attack attempt. Still, the major challenge of assigning accurate attribute values to attack tree nodes is difficult to overcome as attacker-specific information tends to be based on a best guess [31].

A comprehensive overview of more than thirty attack and defence modelling approaches based on directed acyclic graphs can be found in a survey paper by Kordy et al. [17]. A more recent survey focusing on fault and attack trees has been published by Nagaraju et al. [24].

### 2.3 Cybercriminal Profiling

Shinder and Tittel [33] define a *profile* to be a set of characteristics likely to be shared by criminals who commit a certain type of crime. The use of profiles during criminal investigations can be traced several hundred years back in time, and though this is not an exact science, Nykodym et al. [25] argue that the track record legitimates the concept. However, they also argue that attackers have more advantages in a cyber setting as they do not have to be physically present at the crime scene.

The two main methods for profiling are known as *inductive* and *deductive* [37]. In the former, a profile database is developed based on information from already committed crime, and offender characteristics are correlated with types of crime. In the latter, forensics evidence is gathered from the crime scene and used to deduce the characteristics of the offender. Most of the established literature comes from the digital forensics field and relates to deductive profiling. We have been mostly interested in inductive profiling as a tool to identify potential offenders before any crime is actually committed. Furthermore, it is well established that likely offenders have  *motive, means and opportunity* (MMO) [26,35] before committing any crime. As attacker costs belongs to the  *means* characteristic, the literature becomes more limited. Warikoo et al. [37] have  *capability factor* as one of their six profile identification metrics, where available resources for e.g. purchasing malware belongs. Preuß et al. [28] created a small set of profiles based on twelve cybercrime cases between 1998 and 2004. Due to the limited sample size, they could not create a structured set of attributes for these, but found that the principle of  *minimum costs and maximum results* were present in all. Casey [4] presents a threat agent library of archetypal cybercriminal agents where  *resources* is one of the eight attributes defining them. Casey's work is used to define  *Attack Resource Level* in the cyberthreat exchange format  *STIX* [16].

## 3 Method

Our research method follows the principles of  *design science*, supporting a pragmatic research paradigm where artefacts are created to solve real-world problems by cycling through research activities related to  *relevance, design and rigor* [11,34]. The problem we try to address is the challenge of quantifying cyberrisks when there is little reliable historical data about attacks. Our artefacts are 1) an approach for creating a  *resource costs model* (RCM), that is used to find the required attacker investments at each stage of the cyber kill chain and 2) an accompanying modelling tool implemented as a web application.

As a part of the relevance cycle, we initially worked with opportunities and problems related to cybersecurity for maritime shipping. We analysed typical vulnerabilities and threats towards eNavigation systems, and made cost estimations for attacking the various underlying technology modules.

During the rigor cycle, past knowledge, as presented in Sect. 2, was examined and we chose to build on practices that already had a significant uptake among practitioners.

Most central to design science research is the design cycle, consisting of artefact construction, evaluation and refinements based on feedback. Initially, we applied “pen-and-paper” variants of the RCM and validated the expressiveness by constructing models of known cyberattacks towards maritime systems. The second iteration produced a *minimum viable product* (MVP) of the tool. Ries [29] defines a MVP as the version of a new product which allows developers to collect the maximum amount of validated learning about customers with the least effort. Our MVP consisted of an info page tutorial and functionality for building basic resource costs models for each attack phase. For the evaluation we recruited eight security experts who modelled a specific use case. These were observed during modelling and debriefed afterwards. The third iteration added the cybercriminal profiling feature, improved the user interface, as well as tweaking flawed features and functions. This evaluation included another eight security professional from the industry and two maritime domain experts.

## 4 Results

### 4.1 The Resource Costs Model

In a *resource cost model* (RCM), each stage in the cyber kill chain represents the root node of a *resource tree*, depicted in Fig. 1, which is similar in structure to an attack tree.

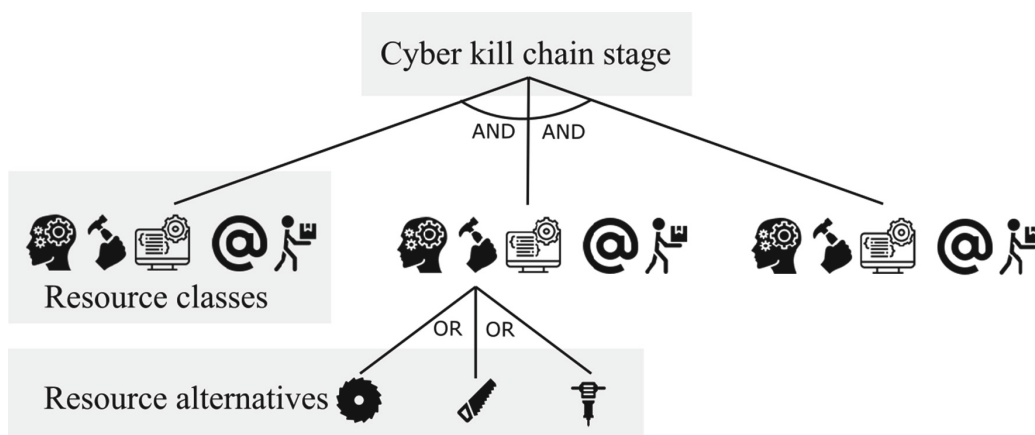


Fig. 1. A resource tree for a single cyber kill chain stage





The second level of the tree defines which resource types are required to complete the parent stage. At this level, all nodes have a conjunctive (*AND*) relationship since an attack would require all necessary resources. A resource can belong to five different classes:



**Skill:** Includes domain knowledge, malware development abilities or utilisation of cybercrime tools or guides.



**Tangible:** Necessary hardware components or other physical objects. This can range from advanced technology to soldering tools.



**Logic:** Commercially available software, data sets or cybercrime tools or services.



**Logic-atomic:** Necessary resources that cannot be broken into smaller parts, e.g. an IP-address, email address or a password.



**Behavioral:** Actions that must be conducted as a part of the attack, for instance bribing, sending out phishing emails or social engineering.

The third level in the tree, *resource alternatives*, are disjunctive (*OR*) leaf nodes that present ways to realize their parent resource class. Each resource alternative is associated with a cost interval and a confidence value. A confidence close to zero communicates that there is little evidence to support the stated cost interval. At the other end of the scale, a confidence of 1 means that there is exhaustive evidence to back the stated cost interval and that the price of the resource is not subject to great variation.

We can express the total cost interval of the attack  $T$  formally by stating that all resources  $R_j$  need to have a valid set  $V$  of resource alternatives. Let  $\alpha$  represent the minimum estimated cost of the cheapest resource alternative and  $\beta$  represent maximum cost of the most expensive resource alternative. From this we can derive the following:

$$T = [(min\ cost = \sum_{\substack{stage \in \\ kill\ chain}} \sum_{i \in V} \alpha_i), (max\ cost = \sum_{\substack{stage \in \\ kill\ chain}} \sum_{i \in V} \beta_i)] \quad (1)$$

By letting  $\phi$  be the average confidence of the  $n$  resource alternatives associated with a resource  $R_j$  and  $c_i$  is the confidence of a resource alternative  $i$  associated with  $R_j$ , we get the following associated confidence  $C$  of the total cost:

$$\phi_j = \frac{\sum_{i \in R_j} c_i}{n} \quad (2)$$

$$C = \prod_{\substack{stage \in \\ kill\ chain}} \prod_R \phi_j \quad (3)$$

In order to mitigate an attack, at least a one of the resources throughout the cyber kill chain must be made too expensive for the adversary. However, the adversary only needs a single resource alternative for each of the resources.

## 4.2 The IRCM Tool

To validate the modelling approach, we have built an interactive installation of the model in the form of a web application called *Interactive Resource Cost Model* (IRCM) tool. This allows the users to model cyberattacks of their choosing, while concurrently deriving the total cost of the attack and probable cybercriminal profiles able to conduct it. An example screenshot from a single resource tree is shown in Fig. 2, while a screenshot of the RCM for the complete cyber kill chain is included in Appendix A.



**Fig. 2.** A screenshot resource tree from the reconnaissance stage

These examples are taken from the maritime domain, where the *Electronic Chart Display and Information System* (ECDIS) is a central component for ship navigation. It displays the vessels position on a chart and integrates information from a number of sensors, such as radar, gyro, GNSS, echo sounder, weather measurements and the anti-collision systems. Malicious manipulation of this position could cause confusion on the ship bridge and potential course alteration could lead to collisions in congested waters [38]. The examples are loosely based on the demonstrated attack against an air-gapped ECDIS system by Lund et al. [20]. This attack was also structured according to the cyber kill chain, but in contrast to an external attack, it was conducted in cooperation with the Royal Norwegian Navy. Also, no information about resource costs were given, so here we have made our own estimations.

As can be seen in Fig. 2, there are four resources defined for the reconnaissance stage. The first one, *ECDIS documentation*, is a tangible class, and the alternatives are to either *purchase* the documentation from the vendor legally, or *steal* it. The second resource is another tangible class, and represents an operational ECDIS unit that can be used to analyse its operating system, software and network traffic. It can be realized in different ways, by *purchasing a unit*

from vendor or the *black market*, or running it as a software *simulation*. These alternatives vary in price, from relatively cheap software (where you pay according to sailing route) to more expensive hardware units in the range of \$10 000 - \$30 000. The third resource is of class logic-atomic, and represents information about the *ship inventory* used to determine which type and where the ECDIS units are installed. To simplify the model, only a single *bribe insider* alternative is used. The final resource is also of type skill, and represents required knowledge about *vulnerabilities* gained through *scanning and testing*.

Both resources and resource alternatives are created by using the tool input data forms. An example screenshot for the ECDIS resource alternative *purchased from vendor* is shown in Fig. 3.

**IRCM** Info Attacker Profiles

**Add a new resource alternative to ECDIS unit**

Name  
Purchase from vendor

Description  
There is a wide range of dedicated ECDIS units available for purchase.

Maximum cost:  
30000

Minimum cost:  
10000

Confidence:  
(Where 0.1 indicates that you have no idea what the cost is and 1 indicates that you are sure of the cost, for example the cost of commercially available hardware)  
0.9

Motivation (number of hours):  
(The time it will take an attacker to acquire or realize this resource alternative. As an example; How many hours will it take the attacker to develop a malware or to acquire a hardware component)  
2

Technical Skill:  
(The academic and technical level an attacker must possess to be able to realize this resource alternative)  
Minimal

Legal Limit:  
(Can the resource alternative be acquired or realized Legally or Illegally?)  
Legally

Access level:  
(Does the resource alternative require Internal or External access level in order to be realized?)  
External

Create Alternative

**Technical skill levels**

**None:** The resource alternative require no expertise or training to be realized

**Minimal:** The resource alternative can be realized through copying code and utilizing existing techniques and tools

**Operational:** The resource alternative require an understanding of the underlying technology and methods used. The requirement to create a new attack or hacking tool falls into this category

**Adept:** The resource alternative require an expertise in technology and attack methods to be realized.

**Fig. 3.** A screenshot from the resource alternative window

The tool has a built-in database of cybercriminal profiles that the model inductively retrieves candidates from. This database is summarized in Appendix B and has been based on profile definitions we have found in the literature [4, 16, 30, 37]. We found out that mapping total attack cost with assumed *wealth* was not a very useful way of doing this. The wealthiest attacker is not

always the most likely one, and attackers have more than one characterizing dimension. Therefore, the tool is able to exclude improbable attacker profiles from the database based on optional information that is assigned to the resources in the RCM. The exclusion rules are based on the following:

- Total minimal *cost* exceeds the financial capacities of the profile [*no cost, low, medium, high*].
- The accumulated time to require all resources exceed its *motivational* limit [*no time, low, medium, high*].
- Any resource alternative that requires a higher *technical skill* level than the profile possesses [*none, minimal, operational, adept*].
- Any resource that requires *moral limits* to be broken [*legally, illegally*].
- Any resource that require an *access level* the profile does not possess [*internal, external*].

The extended ECDIS attack example in Appendix A shows aggregated model information based on input contained in the individual resource tree for each attack stage. The cost interval has a broad range, mostly due to the choice of purchasing ECDIS hardware unit versus other cheaper alternatives in both the *reconnaissance* and *delivery* stages. Besides from these, the overall resource costs related to tangible and skill are relatively low. By analysing the model, we find that there are significant costs related to the *delivery* stage as the attacker would need physical presence at the ship and gain access to the bridge or bribe an insider. It is the air-gapping of the ECDIS that provides the main security measure by making delivery costly. When considering opening up for online software and chart updates, it is clear that additional secure measures will be needed to preserve an expensive attack vector. The confidence value is also very low, but would have been much higher if we had modelled the attack with a specific ECDIS unit in mind where costs are more certain. Also, a higher number of resources will automatically yield a lower confidence, which is natural since acquiring many resources increases uncertainty. The main benefit of the confidence is for attack comparison, which is not shown in these examples. Given the various exclusion rules that have been applied to the model, the most probable attacker profile in this case is *cyber warrior* (described in Appendix B). The cyber warrior profile is not limited by financial requirements of this attack, has a high technical skill level and has little concern for moral limits.

## 5 Discussion

Hong and Kim [12] have pointed to the inherit challenge with graph-based attack models, namely the ability to scale. A purely tree-based model will generate large, bewildering attack trees for complex attacks. In turn, this creates a conflict between analysis and comprehensibility [7]. Hence, some sort of decomposition is needed. We chose to combine two modelling techniques to amplify their advantages and overcome some of their shortcomings. The cyber kill chain allows us

to divide the attack into seven consecutive steps, and by breaking the chain in the early stages we don't have to embellish the later ones. The relatively small resource tree for each of the stages breaks down composite resource requirements into atomic ones, which can be more accurately estimated. This was the main takeaway from the first iteration of the design cycle. Secondly, we experienced that deriving a cost interval, rather than a single estimate, provides more confident information regarding the availability of an attack. A cheap, more available resource alternative set may provide a less stealthy attack than an expensive alternative. By determining both the minimum and maximum cost, we include both the risk willing and risk averse offenders. A large cost interval does not necessarily imply an inaccurate cost estimate, but rather that the evaluated attack can be carried out with a wide span of sophistication and possible impact on the target.

The second iteration involved eight expert end users from a research institute who were observed using the MVP of the tool and debriefed afterwards. Seven out of these eight expressed that the main difficulty was to understand the difference between *resource* and *resource alternative* in the models. We were also able to observe that classifying resources was not straightforward, and the users spent some time navigating between the information page and the modelling interface to check definitions and the tutorial example. Both of these issues improved quickly with hands-on experience and by refining the info page. It was stated during the debrief that “especially interesting is the fact that making only a single resource unavailable, thus breaking the kill chain, will mitigate the entire attack” and all independently agreed that the structured visualisation of the required resources would raise the awareness of the cyberthreat. Some also expressed that many of the resources are impossible to make unavailable, which is true of course. In the MVP, we used *attack trees* as the tree structure term, and this caused some confusion since the RCM focus on resource required to perform the attack and not the attack actions, hence we changed this to *resource tree*.

The third iteration had a focus on inducing criminal profiles from the models and made several improvements to the MVP. We recruited eight professionals from the security industry and two maritime domain experts as end users. Feedback showed that the approach improves the understanding of attacks. The cheapest attack options were considered the most probable, which is helpful when identify mitigation efforts. One of the domain experts encouragingly commented: “It is still a lot of guesswork, but it is systematic guesswork”. Being able to document and provide traceability to threat estimations is vital for industries which require safety and security certification of components. More details of these evaluations can be found in the report by Haga [9]. Parallel to this, Walde and Hanus [36] successfully employed the RCM to plan the purchase of necessary components in order to demonstrate a GNSS spoofing attack.

As already mentioned, the wealthiest attacker is not always the most likely one, therefore we are using five identifying attributes as exclusion rules. A known limitation is that none of these say much about the *motive* of the offender, that is *why* she would commit the crime. This has been out of our scope, but could be extended by looking at the attack impact and attacker reward. Those considerations would have to be determined on a case-by-case basis, requiring additional knowledge dimensions. There is a general criticism towards the cyber kill chain that it focuses too much on the perimeter and malware attack vector [27], and we have seen supportive evidence of that too. Therefore, future improvements could be to include other sets of stages more suitable to describe attacks such as for instance related to social engineering, denial-of-service or code injection.

## 6 Conclusion

Through the iterative nature of design science we have made many improvements to the RCM modelling approach and the accompanying tool. However, we still consider this work to be in progress with many potential improvements related to usefulness and usability. We are also planning to extend the user testing and evaluation, particularly in the field of maritime cybersecurity, but also in other domains to ensure that the artefacts could have a wider usage than just the maritime context. Nevertheless, there is no silver bullet to threat modelling. We are trying to address the real-world problem of missing historical incident data, which is a particular concern for new technology. Attacker costs is one aspect that could be useful during threat estimations, but this must be seen in combination with possible attacker reward as well. In addition, defence costs must be compared with possible loss to make an overall risk assessment.

The RCM has its strength and provides best accuracy with specific attacks; when there are few resources and resource alternatives. Hence, we would not recommend this approach when you want to represent attacks with many possible attack vectors of different types. In such cases, several RCMs could be created and compared, but this quickly becomes a tedious task. As always, the analyst should choose the right tool for the job at hand.

**Acknowledgment.** The research leading to these results has partially been performed by the Cyber Security in Merchant Shipping Service Evolution (CySiMS-SE) project, which received funding from the Research Council of Norway under Grant No. 295969.

## A Tool screenshots

See Figs. 4 and 5.

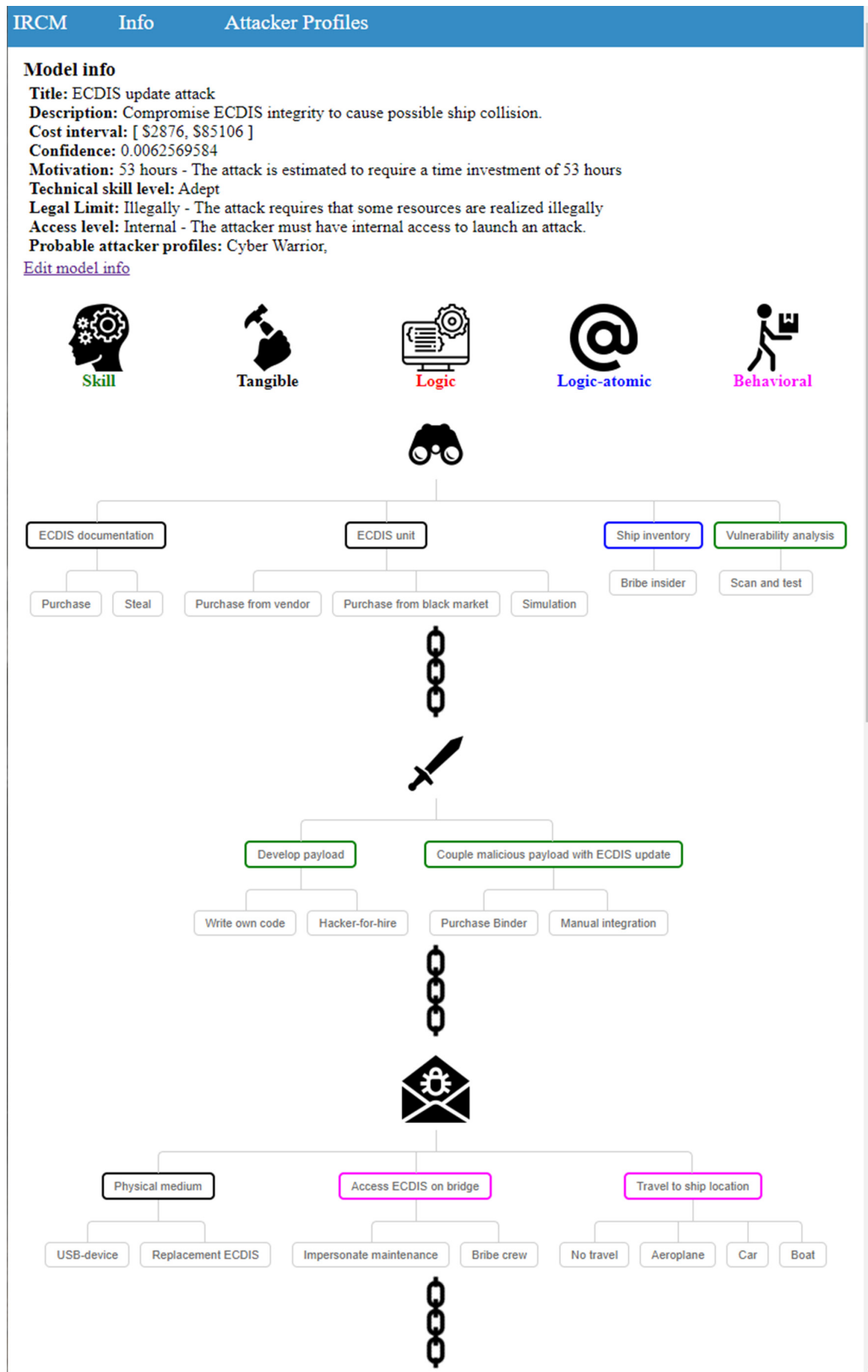
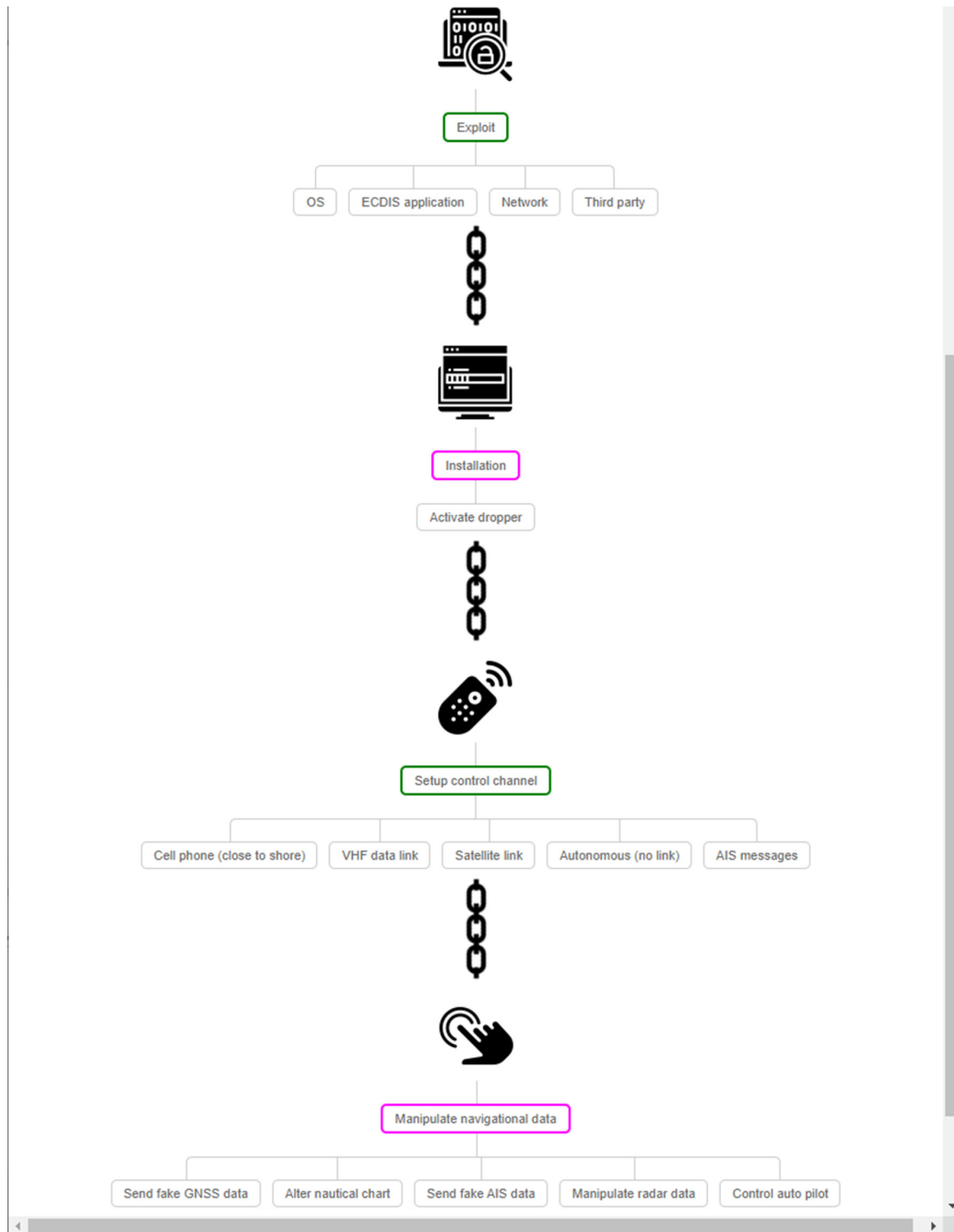


Fig. 4. A screenshot from the first three stages; *Reconnaissance*, *Weaponization* and *Delivery*.





**Fig. 5.** A screenshot from the last four stages; *Exploitation, Installation, Command and Control and Actions on Objectives*








## B Cybercriminal profiles


 **Script kiddie (SK)** has a low level of motivation, thus *time consuming* attacks are not attractive to this profile. The technical skills are limited to *minimal* and the profile only accepts a *minimal* cost. Script kiddies will only utilize resources that can be realized *legally* and have *external* access.


 **Hacktivist (H)** has a medium to high level of motivation anchored in the political cause they represent, thus they may conduct *time consuming*, targeted attacks. The technical skills of a hacktivist is limited to *minimal*. In order to fight for their cause, the hacktivist accepts *some* expenses. The hacktivist is willing to require resources *illegally* and have *external* access level.


 **Vandal (V)** has a low to medium motivation and will only invest a *limited* amount of time in attention seeking attacks. The technical skills of the vandal is limited to *minimal* and the profile accepts a *low* cost. Vandals will only utilize resources that can be realized *legally* and have *external* access.


 **Petty criminal (PC)** has a medium motivation level, willing to invest *some* time in attacks that bring financial gain. They possess *operational* technical skills and accepts a *medium* cost. The petty criminal is willing to require resources *illegally* and has *external* access level.

 **Mobster (M)** has a medium to high level of motivation given that financial gain is possible, thus they may conduct *time consuming* attacks. The technical skills are *operational* and the profile accepts *costly* attacks. Mobsters won't second guess *illegal* resources and have *external* access level.

 **Cyberwarrior (CW)** is a state-sponsored actor with a high motivation level, thus will conduct persistent, *highly time consuming* attacks. The cyberwarrior has *adept* technical skills for launching any attack. In addition, the cyberwarrior is *not limited* by any costs and disposes resources that may be required *illegally*. As an immediate result of the *adept* skill level, the cyberwarrior has *internal* access.

 **Terrorist (T)** tends to be highly motivated and well-funded, thus can conduct *time consuming* and *costly* cyberattacks to front beliefs. The technical skills are limited to *minimal*. The Terrorist is willing to require resources *illegally* and have *external* access level.

 **Internal - Hostile (IN-H)** has a medium motivation level and may launch attacks that require *some* time. The profile knows the system well, which yields an *operational* technical skill. *Some* expenses are acceptable, limited to *legally* acquired resources. Internals have *internal* access level by default.

 **Internal - Non-hostile (IN-NH)** launches cyberattacks by accident, thus *not* motivated at all to invest any time or money in a cyberattack and will only possess resources that can be *legally* realized. Given that accidental cyberattacks are possible yields an *operational* skill level and an *internal* access level.

## References

1. Assante, M.J., Lee, R.M.: The industrial control system cyber kill chain. SANSInstitute InfoSec Reading Room **1** (2015)
2. Bagnato, A., Kordy, B., Meland, P.H., Schweitzer, P.: Attribute decoration of attack-defense trees. *Int. J. Secure Softw. Eng. (IJSSE)* **3**(2), 1–35 (2012)
3. Buldas, A., Laud, P., Priisalu, J., Saarepera, M., Willemson, J.: Rational choice of security measures via multi-parameter attack trees. In: Lopez, J. (ed.) *CRITIS 2006*. LNCS, vol. 4347, pp. 235–248. Springer, Heidelberg (2006). <https://doi.org/10.1007/11962977-19>
4. Casey, T.: Threat agent library helps identify information security risks. Intel White Paper **2** (2007)
5. Cohoen, L.E., Felson, M.: Social change and crime rate trends: a routine activity approach. *Am. Sociol. Rev.* **44**(4), 588–608 (1979)
6. Ekblom, P., Tiley, N.: Going equipped. *Br. J. Criminol.* **40**(3), 376–398 (2000)
7. Gadyatskaya, O., Trujillo-Rasua, R.: New directions in attack tree research: catching up with industrial needs. In: Liu, P., Mauw, S., Stølen, K. (eds.) *GramSec 2017*. LNCS, vol. 10744, pp. 115–126. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-74860-3\\_9](https://doi.org/10.1007/978-3-319-74860-3_9)
8. Grabosky, P.N.: Virtual criminality: old wine in new bottles? *Soc. Legal Stud.* **10**(2), 243–249 (2001)
9. Haga, K.: Breaking the cyber kill chain by modelling resource costs. Master’s thesis, NTNU, Trondheim, Norway (2020)
10. Hahn, A., Thomas, R.K., Lozano, I., Cardenas, A.: A multi-layered and kill-chain based security analysis framework for cyber-physical systems. *Int. J. Crit. Infrastruct. Prot.* **11**, 39–50 (2015)
11. Hevner, A., Chatterjee, S.: Design science research in information systems. In: *Design Research in Information Systems*, pp. 9–22. Springer, Boston (2010). [https://doi.org/10.1007/978-1-4419-5653-8\\_2](https://doi.org/10.1007/978-1-4419-5653-8_2)
12. Hong, J.B., Kim, D.S.: Performance analysis of scalable attack representation models. In: Janczewski, L.J., Wolfe, H.B., Sheno, S. (eds.) *SEC 2013*. IAICT, vol. 405, pp. 330–343. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-39218-4\\_25](https://doi.org/10.1007/978-3-642-39218-4_25)
13. Hutchins, E.M.: The cyber kill chain. Technical report, Lockheed Martin (2020). <https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html>. Accessed 12 Apr 2020
14. Hutchins, E.M., Cloppert, M.J., Amin, R.M.: Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Issues Inf. Warfare Secur. Res.* **1**(1), 80 (2011)
15. Jensen, P.G., Larsen, K., Legay, A., Poulsen, D.: Quantitative evaluation of attack defense trees using stochastic timed automata. In: *International Workshop on Graphical Models for Security*, pp. 75–90. HAL Id: hal-01640091 (2017)
16. Jordan, B., Piazza, R., Wounder, J.: Stix version 2.0. part 1: Stix core concepts. Technical report, OASIS Committee Specifications 01 (2017) <http://docs.oasis-open.org/cti/stix/v2.0/stix-v2.0-part1-stix-core.html>. Accessed 13 Apr 2020
17. Kordy, B., Piètre-Cambacédès, L., Schweitzer, P.: Dag-based attack and defense modeling: don’t miss the forest for the attack trees. *Comput. Sci. Rev.* **13**, 1–38 (2014)
18. Kshetri, N.: The simple economics of cybercrimes. *IEEE Secur. Privacy* **4**(1), 33–39 (2006)

19. Kumar, R., Ruijters, E., Stoelinga, M.: Quantitative attack tree analysis via priced timed automata. In: Sankaranarayanan, S., Vicario, E. (eds.) FORMATS 2015. LNCS, vol. 9268, pp. 156–171. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-22975-1\\_11](https://doi.org/10.1007/978-3-319-22975-1_11)
20. Lund, M.S., Hareide, O.S., Jøsok, Ø.: An attack on an integrated navigation system. *NECESSE* **3**(2), 149–163 (2018)
21. Manky, D.: Cybercrime as a service: a very modern business. *Comput. Fraud Secur.* **2013**(6), 9–13 (2013)
22. McQueen, M.A., Boyer, W.F., Flynn, M.A., Beitel, G.A.: Quantitative cyber risk reduction estimation methodology for a small scada control system. In: Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS 2006), vol. 9, pp. 226–226. IEEE (2006)
23. Meadows, C.: A representation of protocol attacks for risk assessment. In: Proceedings of the DIMACS Workshop on Network Threats, pp. 1–10 (1998)
24. Nagaraju, V., Fiondella, L., Wandji, T.: A survey of fault and attack tree modeling and analysis for cyber risk management. In: 2017 IEEE International Symposium on Technologies for Homeland Security (HST), pp. 1–6. IEEE (2017)
25. Nykodym, N., Taylor, R., Vilela, J.: Criminal profiling and insider cyber crime. *Comput. Law Secur. Rev.* **21**(5), 408–414 (2005)
26. Pendse, S.G.: Ethical hazards: a motive, means, and opportunity approach to curbing corporate unethical behavior. *J. Bus. Ethics* **107**(3), 265–279 (2012)
27. Pols, P.: The unified kill chain: Designing a unified kill chain for analyzing, comparing and defending against cyber attacks. Cyber Security Academy (2017)
28. Preuß, J., Furnell, S.M., Papadaki, M.: Considering the potential of criminal profiling to combat hacking. *J. Comput. Virol.* **3**(2), 135–141 (2007)
29. Ries, E.: The lean startup : how constant innovation creates radically successful businesses. Portfolio Penguin (2011)
30. Rogers, M.K.: The psyche of cybercriminals: a psycho-social perspective. In: Ghosh, S., Turrini, E. (eds.) Cybercrimes: A Multidisciplinary Analysis, pp. 217–235. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-13547-7\\_14](https://doi.org/10.1007/978-3-642-13547-7_14)
31. Saini, V., Duan, Q., Paruchuri, V.: Threat modeling using attack trees. *J. Comput. Sci. Colleges* **23**(4), 124–131 (2008)
32. Schneier, B.: Attack trees. *Dr. Dobb's J.* **24**(12), 21–29 (1999)
33. Shinder, D.L., Tittel, E.: Chapter 3 - understanding the people on the scene. In: Scene of the Cybercrime, pp. 93–146. Syngress, Burlington (2002)
34. Simon, H.A.: *The Sciences of the Artificial*, 3rd edn. MIT Press, Cambridge (1996)
35. Van Ruitenbeek, E., Keefe, K., Sanders, W.H., Muehrcke, C.: Characterizing the behavior of cyber adversaries: the means, motive, and opportunity of cyberattacks. In: 40th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Supplemental (DSN 2010), pp. 17–18 (2010)
36. Walde, A., Hanus, E.G.: The feasibility of AIS- and GNSS-based attacks within the maritime industry. Master's thesis, NTNU, Trondheim, Norway (2020)
37. Warikoo, A.: Proposed methodology for cyber criminal profiling. *Inf. Secur. J. Global Perspect.* **23**(4–6), 172–178 (2014)
38. Wingrove, M.: Security flaws open ECDIS to cyber crime. Technical report, Riviera (2018). <https://www.rivieramm.com/opinion/opinion/security-flaws-open-eedis-to-cyber-crime-24334>. Accessed 20 Apr 2020






**L: ‘A Systematic Mapping Study on Cyber Security Indicator Data’**

Included is the published material [28], following the terms and conditions of the Creative Commons Attribution (CC BY) license used by MDPI.

L

Review

# A Systematic Mapping Study on Cyber Security Indicator Data

Per Håkon Meland <sup>1,\*</sup> , Shukun Tokas <sup>2</sup> , Gencer Erdogan <sup>2</sup> , Karin Bernsmed <sup>1</sup>  and Aida Omerovic <sup>3</sup> 

<sup>1</sup> Software Engineering, Safety and Security, SINTEF Digital, Strindvegen 4, NO-7465 Trondheim, Norway; karin.bernsmed@sintef.no

<sup>2</sup> Software and Service Innovation, SINTEF Digital, Forskningsveien 1, NO-0314 Oslo, Norway; shukun.tokas@sintef.no (S.T.); gencer.erdogan@sintef.no (G.E.)

<sup>3</sup> Norwegian Computing Center, Gaustadalleen 23a, NO-0373 Oslo, Norway; aida@nr.no

\* Correspondence: per.h.meland@sintef.no

**Abstract:** A security indicator is a sign that shows us what something is like or how a situation is changing and can aid us in making informed estimations on cyber risks. There are many different breeds of security indicators, but, unfortunately, they are not always easy to apply due to a lack of available or credible sources of data. This paper undertakes a systematic mapping study on the academic literature related to cyber security indicator data. We identified 117 primary studies from the past five years as relevant to answer our research questions. They were classified according to a set of categories related to research type, domain, data openness, usage, source, type and content. Our results show a linear growth of publications per year, where most indicators are based on free or internal technical data that are domain independent. While these indicators can give valuable information about the contemporary cyber risk, the increasing usage of unconventional data sources and threat intelligence feeds of more strategic and tactical nature represent a more forward-looking trend. In addition, there is a need to take methods and techniques developed by the research community from the conceptual plane and make them practical enough for real-world application.

**Keywords:** threat intelligence; data-driven decision making; risk management; data sources; trends



**Citation:** Meland, P.H.; Tokas, S.; Erdogan, G.; Bernsmed, K.; Omerovic, A. A Systematic Mapping Study on Cyber Security Indicator Data. *Electronics* **2021**, *10*, 1092. <https://doi.org/10.3390/electronics10091092>

Academic Editors: Changhoon Lee, Yu Chen and Jake (Jaeik) Cho

Received: 12 April 2021

Accepted: 30 April 2021

Published: 5 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cyber risk estimates today tend to be based on gut feeling and best guesses. Improved justification and traceability can be achieved through data-driven decisions, but this is not straightforward. With evolving technology and constantly emerging attack methods (and motivations), basing security decisions on past incidents is typically referred to as “driving forward by looking in the rear-view mirror” [1] and cannot be considered reliable. As a remedy to historical data and guesswork, Anderson et al. [2] suggested in 2008 to use forward-looking indicators as an alternative source of decision data, but now, more than a decade later, have we really succeeded in doing this?

The purpose of this paper is to present a systematic mapping study of the literature related to cyber security indicator data. As defined by Kitchenham and Charters [3] and Petersen et al. [4], systematic mapping studies provide an overview of a research area through classification of published literature on the topic. This is somewhat different from systematic literature reviews, which focus more on gathering and synthesizing evidence [4], typically from a smaller set of publications. We identified relevant research and classified their approaches according to a scheme. This contributes to a broad overview of the research field, showing concentrations of effort and revealing areas that need more attention. We then have the possibility to debate if we still base our risk estimates on guts, guesses and past incidents, or whether we have managed to move the field forward, i.e., towards making informed cyber security decisions from relevant indicators. To guide our investigation, we have defined the following research questions:

1. What is the nature of the research using security indicators?

2. What is the intended use of the data?
3. What is the origin of the data for the indicators?
4. What types of the data are being used?
5. What is the data content of the indicators?

The main contributions of this study are: (1) a broad overview of research efforts in the domain of cyber security indicator data; (2) a detailed and reusable classification scheme that can be used to capture new trends in this area using consistent terminology; (3) an analysis of trends within the literature from 2015–2020; and (4) identification of focus areas for further research.

The target audience for this work are researchers and practitioners who want to establish better data-driven practices for cyber risk estimates.

The rest of the paper is structured as follows. Section 2 presents background information about the underlying concepts that are central to our research focus. Section 3 gives an overview of related work and Section 4 presents the methodology used to conduct our systematic mapping study, including search strings, inclusion/exclusion criteria and an overview of the screening process of papers. Section 5 presents the classification scheme that is used to classify primary studies as well as the mapping results. In Section 6, we discuss the result with respect to the research questions, compare our findings with existing research work and recommend possible directions for future work. Finally, Section 7 concludes the paper.

## 2. Background

The following describes terminology and concepts that are central to our mapping study. An *indicator* is defined by Oxford Advanced Learner's Dictionary [5] as "a sign that shows you what something is like or how a situation is changing". An indicator can for instance be observations of mechanisms and trends within the cybercrime markets, as suggested by Pfleeger and Caputo [6], and indicate relevant cyber threats. One or more *data sources* can be used to determine the status of an indicator. For instance, statistics from a dark net marketplace could be a remote data source, while a system log could be a local data source. There are many possible data sources related to cyber threats, including sharing communities, open source and commercial sources [7]. The term used in the context of sharing such information is usually *threat intelligence*, which is any evidence-based knowledge about threats that can inform decisions [8]. The term can be further defined into the following sub-domains [9,10]:

- *Strategic threat intelligence* is high-level information used by decision-makers, such as financial impact of attacks based on historical data or predictions of what threat agents are up to.
- *Operational threat intelligence* is information about specific impending attacks against the organization.
- *Tactical threat intelligence* is about how threat actors are conducting attacks, for instance attacker tooling and methodology.
- *Technical threat intelligence (TTI)* is more detailed information about attacker tools and methods, such as low-level indicators that are normally consumed through technical resources (e.g., intrusion detection systems (IDS) and malware detection software).

To compare or possibly join data source contents, *metrics* can be useful. Mateski et al. [11] defined a metric to be a standard of measurement and something that allows us to measure attributes and behaviors of interest. An example of a metric is the number of malware sales. A *measure* is a specific observation for a metric, for instance the value 42 for a given week. According to Wang [12], security metrics should be quantitative, objective, employ a formal model, not be boolean (0, 1) and reflect time dependence. There is a plethora of possible security metrics, for instance Herrmann [13] presented more than 900 different ones in her book. The challenge is to find the ones that represent practically useful security indicators.

### 3. Related Work

We are aware of several review papers, survey papers and mapping studies that partly overlap with ours and provide supplementary material. For instance, Humayun et al. [14] performed a systematic mapping study of common security threats and vulnerabilities from 78 articles, covering studies spanning over a decade (2007–2018). A direct comparison of the study by Humayun et al. [14] and our study is not straightforward, mainly because of the different objectives; for example, Humayun et al. [14] focused on an analysis of publication venue, demography of researchers and key targets of cyber attacks. However, there are common features in the two studies, such as the research methodology, choice of academic databases and domain (i.e., cyber security). They also gave an overview of other mapping studies and systematic literature reviews in the cyber security area. Beyond these, there are many related surveys and reviews that we highlight in the following.

In a publication from 2107, Grajeda et al. [15] analyzed 715 research articles from the years 2010 to 2015 with respect to the utilization of datasets for cybersecurity and cyber forensics. They found 70 different datasets and organized them into 21 categories. The datasets were collected and analyzed from both peer-reviewed articles and Google search (for the datasets that may not have appeared in selected articles). Taking a broader perspective, Zheng et al. [16] analyzed their use or creation in nearly 1000 academic papers published between 2012 and 2016. They created a taxonomy for describing the datasets and used machine learning to classify the papers accordingly.

Griffioen et al. [17] evaluated the quality of 17 open source cyber threat intelligence feeds over a period of 14 months and 7 additional feeds over 7 months. Within these, they found that the majority of indicators were active for at least 20 days before they are listed, and that some data were biased towards certain countries. Tundis et al. [18] also surveyed existing open source threat intelligence sources, and, based on interviews with 30 experts (i.e., cyber security professionals and academic researchers), they proposed an approach for the automated assessment of such sources.

In 2016, Pendleton et al. [19] surveyed system security metrics, pointing to big gaps between the existing metrics and desirable metrics. More recently, Cadena et al. [20] carried out a systematic mapping study of metrics and indicators of information security incident management based on 10 primary studies for the period from 2010 to 2019. Our study and that of Cadena et al. [20] share the same motivation, i.e., to support informed security decision-making, but the two differ in addressing terms of research focus. For example, we look into classifying data source, data content, data usage, etc., whereas their focus was on attributes related to cost, quality, service and standards.

In 2018, Husák et al. [21] published a survey of prediction and forecasting methods in cyber security. They also looked at input data for these methods and observed that there are many alternatives with different levels of abstraction. They found that evaluations tend to be based on datasets with high age, which do not necessarily reflect current cyber security threats. Other public datasets are scarcely used or artificially created by the authors to evaluate their own proposed methods. Similarly, Sriavstava et al. [22] found in their review that outdated datasets are used to evaluate machine learning and data mining methods. Sun et al. [23] published in 2019 their survey on datasets related to cyber incident prediction. Nineteen core papers were categorized according to the six data types: *organization's report and dataset*, *network dataset*, *synthetic dataset*, *webpage data*, *social media data* and *mixed-type dataset*.

From their literature survey, Laube and Böhme [24] created a framework for understanding defenders' strategies of privately or publicly sharing cyber security information. They found that, although many theoretical works assume sharing to be beneficial, there is little actual empirical validation.

Diesch and Krcmar [25] investigated the link between information security metrics and security management goals through a literature study. After eliminating duplicates, they found 195 technical security metrics based on 26 articles. They questioned whether all of these are really useful. Kotenko et al. [26] showed how different types of source data





are used in attack modeling and security evaluation. They also provided a comprehensive selection of security metrics.

Gheyas et al. [27] performed a systematic literature review on prediction of insider threats based on 37 articles published between 1950 and 2015. They found that only a small percentage of studies used original real-world data. Tounsi and Rais [9] conducted a survey in 2017 that classified and distinguished existing threat intelligence types and evaluated which were the most popular open source/free threat intelligence tools. They also highlighted some of the problems with technical threat intelligence, such as quality, short-livedness and the overwhelming amount of data, much of it with limited usefulness. Another literature study on threat intelligence by Keim and Mohapatra [28] compared nine of the available open source platforms. They pointed out challenges related to a lack of standardization and ability to select data based on creation date. Samtani et al. [29] reviewed the cyber threat intelligence platforms provided by 91 companies (mostly based in the US). More than 90% of the companies relied either solely or primarily on internal network data. They noted that the Darknet was slowly emerging as a new viable data source for some of the companies. In a literature review on the use of Bayesian Network (BN) models in cyber security, Chockalingam et al. [30] identified the utilized type of data sources. Here, most models used expert knowledge and/or data from the literature, while only a few relied on inputs from vulnerability scanners and incidents data. Furthermore, they found that 13 out of 17 BN models were used for predictive purposes.

#### 4. Methodology

We followed the guidelines and recommendations on systematic mapping studies or scoping studies as proposed by Kitchenham and Charters [3] and Peterson et al. [4,31]. In the planning phase, we established a review protocol, which is an essential element when conducting secondary studies. The review protocol describes the research questions (see Section 1) and methods for conducting the secondary study, such as how the primary studies should be located, appraised and synthesized [32]. Especially when several researchers are involved, a clearly defined protocol reduces the possibility of researcher bias and misconceptions. The following briefly describes the contents of the protocol and implementation.

##### 4.1. Search Keywords

Based on our research questions, we defined an initial set of search keywords, which were used to identify the top relevant papers based on a Google Scholar search. We studied these in detail and applied a *snowballing technique* to find additional papers and a few instances of grey literature that we knew would be relevant. Snowballing refers to using the reference list of a paper, or the citations of the paper, to identify additional papers [33]. The resulting set of 18 core papers were then used as a tool to identify and extract a larger set of keywords. These keywords were then used as basis for defining search strings. As shown in Table 1, we separated between primary keywords to look for in the title and secondary ones for the title, abstract and list of keywords defined by the authors of the primary studies.

**Table 1.** Primary and secondary keywords.

Title Keywords	Title, Abstract, Author Defined
“cyber security”, “information security”, “cyber risk”, “cyber threat”, “threat intelligence”, “cyber attack”	“predict”, “strategic”, “tactical”, “likelihood”, “probability”, “metric”, “indicator”

We tested the keywords by checking if they would re-discover the core papers they were derived from. We also removed some superfluous keywords that did not seem to increase the result set. A general observation from experimenting with search strings was

that combinations with only the keyword “security” in the title would be too ambiguous, returning irrelevant results related to the protection of food, animals, borders and climate. Hence, we developed search strings that would either contain keywords “cyber security” or “information security” to improve accuracy of search results.

#### 4.2. Inclusion Criteria

To limit the result set and support the screening process, we defined a set of inclusion criteria, stating that the studies must be:

- related to actual use of indicator data for cyber security risks;
- published between 2015 and 2020 (the selection does not include studies indexed after September 2020);
- written in English; and
- peer-reviewed.

Similarly, our exclusion criteria stated that the studies should not be:

- in the form of patents, general web pages, presentations, books, thesis, tutorials, reports or white papers;
- purely theoretical in nature and with no use of data;
- about visual indicators for tools (e.g., browser extensions);
- addressing topics related to failures, accidents, mistakes or similar;
- repeated studies found in different search engines; or
- inaccessible papers (not retrievable).

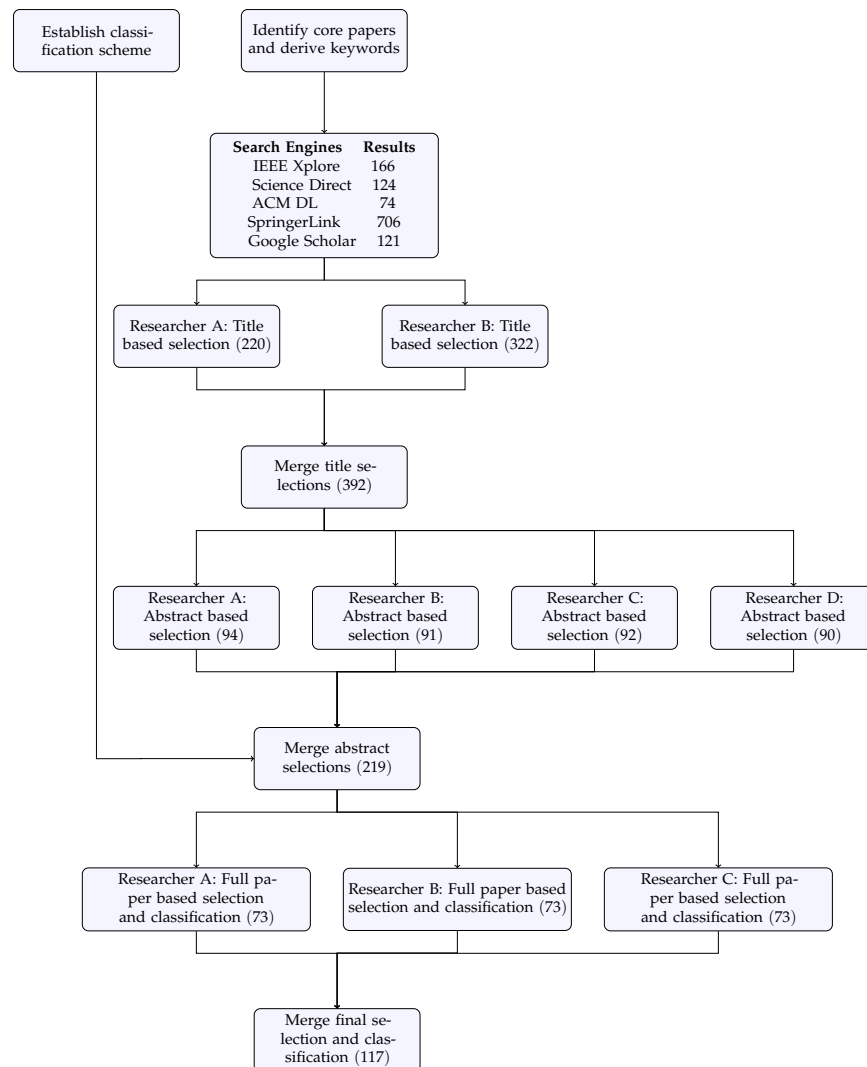
#### 4.3. Database Selection and Query Design

In our study, we chose five online databases: IEEE Xplore, Science Direct, ACM Digital Library, SpringerLink and Google Scholar. These were selected because they are central sources for literature related to computer science and cyber security. Google Scholar is not a literature database by itself, but indexes other databases, so there was bound to be some overlap. For each of the databases, we iteratively defined the search string and conducted manual searches within the database, based on the keywords in Table 1. As Brereton et al. [32] observed, the databases are organized around completely different models and have different search functionalities. It was therefore impossible to use the exact same search strings for all five databases, and we had to tailor the search strings individually. The full definitions of the final search strings that we eventually applied can be found in Appendix A. Most databases order results by relevance, and we therefore applied “ten irrelevant papers in a row” as a stopping criterion. In this way, we did not have to go through the complete result set for all search strings.

#### 4.4. Screening and Classification Process

An overview of the search and screening process is given in Figure 1. This process was initiated during September 2020. Researchers A and B independently ran through every search string for all databases and extracted primary studies based on titles. Each of the two result sets were then assessed by the other researcher. The strategy here was that Researcher B voted on papers selected by Researcher A, while Researcher A voted on papers selected by Researcher B. Duplicates were removed and only those studies with votes from both Researchers A and B were selected for the next stage of the screening. This also included papers for which inclusion/exclusion was hard to decide based on title alone. In total, 392 papers were selected at this stage based on title-screening, for the next stage of abstract/summary-based screening. Due to the number of primary studies, four researchers (Researchers A–D) were involved, and we had to calibrate how papers were selected. To do this, 20 papers were randomly picked out for a test screening where all researchers read the abstracts and made a selection. Afterwards, they compared results and discussed deviations to establish a common practice. Following this, the complete set from the title stage were randomized and divided into four groups, one for each researcher. There was no duplication of efforts (double reading) at this stage, and each researcher got a

unique set to screen based on abstract using our inclusion/exclusion criteria. The result set from the abstract stage yielded 219 primary studies.



**Figure 1.** Mapping study flow chart.

Parallel to the screening process thus far, all researchers had been working on developing a classification scheme to address the research questions. It consisted of 46 parameters, which were partly adopted from related work and partly based on what we had observed in the core papers and selected abstracts. To test the classification scheme itself and to calibrate the researchers for classification, we randomly selected 20 primary studies that Researchers A–C read in full and classified accordingly. As before, the researchers compared and discussed their efforts in a joint session.

In the final stage, the complete set of primary studies from the abstract stage were randomized into three unique groups, fully read, classified and merged. This final result set included 117 primary studies, from which the results in Section 5 were derived. The complete list of the selected primary studies is provided in Appendix B.

### 5. Results

As mentioned in Section 1, systematic mapping studies provide an overview of a research area through classification of published literature on the topic. Thus, in the following, we first present the classification scheme used to categorize the primary studies, and then we present the mapping results with respect to the classification scheme.

#### 5.1. Classification Scheme

The Cyber Security Indicator Data (CSID) classification scheme is illustrated in Figure 2. It covers seven main categories: research type, data openness, data usage, domain, data source, data type and data content. In the following, we describe each category as well as their sub-categories.

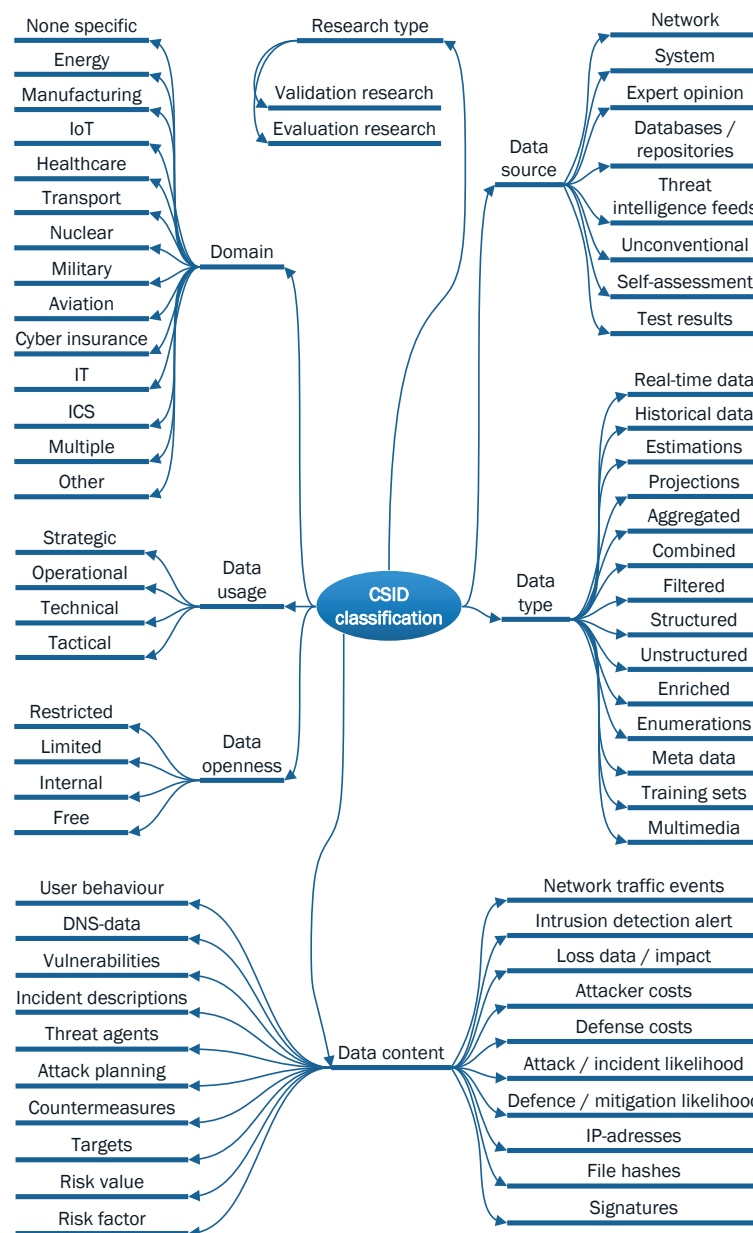


Figure 2. The Cyber Security Indicator Data (CSID) classification scheme.

**Research** type represents different research approaches. Each primary study included in our systematic mapping study is associated with one research approach. As Petersen et al. did in their mapping study [31], we chose to use an existing classification of research approaches by Wieringa et al. [34]. However, based on the exclusion criteria, we disregarded *solution proposal*, *philosophical*, *opinion* and *personal experience* papers and focused on mapping *validation research*, which describes novel techniques with example experiment/lab data, and *evaluation research*, showing how techniques are used in practice with real data and an evaluation.

**Data openness** represents the availability of data reported in the primary studies. We distinguish between the following categories of data openness: *free* in the sense that the data are completely open and freely available; *limited* availability where a membership is required to access data; *restricted* access where data are made available to, e.g., authorities; and *internal* access meaning that the data are only accessible from own system(s). We also considered a fifth category, *commercial*, where access to data requires payment. However, none of the primary studies reported on commercially accessible data and this category is therefore disregarded.

**Data usage** refers to the intended use of data. We consider four categories of data usage: *strategic*, *operational*, *tactical* and *technical*. These categories correspond to the four sub-domains of threat intelligence described in Section 2. Each primary study was associated with one data usage category.

**Domain** refers to an application domain, including *energy*, *manufacturing*, *IoT*, *health-care*, *transport*, *nuclear*, *military*, *aviation*, *cyber insurance*, *IT* and *industrial control systems*. In addition, we included three categories to group the primary studies not addressing a specific domain (*none specific*), a combination of different domains (*multiple*) and finally *other* domains.

**Data source** indicates where the data used in the primary studies originate from. We consider eight non-exclusive data source categories in our classification scheme. *Network* data come from network resources such as firewalls, routers, gateways and DNS-logs. *System* data come from computer resources, typically from internal systems in an organization. *Expert opinion* are indicative variables such as consensus, experience and self-proclamation. *Databases/repositories* provide general data obtained via, e.g., queries. *Threat intelligence feeds* are obtained through subscription-based push services. *Unconventional* data are open source indicators that are either not directly related to the target or not made to predict threats, such as data from marketplaces, forums, blogs and social media. *Self-assessment* data are obtained from internal forms or surveys. *Test results* come from internal tests, typically obtained from tools for penetration testing, vulnerability scanners, etc.

**Data type** refers to the nature of the data. We consider 14 non-exclusive categories of data type. *Real-time data* are obtained from real-time events via, e.g., sensors. *Historical data* can be log data and recorded frequencies of particular events. *Estimations* are based on incomplete data. *Projections* are made to reflect future values. *Aggregated* data are based on similar content, e.g., aggregated cost. *Combined* data emerge when different data types are used to create other data. *Filtered* data are obtained when values have been removed or masked for some reason, e.g., to preserve anonymity. *Structured* data are clearly defined data types whose pattern makes them easily searchable and interpretable. *Unstructured* data are more difficult to find and interpret, such as audio, video and social media postings. *Enriched* data are improved in some way, e.g., by adding missing details. *Enumerations* are catalogues of publicly known information, such as the Common Weakness Enumeration (CWE) [35]. *Meta data* are data about data, include ontologies and language specification. *Training sets* cover artificial data used for testing, training or simulation. *Multimedia* are mostly temporal media such as video and audio.

**Data content** refers to the metrics provided by the data sources. We consider 20 non-exclusive categories of data content. *Network traffic events* are recorded events in the network layer that can indicate an attack. An *intrusion detection alert* originates from either network or computer resources. *Loss data/impact* are about the measured effects/costs

of an attack. *Attacker costs* reflect the required investments to successfully perform an attack. *Defence costs* reflect the required investments to successfully mitigate an attack. *Attack/incident likelihood* is a measurement of the (qualitative or quantitative) likelihood of a successful attack or incident. *Defence/mitigation likelihood* is the (qualitative or quantitative) likelihood of a successful defence or mitigation of an attack. *IP-addresses* include blacklisted ones or those with suspicious activity. *File hashes* are used to identify malicious files, such as malware. *Signatures* are code signatures that may be used to identify, e.g., a virus. *User behavior* reflects content about how people interact in a system, e.g., by monitoring the behavior of employees. *DNS-data* can for instance be poisoned DNS servers or addresses. *Vulnerabilities* are descriptions of such found in software/hardware. *Incident descriptions* reflect real security incidents and breaches. *Threat agents* are descriptions of attributing threat agents. *Attack planning* is information obtained from discussions in forums and social media. *Countermeasures* describe recommended preventive or reactive countermeasures for certain threats. *Targets* are descriptions of identified targets exposed to attacks. *Risk value* means the combined likelihood and impact values, i.e., for a specific domain, organization type or size. *Risk factor* contains values related to risks, such as probability, likelihood, frequency, uncertainty, confidence, consequence or impact.

### 5.2. Mapping Results

In the following, we present the result of our systematic mapping study with respect to the classification scheme described in Section 5.1. A CSV dataset, which includes this scheme and the details of our current classification of primary studies, is available as open research data [36] in order to provide openness, traceability and possible extensions of our work.

As shown in Figure 3, there has been a linear growth in the number of primary studies per year in the period 2015–2020. From being a relatively narrow field with only a handful publications, the increase shows that research on security indicator data is becoming popular. We do not have an exact number for 2020 since the study was conducted before the end of that year. However, the dotted regression line has an annual slope of 7.2, which yields about 40 new publications for 2020.

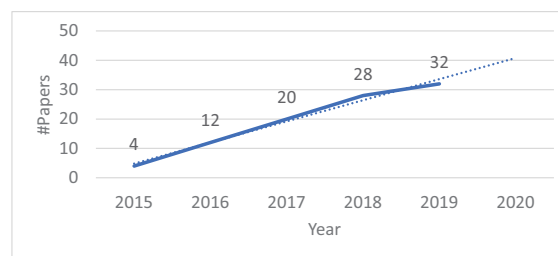


Figure 3. Number of papers per year.

Figure 4 shows a bubble chart illustrating a matrix comprised of the four *data usage* categories (strategic, operational, tactical and technical) and the 14 *domain* categories (energy, manufacturing, IoT, etc., including none specific, multiple and other). Each of the 117 primary studies are grouped in the bubble chart based on a pair of categories ( $x, y$ ), where  $x$  represents a category of domain application and  $y$  represents a category of data usage. The numbers in the matrix represent the number of primary studies that fall under each pair of categories, which is also reflected by the size of the bubbles.

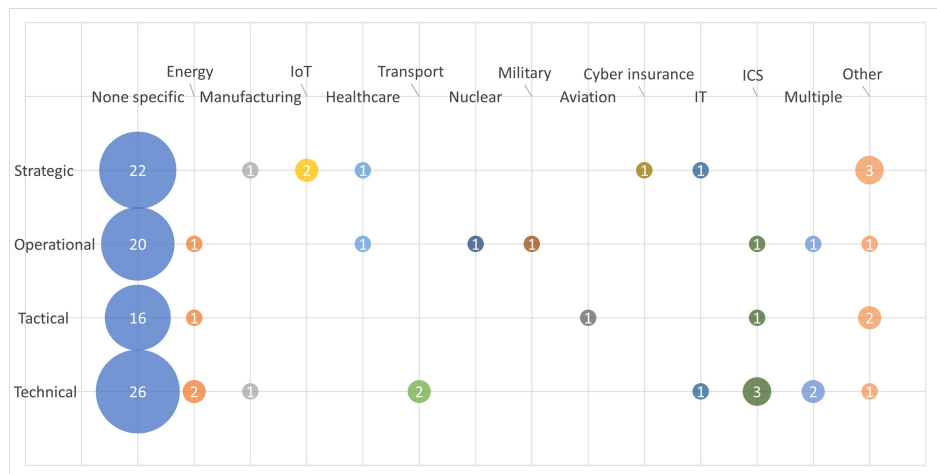


Figure 4. Data usage versus domain.

We can also see from Figure 4 that the majority of the primary studies (84 out of 117) do not address any specific usage domains. Moreover, 26 of these 84 primary studies use technical data, 22 use strategic data, 20 use operational data and 16 use tactical data. Considering the primary studies across all domains from the data usage perspective shows that most of the primary studies use technical data (38), followed by strategic data (31), operational data (27) and tactical data (21). Besides the domain categories *none specific*, *multiple* and *other*, the remaining domain categories are addressed by at least one primary study.

As explained in Section 5.1, we group the primary studies with respect to research type facets. The diagram in Figure 5 shows that the primary studies mostly belong to *validation research* (87 papers), with much less representation within *evaluation research* (30 papers).

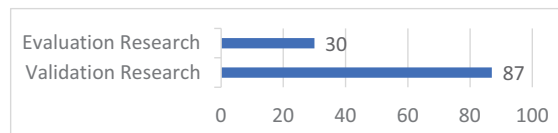


Figure 5. Research type facet.

In terms of data openness, we discovered that the data used in the primary studies mainly fall under the categories *free* or *internal* (see Figure 6). In total, 56 out of 117 (48%) primary studies use data that are *free*, while 46 out of 117 (39%) use *internal* data. From the remaining primary studies, only 12 (10%) use *limited* data and 3 (3%) use *restricted* data. When the study used more than one type of data openness, we classified according to the strictest one.

With respect to the origin of data, we see from Figure 7a that the two most popular data sources are network related data obtained from resources such as firewalls, routers and gateways, as well as system related data obtained from computer resources. Unconventional data, threat intelligence feeds, databases/repositories and expert opinion (see Section 5.1) are other popular resources of data. Note that the data source categories shown in Figure 7a are categories addressed by 20 or more primary studies. The remaining data source categories were addressed by few primary studies (less than 20) and therefore do not represent any significance compared to the counts for the categories shown in Figure 7a. In addition, note that several primary studies include more than one data source.

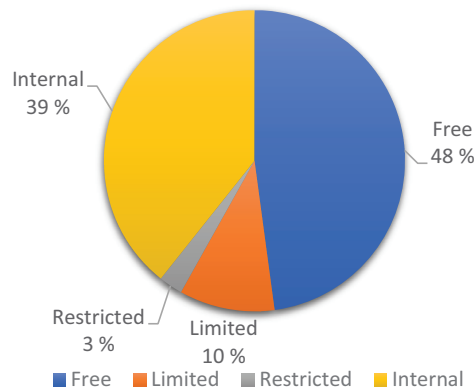
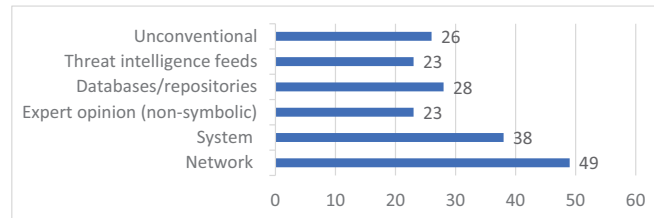
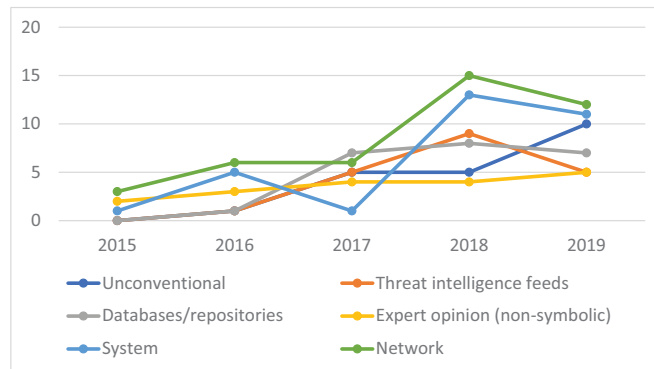


Figure 6. Data openness.



(a)



(b)

Figure 7. (a) Data source categories addressed by 20 or more primary studies; and (b) number of primary studies addressing data source categories in the period.

Figure 7b shows the trend for each category over time. We see that the number of papers addressing the categories *system* and *network* have increased the most since 2017, and we also see that the category *unconventional* has increased significantly since 2018.

We applied a similar strategy for presenting the mapping results as described above for the data type and data content categories. Figure 8a illustrates the data type categories addressed by 20 or more primary studies. In this case, we see a pattern of the three most popular groups of data type categories. Figure 8a shows that *structured* and *historical data* are the most popular data type categories, followed by *unstructured*, *combined* and *real-time data* in a shared second place, and finally *training sets* and *estimations* in a shared third place. In terms of the trend for each category over time, Figure 8b shows that *structured* and *historical data* are also the categories that have been increasing the most. Moreover, the categories *unstructured* and *training sets* have increased significantly since 2018.



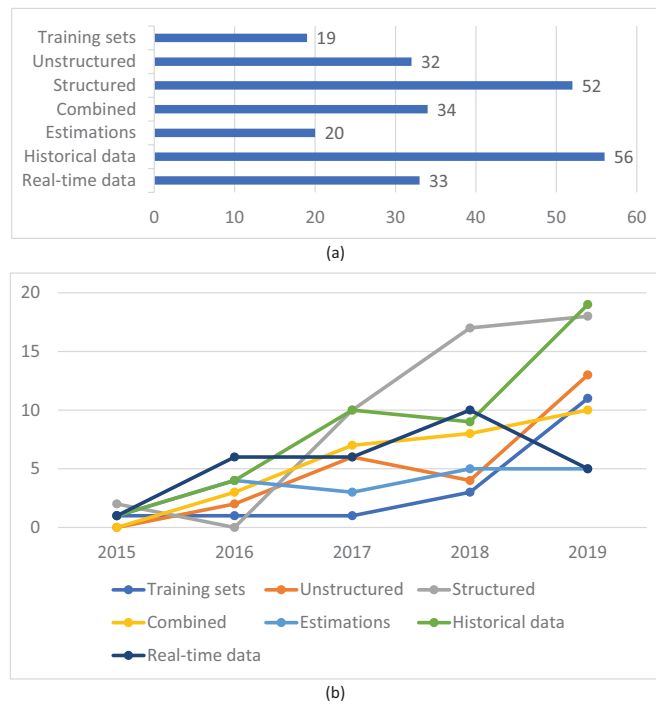


Figure 8. (a) Data type categories addressed by 20 or more primary studies; and (b) number of primary studies addressing data type categories in the period.

With respect to data content categories, Figure 9a shows that *network traffic event* is the dominating category, followed by *incident descriptions* and *vulnerabilities* in a shared second place, and finally *risk factors* and *IP-addresses* in a shared third place. As for data content categories (cf. Figure 9b), studies on network traffic events have had an increasing trend since 2015, while the remaining categories follow more or less a flat trend since 2015.

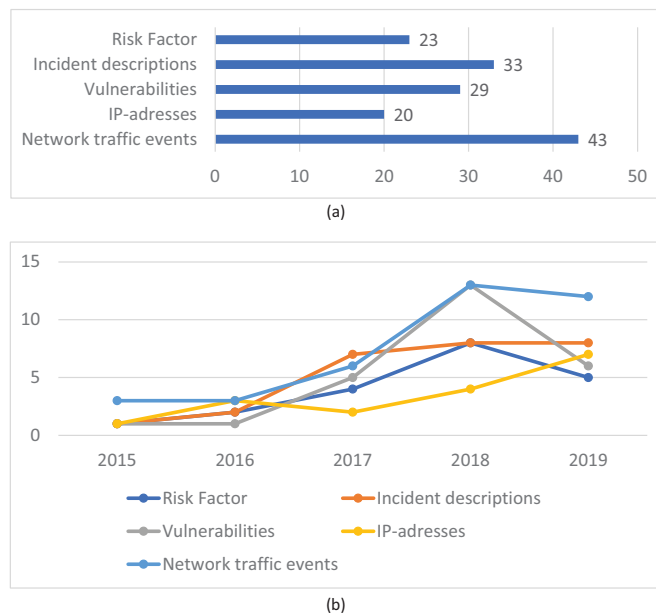


Figure 9. (a) Data content categories addressed by 20 or more primary studies; and (b) number of primary studies addressing data content categories in the period.

In summary, the observations in Figures 7–9 show that data sources are mainly from network resources such as firewalls, routers and gateways. The data types are mainly structured and historical data, and the data content is mainly related to network traffic events. In terms of trends for data sources, we see an increasing number of papers using system, network and unconventional data sources. Moreover, trends for data types show an increasing number of papers using structured, historical, unstructured and training set data. Finally, trends for data content show that network traffic events is the most increasing category.

Finally, we investigated the average number of data source, data type and data content categories that were considered by the primary studies within the reported period. This average trend will help us understand whether the number of categories used by the primary studies are increasing over time. As illustrated in Figure 10, the usage of data source categories is following a flat trend with the lowest average 1.7 in 2017 and 2019 and the highest average 2.0 in 2018. However, the usage of data type and data content categories are increasing following a linear trend. With respect to data type categories, the lowest average is 1.8 in both 2015 and 2016 and the highest average is 3.0 in 2019. With respect to data content categories, the lowest average is 1.8 in 2016 and the highest average is 3.1 in 2018. Thus, while using multiple data sources has not increased much over the years, the usage of multiple data types and data content is increasing following a linear trend.

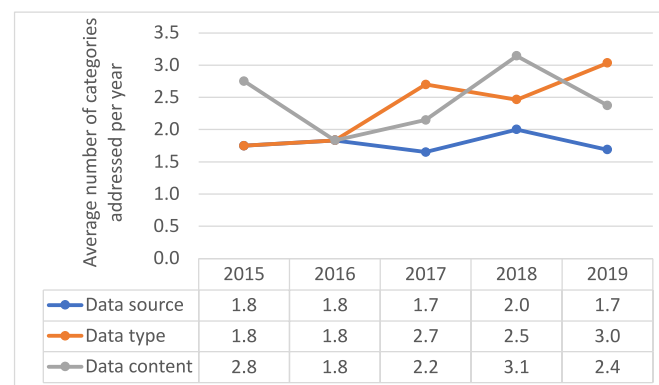


Figure 10. Average number of data source/data type/data content categories per year.

## 6. Discussion

In this section, we discuss our results with respect to the research questions. We compare our findings with previous work in order to find similarities, address our main limitations and recommend future research.

### 6.1. RQ 1: What Is the Nature of the Research Using Security Indicators?

As shown in Figure 5, the majority of the papers included in our systematic mapping study were validation research papers (87 out of 117). This is not surprising since, as pointed out by Wieringa et al. [34], the core business of engineering research is to propose new techniques and investigate their properties. However, this implies that most studies lack empirical evaluation with real-world application. It seems to be easier to publish methods and techniques on a conceptual level than to apply them in practice. This is in line with what Pendleton et al. found for security metrics [19], i.e. researchers often encounter a lack of real data for verification and validation.

### 6.2. RQ 2: What Is the Intended Use of the Data?

The results show that the selected studies are rather evenly distributed in the given data usage categories. In some studies, the data are used for more than one usage category; in such cases, we classified the paper by choosing the broader category. For example, for

technical as well as strategic usage, the study is classified for strategic use as it covers the technical usage. The usage patterns indicate an inclination towards using Technical (38) threat intelligence, which is followed by using Strategic (31), Operational (27) and Tactical (21) data. We consider it positive that the data are used at four levels for informed decision making. However, the studies are sparsely distributed in a wide range of usage domains, with approximately 72% of the selected studies, i.e., 84 of 117, not addressing a specific domain. The sparse distribution of studies within specific domains, mostly 1–2 studies per domain, indicates that research in tapping the potential of threat intelligence at various levels is still in its beginning stages. Chockalingam et al. [30] argued that domain-specific empirical data sources are needed to develop realistic models in cyber security. It can therefore be inferred that more research is needed in domain-specific data usage to contribute to utilizing comprehensive threat intelligence.

### 6.3. RQ 3: What Is the Origin of the Data for the Indicators?

Our results show that the two most popular data origins were from networks and systems. Unconventional data, threat intelligence feeds, databases/repositories and expert opinion were also quite commonly used (see Figure 7). We consider it positive that real-world data have been increasingly used in the last few years, in particular since the majority of earlier studies are not using real-world data. For example, related to digital forensics, Grajeda et al. [15] showed that the clear majority of datasets are experimentally generated (56.4%), with real-world user generated in second place (36.7%). Furthermore, Gheyas et al. [27] showed that only a small percentage of studies up until 2015 used original real-world data for the prediction of insider threats. Chockalingam et al. [30] also showed in 2017 that most Bayesian Network models used expert knowledge and/or data from the literature as their data sources.

An interesting observation regarding the origin of the data is that each of the primary studies used, on average, more than one data source for deriving their indicators (Figure 10). For example, the approach presented by Erdogan et al. [37] reports four data sources as input for cyber-risk assessment (network layer monitoring indicators, application layer monitoring indicators, security test results and business-related information obtained from stakeholders). While we did not record whether these previous studies have shared the datasets openly with others, the benefits of collecting and sharing such data are pointed out by Moore et al. [38] and Zheng et al. [16].

Close to half (48%) of the input data from the primary studies were free, meaning publicly available. That is somewhat lower than what Zheng et al. [16] registered (76%). This could be explained by the fact that many studies used more than one type of data source, and we classified these according to the strictest type (typically internal).

### 6.4. RQ 4: What Types of Data Are Being Used?

The trends related to data type indicate that the community is increasingly becoming better in taking advantage of structured and historical data in particular. Wagner et al. [39] showed a precipitously increasing research interest in cyber threat intelligence sharing up until 2016, followed by a slight decline in the following years. One could assume that this is due to improved maturity and uptake of standardized languages for sharing threat intelligence, such as Mitre's STIX [40]. However, studies by Ramsdale et al. [41] and Bromander et al. [42,43] show the contrary and that, in practice, threat intelligence providers are opting for custom or simple formats. We did not classify primary studies according to specific sharing standards or enumerations, and this could be a future extension to the scheme. Mavroeidis and Bromander [44] provided an overview of those already used for sharing threat intelligence. It is also outside of our analysis whether the increasing number of papers are using different data source instances or if they are using the same ones.

The results indicate a recent sharp growth in publications applying unstructured data. We believe this is directly related to the increased usage of unconventional data sources, such as social media. This is in accordance with findings by Husák et al.'s [21] in their

survey of prediction and forecasting methods in cyber security, showing recent approaches based on non-technical data from sentiment analysis on social networks or changes in user behavior.

#### 6.5. RQ 5: What Is the Data Content of the Indicators?

As mentioned in our results, network traffic dominates among the data content types, which conforms with the popular corresponding data source/origin (network) and data usage (technical) classifications. We also found that many of the primary studies did not really give precise information about what kind of network traffic they were using, which is partly the reason we find a high concentration here. For some primary studies, we could classify more precisely towards IP-addresses or DNS-data. In 2016, Pendleton et al. [19] recommended that security publications should explicitly specify their security metrics, but we did not find much evidence of this actually being done. Data about incidents and vulnerabilities also have a technical content, and, as Tounsi and Rais [9] pointed out, these are easy to quantify, share, standardize and determine immediate actions from. Although not directly comparable, Grajeda et al. [15], found utilization of datasets related to malware (signatures), network traffic and chat logs (attack planning and targets), but these were not dominating for forensics. Within the datasets catalogued by Zheng et al. [16], there were content related to vulnerabilities, exploits (incident descriptions), cybercrime activities (attack planning and targets), network traces (network traffic events), user activities (user behavior), alerts (intrusion detection alert) and configurations (countermeasures). Here, the technical content types dominated as well.

#### 6.6. Limitations and Recommendations for Future Research

While a systematic mapping study captures focus areas and trends within the literature, it does not dig into the details and quality of results from the primary studies. Hence, we cannot give any recommendations on which data and indicator types work better than others. That would require a more focused literature review, but it is our impression that the current literature does not contain appropriate and comparable parameters to make such benchmarks.

Due to the empirical nature of systematic mapping studies, threats to validity such as construct validity or internal validity are present. To mitigate threats to validity concerning selection, screening and classification of studies, we defined a detailed screening strategy and screening and classification process. In addition, we carried out a calibration exercise to address variances between researchers. To a considerable degree, the aforementioned measures confirm the validity of the search, screening and classification processes. We also acknowledge that relevant publications may have been overlooked due to missing search keywords, delayed indexing by search engines or human mistakes in the screening process. Despite actions taken to calibrate the participating researchers and reduce systematic errors, the mapping is based on subjective interpretations of paper contents. Due to limited resources, we did not have the opportunity to undertake double review of the complete set of full papers. However, we would argue that we included such a large body of primary studies that the mapping still shows an accurate and precise overall picture.

Our classification scheme is more detailed or has a different focus than what is seen in related work (e.g., Sun et al. [23], Grajeda et al. [15] and Zheng et al. [16]). It is also highly reusable and can be applied to capture new trends by doing a similar study in the future. Furthermore, it would be interesting to include more grey literature (e.g., technical reports, white papers, theses and web pages) to capture use of cyber security indicators that are not driven by academic research. According to Garousi et al. [45], such multivocal literature reviews can be valuable in closing the gap between academic research and practice. This kind of work would require more use of manual search and snowballing, which unfortunately is quite resource demanding.



## 7. Conclusions

We conducted a systematic mapping study on the use of cyber security indicator data in the academic literature to structure the research area. The number of publications has had a linear growth over the past five years, and the dominant approach is validation research based on free (public) or internally developed indicators. The usage patterns show a slight inclination towards technical threat intelligence, with little use of domain specific data. We can see a trend where data originating from network or system resources are increasing the most, followed by unconventional data, threat intelligence feeds, databases/repositories and expert opinion. On average, more than one data source is used to derive indicators in each paper. Our results show that the research community is eagerly developing new methods and techniques to support security decisions. However, many proposed techniques are on the conceptual level, with little or no empirical evaluation, thus may not yet be mature enough for real-world application. With indicators that are rather technical in nature, we can quickly share information about present security events, increase situational awareness and act accordingly. This allows contemporary cyber risk estimates to become more data-driven and less gut-driven. At the same time, such indicators tend to be short-lived. The increasing usage of unconventional data sources and threat intelligence feeds of more strategic and tactical nature represent a more forward-looking trend. We cannot really say whether or not we have become better at anticipating attacks, but at least it seems the research community is trying.

**Author Contributions:** Conceptualization, P.H.M., K.B. and A.O.; methodology, P.H.M., G.E. and A.O.; validation, P.H.M., S.T. and G.E.; investigation, P.H.M., S.T., G.E., K.B. and A.E.; resources, P.H.M. and S.T.; data curation, P.H.M., S.T., G.E., K.B. and A.E.; writing—original draft preparation, P.H.M., S.T., G.E., K.B. and A.E.; writing—review and editing, P.H.M., S.T., G.E., K.B. and A.E.; visualization, P.H.M., S.T. and G.E.; supervision, P.H.M.; and project administration, G.E. and A.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by CyberSec4Europe, which is funded by the European Union under the H2020 Programme Grant Agreement No. 830929.

**Data Availability Statement:** The data presented in this study are available at <https://doi.org/10.5281/zenodo.4639585>.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Search String Definitions

For all databases, we tried to create as equivalent searches as possible. However, we had to consider differences in features and functionality. The sections below show how we implemented the queries for each of the databases.

### Appendix A.1. IEEE Xplore

The *Command Search* feature of this database allows query strings consisting of data fields and operators (in caps). We also applied a filter to limit the result to publications including and between 2015 and 2020. The following search string was applied:

```
(("Document Title":"cyber security" OR
title:"information security" OR
title:"cyber risk" OR
title:"cyber threat" OR
title:"threat intelligence"
OR title:"cyber attack") AND
("All Metadata":"predict" OR
Search_All:"strategic" OR
Search_All:"tactical" OR
Search_All:"likelihood" OR
Search_All:"probability" OR
Search_All:"metric" OR
```

```
Search_All:"indicator"))
```

#### Appendix A.2. Science Direct

We made use of the search form instead of a query string for this database. The advanced search feature allowed us to specify keywords for the *title* and another set for the *title*, *abstract* and *author-specified keyword*. However, the space between keywords implicitly meant an AND-operator, while what we really needed was OR. This meant that we had to submit 42 search forms, one for each primary keyword for the title in combination with every secondary keyword for the range 2015–2020.

#### Appendix A.3. ACM Digital Library

This database allowed searching for specific keywords in title, abstract and author specified keywords. The following search string was applied:

```
[[Publication Title: "cyber security"] OR
[Publication Title: "information security"] OR
[Publication Title: "cyber risk"] OR
[Publication Title: "cyber threat"] OR
[Publication Title: "threat intelligence"] OR
[Publication Title: "cyber attack"]] AND
[[Abstract: predict] OR [Abstract: strategic] OR
[Abstract: tactical] OR [Abstract: likelihood] OR
[Abstract: probability] OR [Abstract: metric] OR
[Abstract: indicator]] AND
[Publication Date: (01/01/2015 TO 12/31/2020)]
```

#### Appendix A.4. SpringerLink

We employed a form-based (advanced) search. The title search did not allow for operators, hence we had to submit six search forms, one for each primary keyword and where at least one of the secondary keywords appeared somewhere. There was no option to search within just the abstract or author defined keywords, hence the result set became large, and we had to use the stopping criteria (results sorted by relevance, stop after 10 irrelevant in a row). The date range was set to 2015–2020.

#### Appendix A.5. Google Scholar

The advanced features of this search engine allowed for specifying title keywords, with additional ones using | as an OR operator. It was important to turn off personalized search results (turn off “signed-in search activity”) so that different researchers would get the same results. If not, the results would have been influenced by their previous search history. We specifically excluded patents and citations and defined the date range 2015–2020. The following search string was applied:

```
allintitle: ("cyber security" |
"information security"| "cyber risk" |
"cyber threat"| "threat intelligence" |
"cyber attack") (Predict | strategic |
tactical | likelihood | probability |
metric | indicator)
```

### Appendix B. The Selected Primary Studies

- Kolosok, Irina and Liudmila Gurina (2014). “Calculation of cyber security index in the problem of power system state estimation based on SCADA and WAMS measurements”. In: International Conference on Critical Information Infrastructures Security. Springer, pp. 172–177.
- Liu, Yang et al. (2015). “Predicting cyber security incidents using feature-based characterization of network-level malicious activities”. In: Proceedings of the 2015



ACM International Workshop on International Workshop on Security and Privacy Analytics, pp. 3–9.

- Llansó, Thomas, Anurag Dwivedi and Michael Smeltzer (2015). “An approach for estimating cyber attack level of effort”. In: 2015 Annual IEEE Systems Conference (SysCon) Proceedings. IEEE, pp. 14–19.
- Shin, Jinsoo, Hanseong Son, Gyunyoung Heo, et al. (2015). “Development of a cyber security risk model using Bayesian networks”. In: Reliability Engineering & System Safety 134, pp. 208–217.
- Zhan, Zhenxin, Maochao Xu and Shouhuai Xu (2015). “Predicting cyber attack rates with extreme values”. In: IEEE Transactions on Information Forensics and Security 10.8, pp. 1666–1677.
- Atighetchi, Michael et al. (2016). “Experimentation support for cyber security evaluations”. In: Proceedings of the 11th Annual Cyber and Information Security Research Conference, pp. 1–7.
- Aziz, Benjamin, Ali Malik and Jeyong Jung (2016). “Check your blind spot: a new cyber-security metric for measuring incident response readiness”. In: International Workshop on Risk Assessment and Risk-driven Testing. Springer, pp. 19–33.
- Chhetri, Sujit Rokka, Arquimedes Canedo and Mohammad Abdullah Al Faruque (2016). “Kcad: kinetic cyber-attack detection method for cyber-physical additive manufacturing systems”. In: 2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD). IEEE, pp. 1–8.
- Dog, Spike E et al. (2016). “Strategic cyber threat intelligence sharing: A case study of ids logs”. In: 2016 25th International Conference on Computer Communication and Networks (ICCCN). IEEE, pp. 1–6.
- Hamid, T et al. (2016). “Cyber security risk evaluation research based on entropy weight method”. In: 2016 9th International Conference on Developments in eSystems Engineering (DeSE). IEEE, pp. 98–104.
- Je, Young-Man, Yen-Yoo You and Kwan-Sik Na (2016). “Information security evaluation using multi-attribute threat index”. In: Wireless Personal Communications 89.3, pp. 913–925.
- Liao, Xiaojing et al. (2016). “Acing the ioc game: Toward automatic discovery and analysis of open-source cyber threat intelligence”. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, pp. 755–766.
- Noble, Jordan and Niall M Adams (2016). “Correlation-based streaming anomaly detection in cyber-security”. In: 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW). IEEE Computer Society, pp. 311–318.
- Singh, Umesh Kumar and Chanchala Joshi (2016). “Network security risk level estimation tool for information security measure”. In: 2016 IEEE 7th Power India International Conference (PIICON). IEEE, pp. 1–6.
- Wagner, Cynthia et al. (2016). “Misp: The design and implementation of a collaborative threat intelligence sharing platform”. In: Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security, pp. 49–56.
- Wang, Jiao et al. (2016). “A method for information security risk assessment based on the dynamic bayesian network”. In: 2016 International Conference on Networking and Network Applications (NaNA). IEEE, pp. 279–283.
- Wangen, Gaute and Andrii Shalaginov (2016). “Quantitative risk, statistical methods and the four quadrants for information security”. In: International Conference on Risks and Security of Internet and Systems. Springer, pp. 127–143.
- Ahrend, Jan M and Marina Jirotko (2017). “Anticipation in Cyber-Security”. In: Handbook of Anticipation. Springer, Cham, pp. 1–28.
- Aksu, M Ugur et al. (2017). “A quantitative CVSS-based cyber security risk assessment methodology for IT systems”. In: 2017 International Carnahan Conference on Security Technology (ICCST). IEEE, pp. 1–8.

- AlEroud, Ahmed and Izzat Alsmadi (2017). "Identifying cyber-attacks on software defined networks: An inference-based intrusion detection approach". In: *Journal of Network and Computer Applications* 80, pp. 152–164.
- Andress, J et al. (2017). "Chapter 10–Information Security Program Metrics". In: *Building a Practical Information Security Program*, pp. 169–183.
- Bernsmed, Karin et al. (2017). "Visualizing cyber security risks with bow-tie diagrams". In: *International Workshop on Graphical Models for Security*. Springer, pp. 38–56.
- Best, Daniel M et al. (2017). "Improved cyber threat indicator sharing by scoring privacy risk". In: *2017 IEEE International Symposium on Technologies for Homeland Security (HST)*. IEEE, pp. 1–5.
- Černivec, Aleš et al. (2017). "Employing Graphical Risk Models to Facilitate Cyber-Risk Monitoring—the WISER Approach". In: *International Workshop on Graphical Models for Security*. Springer, pp. 127–146.
- Cheng, Ran, Yueming Lu and Jiefu Gan (2017). "Environment-Related Information Security Evaluation for Intrusion Detection Systems". In: *International Conference on Communicatins and Networking in China*. Springer, pp. 373–382.
- Dalton, Adam et al. (2017). "Improving cyber-attack predictions through information foraging". In: *2017 IEEE International Conference on Big Data (Big Data)*. IEEE, pp. 4642–4647.
- Doynikova, Elena and Igor Kotenko (2017). "Enhancement of probabilistic attack graphs for accurate cyber security monitoring". In: *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*. IEEE, pp. 1–6.
- Kandias, Miltiadis et al. (2017). "Stress level detection via OSN usage pattern and chronicity analysis: An OSINT threat intelligence module". In: *Computers & Security* 69, pp. 3–17.
- Khandpur, Rupinder Paul et al. (2017). "Crowdsourcing cybersecurity: Cyber attack detection using social media". In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pp. 1049–1057.
- Lee, Kuo-Chan et al. (2017). "Sec-Buzzer: cyber security emerging topic mining with open threat intelligence retrieval and timeline event annotation". In: *Soft Computing* 21.11, pp. 2883–2896.
- Liu, Ruyue et al. (2017). "A Research and Analysis Method of Open Source Threat Intelligence Data". In: *International Conference of Pioneering Computer Scientists, Engineers and Educators*. Springer, pp. 352–363.
- Polatidis, Nikolaos, Elias Pimenidis, Michalis Pavlidis and Haralambos Mouratidis (2017). "Recommender systems meeting security: From product recommendation to cyber- attack prediction". In: *International Conference on Engineering Applications of Neural Networks*. Springer, pp. 508–519.
- Price-Williams, Matthew, Nick Heard and Melissa Turcotte (2017). "Detecting periodic subsequences in cyber security data". In: *2017 European Intelligence and Security Informatics Conference (EISIC)*. IEEE, pp. 84–90.
- Qamar, Sara et al. (2017). "Data-driven analytics for cyber- threat intelligence and information sharing". In: *Computers & Security* 67, pp. 35–58.
- Ślezak, Dominik et al. (2017). "Scalable cyber-security analytics with a new summary-based approximate query engine". In: *2017 IEEE International Conference on Big Data (Big Data)*. IEEE, pp. 1840–1849.
- Stine, Ian et al. (2017). "A cyber risk scoring system for medical devices". In: *International Journal of Critical Infrastructure Protection* 19, pp. 32–46.
- Teoh, TT et al. (2017). "Analyst intuition based Hidden Markov Model on high speed, temporal cyber security big data". In: *2017 13th International Conference on*





- Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD). IEEE, pp. 2080–2083.
- Wagner, Thomas D et al. (2017). “Towards an Anonymity Supported Platform for Shared Cyber Threat Intelligence”. In: International Conference on Risks and Security of Internet and Systems. Springer, pp. 175–183.
  - Yaseen, Amer Atta and Mireille Bayart (2017). “Cyber-attack detection with fault accommodation based on intelligent generalized predictive control”. In: IFAC-PapersOnLine 50.1, pp. 2601–2608.
  - Aditya, K, Slawomir Grzonkowski and Nhien-An Le-Khac (2018). “Riskwriter: Predicting cyber risk of an enterprise”. In: International Conference on Information Systems Security. Springer, pp. 88–106.
  - Almohannadi, Hamad et al. (2018). “Cyber threat intelligence from honeypot data using elasticsearch”. In: 2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA). IEEE, pp. 900–906.
  - Araujo, Frederico et al. (2018). “Cross-Stack Threat Sensing for Cyber Security and Resilience”. In: 2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W). IEEE, pp. 18–21.
  - Barboni, Angelo, Francesca Boem and Thomas Parisini (2018). “Model-based detection of cyber-attacks in networked MPC-based control systems”. In: IFAC-PapersOnLine 51.24, pp. 963–968.
  - Böhm, Fabian, Florian Menges and Günther Pernul (2018). “Graph-based visual analytics for cyber threat intelligence”. In: Cybersecurity 1.1, p. 16.
  - Cho, Hyeisun et al. (2018). “Method of Quantification of Cyber Threat Based on Indicator of Compromise”. In: 2018 International Conference on Platform Technology and Service (PlatCon). IEEE, pp. 1–6.
  - Ghazi, Yumna et al. (2018). “A supervised machine learning based approach for automatically extracting high-level threat intelligence from unstructured sources”. In: 2018 International Conference on Frontiers of Information Technology (FIT). IEEE, pp. 129–134.
  - Gokaraju, Balakrishna et al. (2018). “Identification of spatio-temporal patterns in cyber security for detecting the signature identity of hacker”. In: SoutheastCon 2018. IEEE, pp. 1–5.
  - Gonzalez-Granadillo, G et al. (2018). “Dynamic risk management response system to handle cyber threats”. In: Future Generation Computer Systems 83, pp. 535–552.
  - Gschwandtner, Mathias et al. (2018). “Integrating threat intelligence to enhance an organization’s information security management”. In: Proceedings of the 13th International Conference on Availability, Reliability and Security, pp. 1–8.
  - Guerrero-Higueras, Ángel Manuel, Noemi DeCastro-Garcia and Vicente Matellan (2018). “Detection of Cyber-attacks to indoor real time localization systems for autonomous robots”. In: Robotics and Autonomous Systems 99, pp. 75–83.
  - Haughey, Hamish et al. (2018). “Adaptive traffic fingerprinting for darknet threat intelligence”. In: Cyber Threat Intelligence. Springer, pp. 193–217.
  - Iqbal, Zafar, Zahid Anwar and Rafia Mumtaz (2018). “STIXGEN-A Novel Framework for Automatic Generation of Structured Cyber Threat Information”. In: 2018 International Conference on Frontiers of Information Technology (FIT). IEEE, pp. 241–246.
  - Kim, Eunsoo et al. (2018). “CyTIME: Cyber Threat Intelligence Management framework for automatically generating security rules”. In: Proceedings of the 13th International Conference on Future Internet Technologies, pp. 1–5.
  - Kim, Nakhyun et al. (2018). “Study of Natural Language Processing for Collecting Cyber Threat Intelligence Using SyntaxNet”. In: International Symposium of Information and Internet Technology. Springer, pp. 10–18.
  - Kottenko, Igor et al. (2018). “AI-and metrics-based vulnerability-centric cyber security assessment and countermeasure selection”. In: Guide to Vulnerability Analysis for Computer Networks and Systems. Springer, pp. 101–130.

- Lee, Chanyoung, Ho Bin Yim and Poong Hyun Seong (2018). "Development of a quantitative method for evaluating the efficacy of cyber security controls in NPPs based on intrusion tolerant concept". In: *Annals of Nuclear Energy* 112, pp. 646–654.
- Moskal, Stephen, Shanchieh Jay Yang and Michael E Kuhl (2018). "Extracting and evaluating similar and unique cyber attack strategies from intrusion alerts". In: 2018 IEEE International Conference on Intelligence and Security Informatics (ISI). IEEE, pp. 49–54.
- Pitropakis, Nikolaos et al. (2018). "An enhanced cyber attack attribution framework". In: *International Conference on Trust and Privacy in Digital Business*. Springer, pp. 213–228.
- Prabhu, Vinayak et al. (2018). "Towards Data-Driven Cyber Attack Damage and Vulnerability Estimation for Manufacturing Enterprises". In: *International Conference on Remote Engineering and Virtual Instrumentation*. Springer, pp. 333–343.
- Radanliev, Petar et al. (2018). "Economic impact of IoT cyber risk-analysing past and present to predict the future developments in IoT risk analysis and IoT cyber insurance". In:
- Shu, Kai et al. (2018). "Understanding cyber attack behaviors with sentiment information on social media". In: *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*. Springer, pp. 377–388.
- Smith, Matthew David and M Elisabeth Paté-Cornell (2018). "Cyber risk analysis for a smart grid: how smart is smart enough? a multiarmed bandit approach to cyber security investment". In: *IEEE Transactions on Engineering Management* 65.3, pp. 434–447.
- Vinayakumar, R, Prabaharan Poornachandran and KP Soman (2018). "Scalable framework for cyber threat situational awareness based on domain name systems data analysis". In: *Big data in engineering applications*. Springer, pp. 113–142.
- Wang, Junshe et al. (2018). "Network attack prediction method based on threat intelligence". In: *International Conference on Cloud Computing and Security*. Springer, pp. 151–160.
- Zieger, Andrej, Felix Freiling and Klaus-Peter Kossakowski (2018). "The  $\beta$ -time-to-compromise metric for practical cyber security risk estimation". In: 2018 11th International Conference on IT Security Incident Management & IT Forensics (IMF). IEEE, pp. 115–133.
- Bo, Tao et al. (2019). "TOM: A Threat Operating Model for Early Warning of Cyber Security Threats". In: *International Conference on Advanced Data Mining and Applications*. Springer, pp. 696–711.
- Doynikova, Elena, Andrey Fedorchenko and Igor Kotenko (2019). "Ontology of metrics for cyber security assessment". In: *Proceedings of the 14th International Conference on Availability, Reliability and Security*, pp. 1–8.
- Dragos, Valentina et al. (2019). "Entropy-Based Metrics for URREF Criteria to Assess Uncertainty in Bayesian Networks for Cyber Threat Detection". In: 2019 22th International Conference on Information Fusion (FUSION). IEEE, pp. 1–8.
- Gautam, Apurv Singh, Yamini Gahlot and Pooja Kamat (2019). "Hacker Forum Exploit and Classification for Proactive Cyber Threat Intelligence". In: *International Conference on Inventive Computation Technologies*. Springer, pp. 279–285.
- Kannavara, Raghudeep et al. (2019). "A threat intelligence tool for the security development lifecycle". In: *Proceedings of the 12th Innovations on Software Engineering Conference (formerly known as India Software Engineering Conference)*, pp. 1–5.
- Keim, Yansi and AK Mohapatra (2019). "Cyber threat intelligence framework using advanced malware forensics". In: *International Journal of Information Technology*, pp. 1–10.



- Al-khateeb, Samer and Nitin Agarwal (2019). "Social cyber forensics: leveraging open source information and social network analysis to advance cyber security informatics". In: Computational and Mathematical Organization Theory, pp. 1–19.
- Li, Yi-Fan et al. (2019). "Multistream classification for cyber threat data with heterogeneous feature space". In: The World Wide Web Conference, pp. 2992–2998.
- Marukhlenko, AL, AV Plugatarev and DO Bobyntsev (2019). "Complex Evaluation of Information Security of an Object with the Application of a Mathematical Model for Calculation of Risk Indicators". In: International Russian Automation Conference. Springer, pp. 771–778.
- Merino, Tim et al. (2019). "Expansion of cyber attack data from unbalanced datasets using generative adversarial networks". In: International Conference on Software Engineering Research, Management and Applications. Springer, pp. 131–145.
- Milajerdi, Sadegh M et al. (2019). "Poirot: Aligning attack behavior with kernel audit records for cyber threat hunting". In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, pp. 1795–1812.
- Milošević, Jezdimir, Henrik Sandberg and Karl Henrik Johansson (2019). "Estimating the impact of cyber-attack strategies for stochastic networked control systems". In: IEEE Transactions on Control of Network Systems 7.2, pp. 747–757
- Mokaddem, Sami et al. (2019). "Taxonomy driven indicator scoring in MISP threat intelligence platforms". In: arXiv preprint arXiv:1902.03914.
- Mukhopadhyay, Arunabha et al. (2019). "Cyber risk assessment and mitigation (CRAM) framework using logit and probit models for cyber insurance". In: Information Systems Frontiers 21.5, pp. 997–1018.
- Noor, Umara et al. (2019). "A machine learning framework for investigating data breaches based on semantic analysis of adversary's attack patterns in threat intelligence repositories". In: Future Generation Computer Systems 95, pp. 467–487.
- Okutan, Ahmet and Shanchieh Jay Yang (2019). "ASSERT: attack synthesis and separation with entropy redistribution towards predictive cyber defense". In: Cybersecurity 2.1, pp. 1–18.
- Papastergiou, Spyridon, Haralambos Mouratidis and Eleni-Maria Kalogeraki (2019). "Cyber security incident handling, warning and response system for the european critical information infrastructures (cybersane)". In: International Conference on Engineering Applications of Neural Networks. Springer, pp. 476–487.
- Pour, Morteza Safaei et al. (2019). "Comprehending the IoT cyber threat landscape: A data dimensionality reduction technique to infer and characterize Internet-scale IoT probing campaigns". In: Digital Investigation 28, S40–S49.
- Riesco, Raúl and Víctor A Villagrà (2019). "Leveraging cyber threat intelligence for a dynamic risk framework". In: International Journal of Information Security 18.6, pp. 715–739.
- Rijswijk-Deij, Roland van et al. (2019). "Privacy-conscious threat intelligence using DNSBloom". In: 2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM). IEEE, pp. 98–106.
- Simran, K et al. (2019). "Deep Learning Approach for Enhanced Cyber Threat Indicators in Twitter Stream". In: International Symposium on Security in Computing and Communication. Springer, pp. 135–145.
- Sliva, Amy, Kai Shu and Huan Liu (2019). "Using social media to understand cyber attack behavior". In: International Conference on Applied Human Factors and Ergonomics. Springer, pp. 636–645.
- Subroto, Athor and Andri Apriyana (2019). "Cyber risk prediction through social media big data analytics and statistical machine learning". In: Journal of Big Data 6.1, pp. 1–19.
- Tonn, Gina et al. (2019). "Cyber risk and insurance for transportation infrastructure". In: Transport policy 79, pp. 103–114.

- Trivedi, Tarun et al. (2019). "Threat Intelligence Analysis of Onion Websites Using Sublinks and Keywords". In: *Emerging Technologies in Data Mining and Information Security*. Springer, pp. 567–578.
- Ullah, Sharif et al. (2019). "Cyber Threat Analysis Based on Characterizing Adversarial Behavior for Energy Delivery System". In: *International Conference on Security and Privacy in Communication Systems*. Springer, pp. 146–160.
- Ustebay, Serpil, Zeynep Turgut and M Ali Aydin (2019). "Cyber Attack Detection by Using Neural Network Approaches: Shallow Neural Network, Deep Neural Network and AutoEncoder". In: *International Conference on Computer Networks*. Springer, pp. 144–155.
- Vielberth, Manfred, Florian Menges and Günther Pernul (2019). "Human-as-a-security-sensor for harvesting threat intelligence". In: *Cybersecurity 2.1*, pp. 1–15.
- Vinayakumar, R, KP Soman, et al. (2019). "Deep learning framework for cyber threat situational awareness based on email and url data analysis". In: *Cybersecurity and Secure Information Systems*. Springer, pp. 87–124.
- Wang, Huaizhi et al. (2019). "Deep learning aided interval state prediction for improving cyber security in energy internet". In: *Energy 174*, pp. 1292–1304.
- Wangen, Gaute (2019). "Quantifying and Analyzing Information Security Risk from Incident Data". In: *International Workshop on Graphical Models for Security*. Springer, pp. 129–154.
- Yang, Wenzhuo and Kwok-Yan Lam (2019). "Automated Cyber Threat Intelligence Reports Classification for Early Warning of Cyber Attacks in Next Generation SOC". In: *International Conference on Information and Communications Security*. Springer, pp. 145–164.
- Zhang, Hongbin et al. (2019). "Network attack prediction method based on threat intelligence for IoT". In: *Multimedia Tools and Applications 78.21*, pp. 30257–30270.
- Almukaynizi, Mohammed et al. (2020). "A Logic Programming Approach to Predict Enterprise-Targeted Cyberattacks". In: *Data Science in Cybersecurity and Cyberthreat Intelligence*. Springer, pp. 13–32.
- Barrère, Martin et al. (2020). "Measuring cyber-physical security in industrial control systems via minimum-effort attack strategies". In: *Journal of Information Security and Applications 52*, p. 102471.
- Chen, Scarlett, Zhe Wu and Panagiotis D Christofides (2020). "Cyber-attack detection and resilient operation of nonlinear processes under economic model predictive control". In: *Computers & Chemical Engineering*, p. 106806.
- Evangelou, Marina and Niall M Adams (2020). "An anomaly detection framework for cyber-security data". In: *Computers & Security 97*, p. 101941.
- Facchinetti, Silvia, Paolo Giudici and Silvia Angela Osmetti (2020). "Cyber risk measurement with ordinal data". In: *Statistical Methods & Applications 29.1*, pp. 173–185.
- Figueira, Pedro Tubio, Cristina López Bravo and José Luis Rivas López (2020). "Improving information security risk analysis by including threat-occurrence predictive models". In: *Computers & Security 88*, p. 101609.
- Huang, Linan and Quanyan Zhu (2020). "A dynamic games approach to proactive defense strategies against advanced persistent threats in cyber-physical systems". In: *Computers & Security 89*, p. 101660.
- Khosravi, Mehran and Behrouz Tork Ladani (2020). "Alerts Correlation and Causal Analysis for APT Based Cyber Attack Detection". In: *IEEE Access 8*, pp. 162642–162656.
- Kour, Ravdeep, Adithya Thaduri and Ramin Karim (2020). "Predictive model for multistage cyber-attack simulation". In: *International Journal of System Assurance Engineering and Management 11.3*, pp. 600–613.
- Krisper, Michael, Jürgen Dobaj and Georg Macher (2020). "Assessing Risk Estimations for Cyber-Security Using Expert Judgment". In: *European Conference on Software Process Improvement*. Springer, pp. 120–134.

- Liao, Yi-Ching (2020). “Quantitative Information Security Vulnerability Assessment for Norwegian Critical Infrastructure”. In: International Conference on Critical Information Infrastructures Security. Springer, pp. 31–43.
- Luh, Robert and Sebastian Schrittwieser (2020). “Advanced threat intelligence: detection and classification of anomalous behavior in system processes”. In: *e & i Elektrotechnik und Informationstechnik* 137.1, pp. 38–44.
- Marin, Ericsson, Mohammed Almukaynizi and Paulo Shakarian (2020). “Inductive and deductive reasoning to assist in cyber-attack prediction”. In: 2020 10th Annual Computing and Communication Workshop and Conference (CCWC). IEEE, pp. 0262–0268.
- Al-Mohannadi, Hamad, Irfan Awan and Jassim Al Hamar (2020). “Analysis of adversary activities using cloud-based web services to enhance cyber threat intelligence”. In: Service Oriented Computing and Applications, pp. 1–13.
- Mohasseb, Alaa et al. (2020). “Cyber security incidents analysis and classification in a case study of Korean enterprises”. In: Knowledge and Information Systems.
- Polatidis, Nikolaos, Elias Pimenidis, Michalis Pavlidis, Spyridon Papastergiou, et al. (2020). “From product recommendation to cyber-attack prediction: generating attack graphs and predicting future attacks”. In: *Evolving Systems*, pp. 1–12.
- Rahman, Md Anisur, Yeslam Al-Saggaf and Tanveer Zia (2020). “A Data Mining Framework to Predict Cyber Attack for Cyber Security”. In: The 15th IEEE Conference on Industrial Electronics and Applications (ICIEA2020). IEEE Xplore.
- Tundis, Andrea, Samuel Ruppert and Max Mühlhäuser (2020). “On the Automated Assessment of Open-Source Cyber Threat Intelligence Sources”. In: International Conference on Computational Science. Springer, pp. 453–467.
- Uyheng, Joshua et al. (2020). “Interoperable pipelines for social cyber-security: Assessing Twitter information Operations during NATO Trident Juncture 2018”. In: *Computational and Mathematical Organization Theory* 26.4, pp. 465–483.
- Zhao, Jun et al. (2020). “TIMiner: Automatically Extracting and Analyzing Categorized Cyber Threat Intelligence from Social Data”. In: *Computers & Security*, p. 101867.

## References

1. Madnick, S. How Do You Prepare for the Unexpected Cyber Attack? *SSRN Electron. J.* **2020**. [CrossRef]
2. Anderson, R.; Böhme, R.; Clayton, R.; Moore, T. Security Economics and the Internal Market. Available online: <https://www.enisa.europa.eu/publications/archive/economics-sec/> (accessed on 23 March 2021).
3. Kitchenham, B.; Charters, S. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*; Technical Report EBSE-2007-01, Joint Report; Keele University: Keele, UK; University of Durham: Durham, UK, 2007.
4. Petersen, K.; Vakkalanka, S.; Kuzniarz, L. Guidelines for Conducting Systematic Mapping Studies in Software Engineering: An Update. *Inf. Softw. Technol.* **2015**, *64*, 1–18. [CrossRef]
5. Lea, D.; Bradbery, J. Oxford Advanced Learner’s Dictionary. 2021. Available online: <https://www.oxfordlearnersdictionaries.com/definition/english/indicator> (accessed on 22 April 2021).
6. Pfleeger, S.L.; Caputo, D.D. Leveraging Behavioral Science to Mitigate Cyber Security Risk. *Comput. Secur.* **2012**, *31*, 597–611. [CrossRef]
7. Brown, S.; Gommers, J.; Serrano, O. From Cyber Security Information Sharing to Threat Management. *WISCS '15: Proceedings of the 2nd ACM Workshop on Information Sharing and Collaborative Security*; Association for Computing Machinery: New York, NY, USA, 2015; pp. 43–49.
8. McMillan, R. Definition: Threat Intelligence. Available online: [https://www.gartner.com/imagesrv/media-products/pdf/webroot/issue1\\_webroot.pdf](https://www.gartner.com/imagesrv/media-products/pdf/webroot/issue1_webroot.pdf) (accessed on 26 March 2021).
9. Tounsi, W.; Rais, H. A Survey on Technical Threat Intelligence in the Age of Sophisticated Cyber Attacks. *Comput. Secur.* **2018**, *72*, 212–233. [CrossRef]
10. Chismon, D.; Ruks, M. Threat Intelligence: Collecting, Analysing, Evaluating. Available online: <https://informationsecurityreport/whitepapers/threat-intelligence-collecting-analysing-evaluating/10> (accessed on 26 March 2021).
11. Mateski, M.; Trevino, C.M.; Veitch, C.K.; Michalski, J.; Harris, J.M.; Maruoka, S.; Frye, J. Cyber Threat Metrics. Available online: <https://fas.org/irp/eprint/metrics.pdf> (accessed on 26 March 2021).
12. Wang, A.J.A. Information Security Models and Metrics. In Proceedings of the 43rd Annual Southeast Regional Conference, (ACM-SE 43), Kennesaw, GA, 18–20 March 2005; pp. 178–184.

13. Herrmann, D.S. *Complete Guide to Security and Privacy Metrics: Measuring Regulatory Compliance, Operational Resilience, and ROI*, 1st ed.; Auerbach Publications: Boston, MA, USA, 2007.
14. Humayun, M.; Niazi, M.; Jhanjhi, N.Z.; Alshayeb, M.; Mahmood, S. Cyber Security Threats and Vulnerabilities: A Systematic Mapping Study. *Arab. J. Sci. Eng.* **2020**, *45*, 3171–3189. [CrossRef]
15. Grajeda, C.; Breitingner, F.; Baggili, I. Availability of Datasets for Digital Forensics—And What is Missing. *Digit. Investig.* **2017**, *22*, S94–S105. [CrossRef]
16. Zheng, M.; Robbins, H.; Chai, Z.; Thapa, P.; Moore, T. Cybersecurity Research Datasets: Taxonomy and Empirical Analysis. In Proceedings of the 11th USENIX Workshop on Cyber Security Experimentation and Test (CSET'18), Baltimore, MD, USA, 13 August 2018.
17. Griffioen, H.; Booi, T.; Doerr, C. Quality Evaluation of Cyber Threat Intelligence Feeds. In Proceedings of the 18th International Conference on Applied Cryptography and Network Security (ACNS'20), Rome, Italy, 19–22 October 2020; pp. 277–296.
18. Tundis, A.; Ruppert, S.; Mühlhäuser, M. On the Automated Assessment of Open-Source Cyber Threat Intelligence Sources. In Proceedings of the 20th International Conference on Computational Science (ICCS'20), Amsterdam, The Netherlands, 3–5 June 2020; pp. 453–467.
19. Pendleton, M.; Garcia-Lebron, R.; Cho, J.H.; Xu, S. A Survey on Systems Security Metrics. *ACM Comput. Surv. CSUR* **2016**, *49*, 1–35. [CrossRef]
20. Cadena, A.; Gualoto, F.; Fuertes, W.; Tello-Oquendo, L.; Andrade, R.; Tapia Leon, F.; Torres, J. Metrics and Indicators of Information Security Incident Management: A Systematic Mapping Study. In *Smart Innovation, Systems and Technologies*; Springer Nature Singapore Private Limited: Singapore, Singapore, 2020; pp. 507–519. [CrossRef]
21. Husák, M.; Komárková, J.; Bou-Harb, E.; Čeleda, P. Survey of Attack Projection, Prediction, and Forecasting in Cyber Security. *IEEE Commun. Surv. Tutor.* **2018**, *21*, 640–660. [CrossRef]
22. Sriavstava, R.; Singh, P.; Chhabra, H. Review on Cyber Security Intrusion Detection: Using Methods of Machine Learning and Data Mining. In *Internet of Things and Big Data Applications: Recent Advances and Challenges*; Springer: Cham, Switzerland, 2020; pp. 121–132. [CrossRef]
23. Sun, N.; Zhang, J.; Rimba, P.; Gao, S.; Zhang, L.Y.; Xiang, Y. Data-Driven Cybersecurity Incident Prediction: A Survey. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 1744–1772. [CrossRef]
24. Laube, S.; Böhme, R. Strategic Aspects of Cyber Risk Information Sharing. *ACM Comput. Surv. CSUR* **2017**, *50*, 1–36. [CrossRef]
25. Diesch, R.; Krcmar, H. SoK: Linking Information Security Metrics to Management Success Factors. In Proceedings of the 15th International Conference on Availability, Reliability and Security (ARES'20), Dublin, Ireland, 25–28 August 2020; pp. 1–10.
26. Kotenko, I.; Doynikova, E.; Chechulin, A.; Fedorchenko, A., AI- and Metrics-Based Vulnerability-Centric Cyber Security Assessment and Countermeasure Selection. In *Guide to Vulnerability Analysis for Computer Networks and Systems: An Artificial Intelligence Approach*; Springer: Cham, Switzerland, 2018; pp. 101–130. [CrossRef]
27. Gheyas, I.A.; Abdallah, A.E. Detection and Prediction of Insider Threats to Cyber Security: A Systematic Literature Review and Meta-Analysis. *Big Data Anal.* **2016**, *1*, 1–29. [CrossRef]
28. Keim, Y.; Mohapatra, A.K. Cyber Threat Intelligence Framework Using Advanced Malware Forensics. *Int. J. Inf. Technol.* **2019**, 1–10. [CrossRef]
29. Samtani, S.; Abate, M.; Benjamin, V.; Li, W., Cybersecurity as an Industry: A Cyber Threat Intelligence Perspective. In *The Palgrave Handbook of International Cybercrime and Cyberdeviance*; Palgrave Macmillan: Cham, Switzerland, 2020; pp. 135–154. [CrossRef]
30. Chockalingam, S.; Pieters, W.; Teixeira, A.; van Gelder, P. Bayesian Network Models in Cyber Security: A Systematic Review. In Proceedings of the 22nd Nordic Conference on Secure IT Systems (NordSec'17), Tartu, Estonia, 8–10 November 2017; pp. 105–122.
31. Petersen, K.; Feldt, R.; Mujtaba, S.; Mattsson, M. Systematic Mapping Studies in Software Engineering. In Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering (EASE'08), Bari, Italy, 26–27 June 2008; pp. 1–10.
32. Brereton, P.; Kitchenham, B.A.; Budgen, D.; Turner, M.; Khalil, M. Lessons from Applying the Systematic Literature Review Process within the Software Engineering Domain. *J. Syst. Softw.* **2007**, *80*, 571–583. [CrossRef]
33. Wohlin, C. Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering. In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE'14), London, UK, 13–14 May 2014; pp. 1–10.
34. Wieringa, R.; Maiden, N.; Mead, N.; Rolland, C. Requirements Engineering Paper classification and Evaluation Criteria: A Proposal and a Discussion. *Requir. Eng.* **2006**, *11*, 102–107. [CrossRef]
35. The MITRE Corporation. Common Weakness Enumeration (CWE). 2021. Available online: <https://cwe.mitre.org/> (accessed on 22 April 2021).
36. Meland, P.H.; Tokas, S.; Erdogan, G.; Bernsmed, K.; Omerovic, Cyber Security Indicators Mapping Scheme and Result. 2021. Available online: <https://doi.org/10.5281/zenodo.4639585> (accessed on 19 March 2021).
37. Erdogan, G.; Gonzalez, A.; Refsdal, A.; Seehusen, F. A Method for Developing Algorithms for Assessing Cyber-Risk Cost. In Proceedings of the 2017 IEEE International Conference on Software Quality, Reliability and Security (QRS'17), Prague, Czech Republic, 25–29 July 2017; pp. 192–199.
38. Moore, T.; Kenneally, E.; Collett, M.; Thapa, P. Valuing Cybersecurity Research Datasets. In Proceedings of the 18th Workshop on the Economics of Information Security (WEIS'19), Boston, MA, USA, 3–4 June 2019; pp. 1–27.

39. Wagner, T.D.; Mahbub, K.; Palomar, E.; Abdallah, A.E. Cyber Threat Intelligence Sharing: Survey and Research Directions. *Comput. Secur.* **2019**, *87*, 101589. [\[CrossRef\]](#)
40. Barnum, S. Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX). *Mitre Corp.* **2012**, *11*, 1–22.
41. Ramsdale, A.; Shiaeles, S.; Kolokotronis, N. A Comparative Analysis of Cyber-Threat Intelligence Sources, Formats and Languages. *Electronics* **2020**, *9*, 824. [\[CrossRef\]](#)
42. Bromander, S.; Muller, L.P.; Eian, M.; Jøsang, A. Examining the “Known Truths” in Cyber Threat Intelligence—The Case of STIX. In Proceedings of the 15th International Conference on Cyber Warfare and Security, Norfolk, VA, USA, 12–13 March 2020; p. 493–XII.
43. Bromander, S.; Swimmer, M.; Muller, L.; Jøsang, A.; Eian, M.; Skjøtskift, G.; Borg, F. Investigating Sharing of Cyber Threat Intelligence and Proposing a New Data Model for Enabling Automation in Knowledge Representation and Exchange. *Digit. Threat. Res. Pract.* **2021**. [\[CrossRef\]](#)
44. Mavroeidis, V.; Bromander, S. Cyber Threat Intelligence Model: An Evaluation of Taxonomies, Sharing Standards, and Ontologies within Cyber Threat Intelligence. In Proceedings of the 2017 European Intelligence and Security Informatics Conference (EISIC'17), Athens, Greece, 11–13 September 2017; pp. 91–98.
45. Garousi, V.; Felderer, M.; Mäntylä, M.V. The Need for Multivocal Literature Reviews in Software Engineering: Complementing Systematic Literature Reviews with Grey Literature. In Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering (EASE'16), Limerick, Ireland, 1–3 June 2016; pp. 1–6.

**M: 'Assessing cyber threats for storyless systems'**

Included is the submitted manuscript of the paper whilst awaiting consideration.



**M**



# Assessing cyber threats for storyless systems

Per Håkon Meland<sup>a,c,\*</sup>, Dag Atle Nesheim<sup>b</sup>, Karin Bernsmed<sup>a</sup>, Guttorm Sindre<sup>c</sup>

<sup>a</sup>SINTEF Digital, Strindvegen 4, 7465 Trondheim, Norway

<sup>b</sup>SINTEF Ocean, Postboks 4762 Torgard, 7465 Trondheim, Norway

<sup>c</sup>Norwegian University of Science and Technology, Høgskoleringen 1, 7491 Trondheim, Norway

## Abstract

A proper assessment of potential cyber threats is vital for security decision-making. This becomes an even more challenging task when dealing with new system designs and industry sectors where there is little or no historical data about past security incidents. We have developed a threat likelihood estimation approach that supports risk management under such circumstances. Quantifiable conditions are determined from the environment in which the system will reside and operate, that is the availability of potential threat actors, their opportunities of performing attacks, the required means that are needed for the attack to succeed, and motivation factors. Our research method follows the principles of practice research where both researchers and practitioners have played central roles in a real-life development project for a maritime communication system. We used a qualitative case study for feature-based evaluation of the approach and associated tool template, and to gather evidence on practical aspects such as suitability for purpose, efficiency and drawbacks from five user groups. The results show that representative participants from the cyber security and maritime community gave positive and consistent scores on the features, and regarded time usage, traceability of the threat assessment and the ability to indicate underlying uncertainty to be very appropriate. The approach has been proven useful for this domain and should be applicable to others as well, but the template requires up-front investments in gathering knowledge that is relevant and reusable in additional context situations.

© 2021 Published by Elsevier Ltd.

**Keywords:** cyber threats, decision-making, estimation, empirical evaluation, case study, maritime communication

## 1. Introduction

Many recent reports show that cyber attacks are becoming more sophisticated and frequent [1, 2, 3, 4]. This makes it a difficult task to decide how much and what kind of security is needed to protect organisations and their systems. Cyber security decision-making is uncertain by nature, and even more so when dealing with new system designs and industry sectors that are undergoing rapid digitalisation, opening themselves up to more exposure. Under such circumstances, we can talk about systems that are *storyless*, meaning that there is little or no (his-)story or knowledge related to past security incidents. Such data, e.g., *attack frequency*, *attack type distribution*, *number of successful/prevented attacks*, are often required input when trying to quantify threat likelihood in traditional methods. With storyless systems, we must seek other ways to assess potential threats and their consequences in order to make informed decisions on risk treatment.

\*Corresponding author

Email addresses: [per.h.meland@sintef.no](mailto:per.h.meland@sintef.no) (Per Håkon Meland), [dag.atle.nesheim@sintef.no](mailto:dag.atle.nesheim@sintef.no) (Dag Atle Nesheim), [karin.bernsmed@sintef.no](mailto:karin.bernsmed@sintef.no) (Karin Bernsmed), [guttorm.sindre@ntnu.no](mailto:guttorm.sindre@ntnu.no) (Guttorm Sindre)

11 The purpose of this paper is to present a systematic approach for assessing threats for storyless systems. The goal  
12 has been to develop something that can be readily applied in real-life projects, being efficient in terms of resource usage  
13 and flexible enough to be adjusted to the best data available. With this approach, we are able to make threat estimations  
14 based on the availability of potential threat actors, their opportunities of performing attacks, the required means  
15 (resources) that are needed for the attack to succeed, and motivation factors. Such estimations are less dependent on  
16 historical events data, and therefore allow us to use a proactive approach for assessing new designs and prototypes.

17 Through a case study performed in relation to a maritime system development project, we have sought answers to  
18 the following research questions:

- 19 1. How can we estimate threat likelihood for a new design?
- 20 2. What are the perceived advantages and disadvantages of such an approach?

21 The project has involved security experts and domain specialists who have participated in actual threat assessments  
22 and evaluated the approach. We hope that this contribution will be a practical and relevant addition to existing risk  
23 management methods, within the maritime as well as other domains.

24 The structure of this paper is as follows. Section 2 provides information about threat modelling, associated con-  
25 cepts, challenges and state of the art. Section 3 explains our research method and case study. Section 4 explains the  
26 approach itself with an illustrative example. Our evaluation results are presented in Section 5, and we discuss our  
27 results and threats to validity in Section 6. Finally, Section 7 concludes the paper.

## 28 2. Background and state of the art

29 As defined by the ISO/IEC 27000 vocabulary [5]; a *threat* is the potential cause of an unwanted incident, which  
30 can result in harm to a system or organization. When assessing threats, we often talk about *threat modelling*. In 2000,  
31 Schneier [6] described threat modelling as a way of imagining the vast vulnerability landscape of a system and ways  
32 to attack it. He also made a point that this is something hard to do and only comes with experience. Two decades later,  
33 a diverse set of security experts published the *Threat Modeling Manifesto* [7] based on the most common concepts  
34 from the literature throughout the years. The manifesto defines threat modelling as “analyzing representations of a  
35 system to highlight concerns about security and privacy characteristics”, where some of the most central questions  
36 one should try to answer are “what are you building?”, “what can go wrong?”, “what to do about it?” and “did you do  
37 a decent analysis job?”.

38 There is no single, ideal and uniform method of assessing threats and associated risks. There are overarching  
39 processes and practices found within standards such as the ISO/IEC 31000- and 27000-series ([8, 9]) and NIST  
40 publications [10, 11], but exactly how to perform this will usually depend on factors such as the wanted perspective,  
41 experience, personal preferences, available information, and local conditions. When there is little quantitative data  
42 available, subjective opinions become central in the assessments. Though security experts and domain specialists can  
43 make good estimates on consequences following a cyber event, determining the likelihood factor is a harder challenge  
44 as that involves a fair share of guesswork. Böhme et al. have pointed out that [12] “models of cyber risk arrival need  
45 to be more predictive.” This is in accordance with Ahrend and Jirotko [13], who state that “cyber security defenders  
46 need to make more informed decisions regarding what threats to mitigate and how to mitigate them” and “to do so  
47 requires defenders to *anticipate* threat actors’ behaviour”. Almukaynizi et al. [14] have shown a growing community  
48 attention towards predicting cyber security events, and argue that predictions should be transparent and interpretable  
49 to allow human-in-the-loop-driven decisions.

50 In the literature we can find different approaches on how to support human-driven predictions of risk factors. For  
51 instance, Hubbard [15] has proposed the HTMA approach (*how to measure anything*) for cyber security risks, which  
52 heavily relies on subjective expert opinions. Santini et al. [16] have extended this approach, adding more objective  
53 data from several sources to progressively improve the risk model. These *key risk indicators* (KRIs) were mainly  
54 based on measurements internal to the organisation, such as malware infections, vulnerabilities, data breaches and  
55 deep web exposure. Figueira et al. [17] have proposed a mixed qualitative-quantitative risk analysis approach, using  
56 regression models instead of data about the past to compute future threat probability. Similar to Santini et al., they base  
57 their estimations on currently known system vulnerabilities. Kissoon [18] also applies regression models to measure  
58 the effectiveness of current implemented cyber security measures in organisations. She uses internal variables such  
59 as risk appetite, security budget and loss after security breach obtained from surveys and interviews. Al-Hadhrani et

60 al. [19] have proposed to use subjective logic based on the criteria vulnerability level and technical attack difficulty to  
61 compensate for the lack of accurate, probabilistic data.

62 The challenge of threat prediction becomes even more apparent with storyless systems, for which there is virtually  
63 no data about existing vulnerabilities, attack frequencies or loss after incidents. Our approach is mainly concerned  
64 with assessing such systems, and also limiting what is known as *Knightian uncertainty*, where risky (quantifiable)  
65 decisions are made based on non-quantifiable conditions [20]. Instead of taking the system-centric view, we determine  
66 quantifiable conditions from the environment in which the system resides and operates. Previous work that has been  
67 using these premises is for instance presented by Buldas et al. [21], who derive cost of attacks from threat models in  
68 order to decide whether the system is a realistic target for gain-oriented attackers. A similar path can also be seen in a  
69 series of papers by Knez et al. [22], Llansó et al. [23], McNeil et al. [24], that describe a *capability-based approach* to  
70 cyber risk management for space missions. They criticise the amount of labour that is needed to describe attack paths  
71 and give likelihood estimation, emphasizing that these are too subjective and do not scale well for complex systems.  
72 They suggest that mitigations should be based on representations of presumed offensive capabilities of attackers and  
73 the defensive capabilities. Recently, ter Beek et al. [25] have developed a framework for quantitative security risk  
74 modelling where the cost of an attack (both successful and failed) are calculated and used as a constraint. Similarly,  
75 Bagnato et al. [26] use different types of data not tied to past events as part of threat model assessments. They also  
76 advocate for the involvement of domain specialists in order to give accurate estimates, and based on a case study they  
77 identified so-called conflicting modelling goals that have practical implications on the quality of the risk analysis.  
78 These were *time usage* for creating models, *reusability* of context dependent data values, *accuracy* and *simplicity*.  
79 Most of these conflicting goals are in line with the later findings from a survey on graphical security models by Hong  
80 et al. [27], pointing to common practical challenges related to *scalability* of complex models, *reusability* and *tool*  
81 *availability*.

82 In most cases we want to make our estimations based on the best data available, which can be a combination  
83 of some historical data and subjective opinions. For instance, through a set of case studies, Paté-Cornell et al. [28]  
84 have presented several ways to gather and use the information available to quantify cyber risk. For extreme events  
85 without data, they suggest using probabilistic analysis of potential scenarios where the limits of statistical data are  
86 completed by expert opinions. Examples of data are potential points of access, vulnerabilities, software update time  
87 and the costs/loss after successful attacks. Buldas et al. [29] have presented a quantitative attribute approach that deals  
88 with incomplete information. This could be applied when there is some historical data and some domain knowledge  
89 available to the model.

90 Related to the maritime domain, Mraković and Vojinović [30] show that regulatory bodies and international or-  
91 ganisations set risk assessment as a necessary first step for preventing unwanted events at sea, with several sets of  
92 guidelines that refer to the NIST publications. Still, these guidelines do not give details on exactly *how* these assess-  
93 ment should be conducted. Looking at the literature, Tam and Jones [31] have proposed an approach called *Maritime*  
94 *Cyber Risk Assessment* (MaCRA). The risk assessment in MaCRA is based on three dimensions: *system vulnera-*  
95 *bilities*, *ease of exploit*, and the *reward* achieved by the attacker. This approach has some similarity to ours: the  
96 vulnerability dimension resembles our opportunity factor, the ease-of-exploit dimension resembles our *means* factor  
97 (does the attacker have the required means to perform the attack, or at what cost can such means be obtained?), and  
98 the reward dimension resembles our *motivation* factor. On the other hand, while our approach has a separate factor  
99 for *threat actors*, actors are discussed inside the dimensions of reward and ease-of-exploit in MaCRA, for instance,  
100 different types of actors (criminals, terrorists, hacktivists) may be pursuing different types of rewards (money, harm  
101 to an enemy, attention to political causes), and the ease-of-exploit will be different depending on the type of attacker  
102 (e.g. experienced hacker vs. novice). However, the bigger differences are in the way of working with the two ap-  
103 proaches. MaCRA is based on a pre-cataloguing of different types of actors and target system components typically  
104 found in the maritime sector, where picking the system configuration will produce rough estimates of risks for various  
105 threats based on historical data. Our approach rather focuses on people working together to produce estimates for the  
106 weight of various factors, looking at threats one by one, to arrive at a numerical estimate for the threat likelihood.  
107 Hence, rather than being pure competitors, it is also possible that the two approaches could complement each other,  
108 using our approach for the estimation of threat values - but with benefits from MaCRA's pre-cataloguing of various  
109 system components where applicable, and using a MaCRA-inspired approach to visualise the gravity various threats  
110 compared to each other in a nice graphical display.

111 Another work especially addressing maritime cyber-security is Kessler et al. [32], providing a taxonomy to aid

112 risk assessment. The taxonomy supports a way of identifying possible threats to the target system (including both  
 113 malicious attacks and natural hazards), categorizing these threats according to four attributes: the type of attack (e.g.,  
 114 GPS jamming), which security goal (of Confidentiality, Integrity, Availability, Possession, Authenticity, Utility) that  
 115 this attack would invalidate (e.g., Availability in the case of GPS jamming), which systems are involved (e.g., GPS),  
 116 and the threat category (e.g., Jamming). Then, estimates of risk for each threat are derived from tables indicating  
 117 the source of the threat (human attacker or natural hazard), and the likelihood, severity and ease. However, unlike  
 118 our approach, Kessler et al. do not propose a more detailed support or work process for estimating the values for the  
 119 likelihood and ease. This is a main difference from our approach, which tries to go in more detail to provide values  
 120 based on e.g. the attackers opportunities to acquire the necessary means for the attack. Also, our approach does not  
 121 look at natural hazards, but instead has a more detailed breakdown of various human attackers, assigning weights for  
 122 various types of attackers.

123 Svilicic et al. [33] have described how to conduct a cyber risk assessment for a specific ship. The basis of their  
 124 analysis was a combination of a ship crew survey and a technical vulnerability analysis of some of the ship's critical  
 125 system components. In contrast to our approach, such an assessment should be more suitable after deployment  
 126 and when the crew have gained operational experience. You et al. [34] have conducted a literature review on risk  
 127 assessment methods from other domains. They conclude that these can be easily adapted to maritime and port security,  
 128 but it is also clear that they will depend on good subjective estimations or historical data.

129 Further background techniques that our approach directly applies are presented alongside the approach itself in  
 130 Section 4.

### 131 3. Method and materials

132 Our research method follows the principles of *practice research* as defined by Goldkuhl [35], where both re-  
 133 searchers and practitioners play central roles in situational inquiry and generalizing knowledge. We have introduced  
 134 new artefacts in the form of an approach for assessing threats and a tool template that supports this activity. Based on  
 135 Kitchenham [36, 37], we have employed the DESMET evaluation method to assess the appropriateness of our arte-  
 136 facts in the context of a "real" project for the maritime industry. This can be described as qualitative case study, where  
 137 the evaluators make subjective assessments of the relative importance of different features and how well a feature is  
 138 implemented. According to Kitchenham, such an evaluation method is suitable when the benefits are observable on a  
 139 single project and difficult to quantify, and the user population is limited. Zelowitz and Wallace [38] argue that feature  
 140 analysis is well-suited for evaluating new technology and provide insight into its use, and Marshall [39] has shown  
 141 that this is an established evaluation method in software engineering. For these reasons we consider feature-based  
 142 evaluation to be appropriate for our study as well.

143 As depicted in Figure 1, we initially developed the approach by combining and adapting existing techniques for  
 144 threat assessments. Our motivation for doing this was to perform internal risk assessment of the storyless system we  
 145 were developing as part of our case study project, which required us to document and justify our security trade-offs.  
 146 As a second step we chose two representative sub-systems to validate the approach, involving security experts and  
 147 domain specialists that were informally debriefed afterwards. The results of this validation have partly been published  
 148 by Haga et al. [40]. Though we were able to validate that the needs and expectations were met from the sample  
 149 of stakeholders, we also saw possibilities for improving the efficiency by reusing some of the model elements and  
 150 associated values. We therefore expanded the approach and created tool templates to support the activities as part of  
 151 step three. We now reapplied the approach to a larger set of sub-systems in step four, involving additional stakeholders  
 152 and performed a more systematic evaluation in step five.

153 Each evaluation session was conducted as semi-structured interviews, which Robson and McCartan [41] consider  
 154 most appropriate for researchers who are closely involved with the overall project. We had selected a set of core  
 155 features that the participants in each session would score according to a Likert scale and comment on as a group.  
 156 Furthermore, we asked questions recommended by DESMET related to:

- 157 • Suitability for purpose - will the overall approach do the job we want it to?
- 158 • Is the approach efficient in terms of resource usage?
- 159 • Drawbacks - is there any aspect that makes the approach less attractive though it does the job?

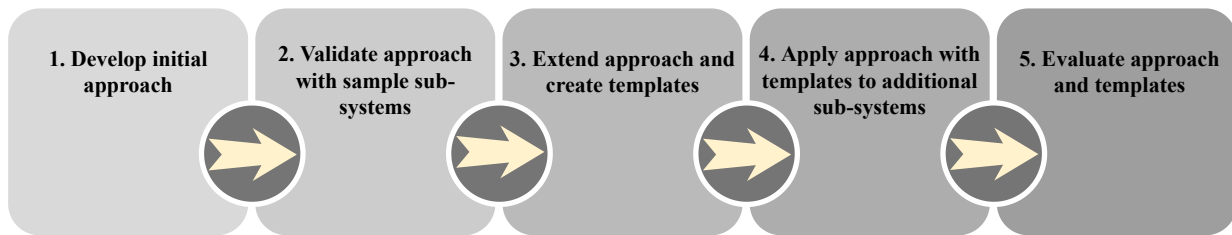


Figure 1. Steps for developing and evaluating the approach.

- Other advantages - are there other attractive aspects of the approach, beside efficiency and fit for purpose?

All participation was voluntary, and the recorded results were anonymised. The details of the actual threat assessment are confidential, but in the following section we give an overview of the case study system to show the context.

### 3.1. Case study: A new maritime communication system

The *maritime* domain is defined as “all areas and things of, on, under, relating to, adjacent to, or bordering on a sea, ocean, or other navigable waterway, including all maritime-related activities, infrastructure, people, cargo, and vessels and other conveyances” [42]. According to Kontovas and Psaraftis [43], the *International Maritime Organisation* (IMO) has recognised that the whole philosophy of using historical data for *Formal Safety Assessment* (FSA) cannot be used for new system designs. Furthermore, it is undesirable to wait for new incidents to happen in order to measure the effects of newly implemented risk controls. We believe the same arguments hold for cyber security risks. Though the maritime domain has a long tradition of safety-focus, ENISA [44] has pointed out that the awareness of cyber security in the maritime community has unfortunately been low. At the same time, the domain is characterised by a complex ICT infrastructure with fast technology development.

Although several studies, such as the ones by Caproli et al. [45], Mraković and Vojinović [30] and Chang et al. [46], give interesting overviews of typical security threats in maritime systems, with some examples of incidents and suggestions of countermeasures, there is little data available to directly quantify the factors relevant for estimating risks. Jacq et al. [47] have proposed a software architecture for monitoring security incidents in maritime systems and setting up a maritime security operations centre to aid vessels in case of attacks. The proposed system would collect data about actual security incidents. If the use of such systems becomes widespread in the future, this would give better data on which to base estimations. Yet at present maritime systems are largely storyless when it comes to cyber-security risk analysis. This yields a need for better support when assessing threats and affirms the domain as interesting from a research perspective.

Our case study has taken place within the context of a research and development project named *Cyber Security in Merchant Shipping Service Evolution* (CySiMS-SE) [48], which lasted from 2019 to 2021. The goal of this project has been to demonstrate and operationalise security for the *VHF Data Exchange System* (VDES) [49] radio and integrate it with the on-board computer architecture. An example use case for this system is for ships to digitally sign and transmit route data to a national coastal administration. A simplified overview of the system is depicted in Figure 2, which shows the main sub-components and how they are connected. On the bridge of the ship, there is a *Global Navigation Satellite System* (GNSS) providing positioning and time data. The VDES is responsible for data transfer to on-shore base stations. A dedicated *Public Key Infrastructure* (PKI) unit is invoked to perform cryptography functions and securely storing the ship’s private key and a cache of public key certificates. A Nav unit integrates digital navigational data and is used by the navigator for planning routes. The GNSS and VDES sub-components are connected to a dedicated IEC 61162-450 [50] compliant network, and traffic needs to go through a firewall to reach either the regular TCP/IP network connected to the PKI-unit and Nav on the bridge or other off-bridge systems, e.g., administrative, crew or entertainment systems. On-shore we can also find a PKI-service that enables enrolment and revocation of certificates, as well as a repository of public key certificates for the flag state.

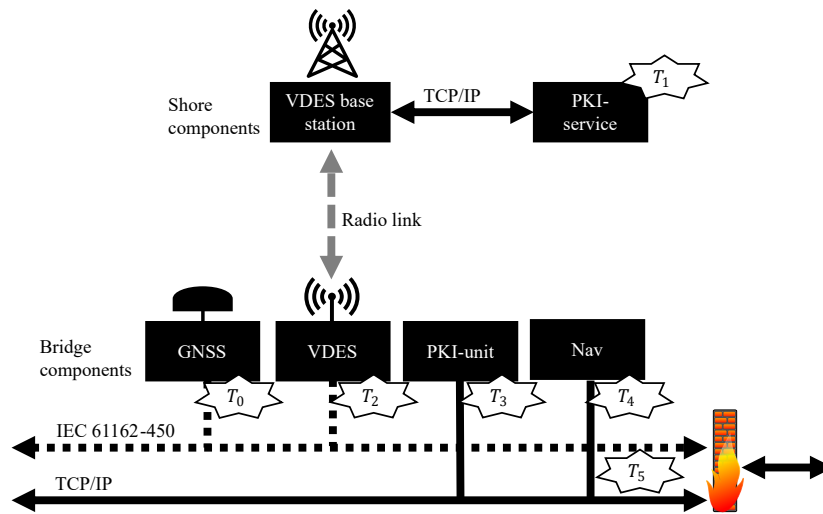


Figure 2. Threats targeting shore-based and on-board bridge components.

197 Based on an analysis [51] of the maritime cyber threat landscape showing that malware infection is the prevalent  
 198 way of compromising systems, the scope of the assessment has been on the unwanted event that one or several of  
 199 the sub-components could become infected and the likelihoods associated to this. The threats we have assessed are  
 200 marked  $T_{1-5}$  in Figure 2, whilst  $T_0$  is used as an example in this paper.

#### 201 4. The threat likelihood approach explained

202 This section explains our approach, which should be seen as a customized version of *OWASP Risk Rating Method-*  
 203 *ology* (OWASPRR) by Williams [52]. Basically, the goal is to “estimate the likelihood of a successful attack from a  
 204 group of possible attackers” based on a model that is simple to use, yet with enough detail to make accurate estimates.  
 205 Williams recommends that the risk rating model should be tailored according to specific organisations, and for our  
 206 approach we have chosen a set of likelihood factors that are more suitable for our use on storyless systems than this  
 207 reference model.

208 Figure 3 shows the four likelihood factors we consider for each threat; threat actors, opportunity, means, and  
 209 motivation. Since we are dealing with intentional attacks, there will always be threat actors actively involved. The  
 210 remaining factors are based on the traditional concept from criminal law, that people who commit crime are likely the  
 211 ones who have *motive*, *means*, and *opportunity* (MMO) to do so [53]. According to Van Ruitenbeek et al. [54], these  
 212 factors are also applicable for analysis in the cyber realm.

213 For each factor we apply the threat template to find a weighted value that gives the following indication:

- 214 • For *threat actors* the weight indicates how large a group the actor represents in comparison to the other actors.
- 215 • For *opportunity* the weight should be based on the threat actor’s spatial, temporal and vulnerability exploiting  
 216 opportunities.
- 217 • For *means* the assessment should consider to what extent the different threat actors have the required means  
 218 needed to perform the attack.
- 219 • The *motivation* weight should be based on what motivation factors and intents that can be associated to each  
 220 threat actor.

221 The weight values are numerical values between 0 and 10 and we derive the overall threat likelihood value from  
 222 the average of these.



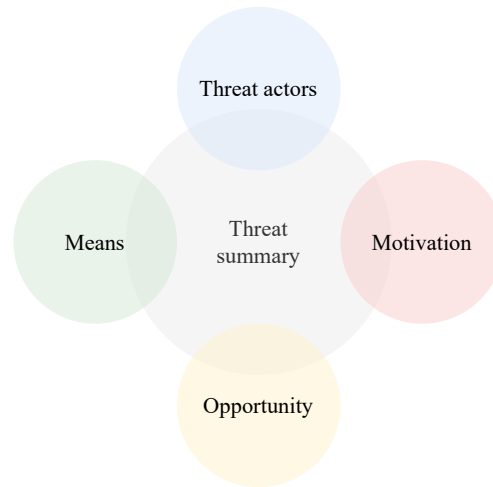


Figure 3. Threat factors used to derive the overall threat assessment.

223 The threat template provides domain knowledge that supports the estimation of the individual threat factors. The  
 224 following sections show how to apply the threat template to the example threat  $T_0$  from the maritime case study. The  
 225 results from each template are used as input to a threat summary, providing traceability and justification for the overall  
 226 threat likelihood. Just as the OWASPRR, we offer a spreadsheet containing the template and a threat summary. This  
 227 tool provides documentation of the threat assessment and enables calculation of the numerical values.

228 Though there is no explicit starting order when working with the different factors, our experience indicates that it  
 229 is natural to begin with threat actors followed by opportunity, means and motivation. All factors can be revisited and  
 230 adjusted iteratively throughout the process.

#### 231 4.1. Identifying threat actors

232 We use *inductive profiling* [55] as a tool to identify potential offenders before any crime is actually committed.  
 233 Shinder and Tittel [56] define a profile to be a set of characteristics likely to be shared by criminals who commit a  
 234 certain type of crime. Our template for threat actors is not only limited to traditional criminals, but also includes  
 235 relevant actors from the maritime operations who could become involved in a cyber attack.

236 Figure 4 shows an excerpt of the taxonomy found within this template. It is not meant to be exhaustive, but serves  
 237 as an inspiration where the assessors can select, add or join elements that are entered into the threat summary. The  
 238 actual threat template contains a more thorough description of each actor based on available literature [57, 58, 59, 60,  
 239 61, 40].

240 Based on the context, we start by picking threat actors that could somehow be involved. In our example we are  
 241 considering a system component on-board the ship, therefore we include profiles among the crew and can disregard  
 242 a lot of the actors tied to land-based operations. The relevant actors are marked with a warning sign in Figure 4.

243 As with the OWASPRR, we use the weight *size* to indicate how large these groups of threat actors are. The  
 244 weights between 0 and 10 are not the actual number of people, but values relative to each other. So for instance, with  
 245 a vessel that has a captain, chief, second and electro-technical officer, these actors are typically given a weight of 1.  
 246 Alternatively, we could merge them into a more generic officer actor with a weight of 2 – 3. There is usually a slightly  
 247 higher number of sailors/ratings on-board, which could yield a weight of 4. It also makes sense to apply a weight of 3  
 248 for technical workers from the shipping company, who could remotely access components or do physical maintenance  
 249 on these. Cyber extortionist is given the highest weight, 8, based on the number of potential online cyber criminals  
 250 we know are out there. Maritime operations are unfortunately often targeted when there are geopolitical conflicts or  
 251 tension between states. In this example, we assign the weight 5 to cyber warrior as the vessel is sailing under a flag  
 252 that has a few hostile nations.



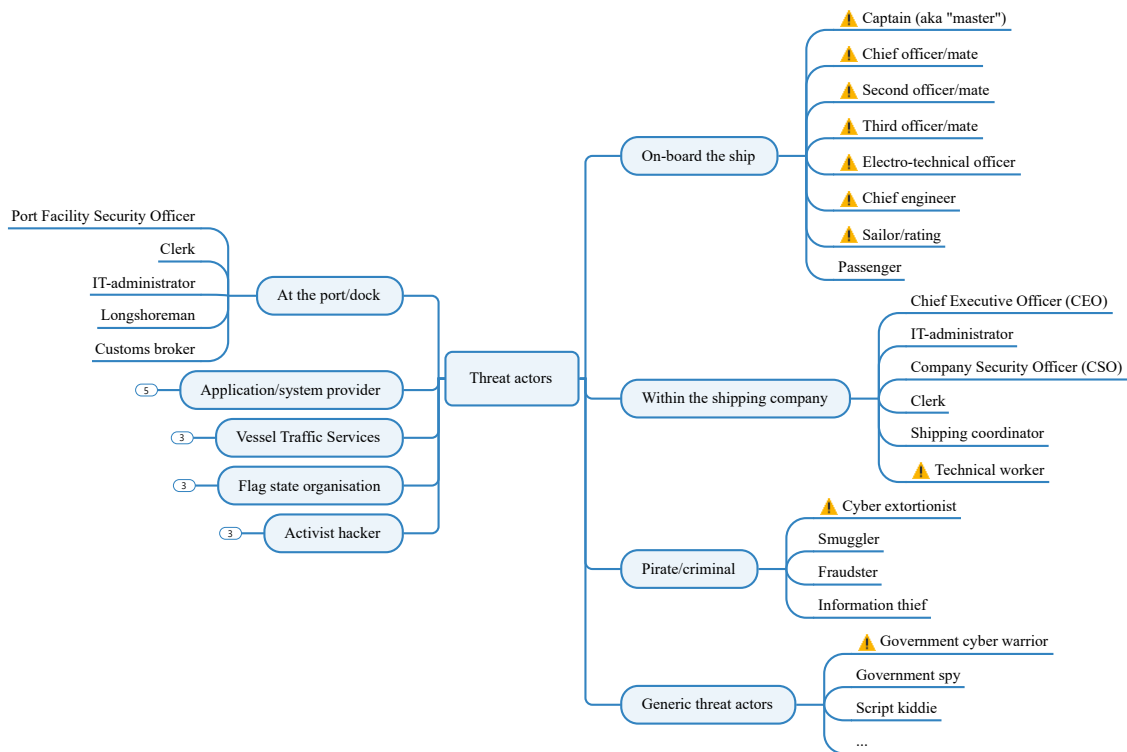


Figure 4. Potential threat actors found in the template.

253 4.2. Finding opportunities

254 Opportunity can be defined as the presence of a favourable combination of circumstances that makes an action  
 255 possible [62]. Opportunity can therefore be used as an indicator for *when* and *where*, and to some extent *how*, the  
 256 threat can manifest itself. If there are vulnerabilities that can be exploited from anywhere, at any time, the opportunity  
 257 weight will be high. If, instead, the adversary must be at the right place at the right time, the weight will be low. In  
 258 practice, not all vulnerabilities can be eliminated, as this would cause excessive security costs and inhibit meaningful  
 259 operations. However, we should strive to make the window of opportunity as small as possible so that the adversary  
 260 cannot easily attack the system without being noticed.

261 In our threat template for opportunity, we take into account that maritime vessels have a changing operational  
 262 environment. We have further divided opportunity into three dimensions. The first one is the spatial dimension,  
 263 which is another name for location. The next opportunity dimension is related to time. In many cases, the spatial and  
 264 temporal characteristics will be interlinked, for instance sailing on autopilot is usually performed at open sea, while  
 265 tugging usually takes place in congested waters. It is possible to have several temporal characteristics for opportunity.  
 266 For instance, a certain attack opportunity may arise while the ship is sailing on autopilot but would need at least 10  
 267 minutes (window size) to succeed.

268 Our third opportunity dimension is related to system vulnerabilities. There must be such vulnerabilities present in  
 269 order to exploit the system. Note that many of these indicators are mostly related to legacy systems, and to a lesser  
 270 degree, new systems still under design/implementation.

271 Figure 5 shows an excerpt from the taxonomy found within the template. Based on the context we choose relevant  
 272 opportunities (marked with a warning sign) for the threat actors and provide a weight with a justification in the threat  
 273 summary.

274 4.3. Deriving the necessary means

275 The required means or resources needed to perform an attack is another factor that helps us determine the threat  
 276 likelihood. While cheap attacks can potentially be implemented by many, more expensive ones require attackers that



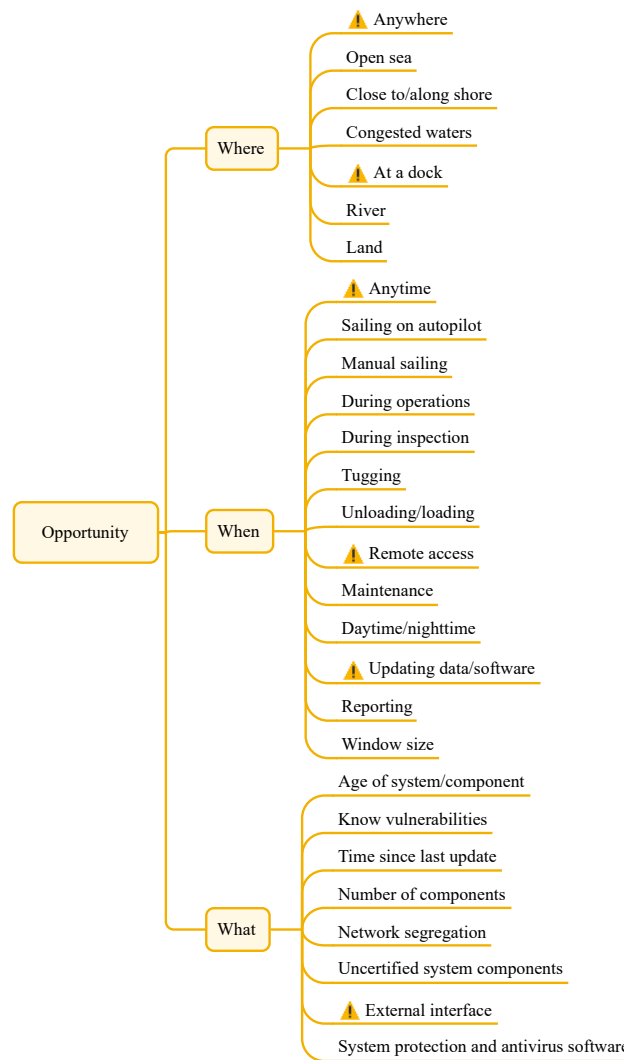


Figure 5. The template provides potential opportunities for the selected threat actors, divided into circumstances related to where, when and what.

277 are more determined to invest. As shown in Section 2, there are different approaches for estimating attacker costs,  
 278 however, most of these are based on known attack paths. With new designs it is more difficult to predict attack paths.

279 We utilise an approach described by Haga et al. [40], which again is based on two methods with an already  
 280 high uptake in the security community, namely the Cyber Kill Chain by Lockheed Martin [63, 64] and attack trees  
 281 by Schneier [65]. Here, a resource tree can be modelled for each consecutive stage of a cyber-attack. These trees  
 282 estimate the fundamental resources that are required to complete this stage and move on to the next one, but differ  
 283 from traditional attack trees since they are not concerned about the details of the attack paths. The tree consists of a  
 284 root node, defining the cyber kill stage, a second level of conjunctive resource classes, and a third level of disjunctive  
 285 resource alternatives. We assign monetary cost values for the resource alternatives along with an optional confidence  
 286 value. For instance, if the attacker would require a certain type of hardware to perform the attack, and the direct  
 287 cost of that item is known, we can assign that value with a confidence value close to 1 (certain). However, in cases  
 288 where we are unsure about the cost, for instance for finding exploitable vulnerabilities, we use a low value such as 0.2  
 289 (uncertain). The cost and confidence values propagate up the trees from the included kill chain stages.

290 Our means template is an alternative to the *Interactive Resource Cost Model* (IRCM) tool by Haga et al. [40].  
 291 Instead of having to model the resource trees from scratch, generic structures are part of the template and only need



292 cost values and optionally confidence. These structures were developed from the validation phase, as we saw that there  
 293 were a lot of common tree elements in the models created for the sample sub-systems. While Haga et al. [40] operate  
 294 with cost intervals for the resource alternatives, our means template simplifies the estimation task by propagating the  
 295 minimum expected costs ( $\alpha$ ) from the alternatives ( $V$ ) for each required resource ( $R_j$ ). The total estimated minimum  
 296 means ( $M$ ) is the sum of all required resources from the included kill stages, which can be formally expressed as:

$$M = \sum_{\substack{\text{stage} \in \\ \text{kill chain}}} \sum_{i \in V} \alpha_i \quad (1)$$

297 As suggested by Haga et al. [40], the overall confidence ( $C$ ) is the product of the average confidence of the  
 298 resource alternatives ( $c_i$ ) to all resources ( $R_j$ ) for the included kill chain stages:

$$C = \prod_{\substack{\text{stage} \in \\ \text{kill chain}}} \prod_R \frac{\sum_{i \in R_j} c_i}{n} \quad (2)$$

299 Figure 6 shows a screenshot excerpt from the means template applied to  $T_0$ , involving the reconnaissance and  
 300 weaponization kill stages. Where resource alternatives or stages are considered irrelevant for the assessment, the cost  
 301 cells can be left blank. Blank confidence values are treated as 1 unless specified otherwise.

302 An essential part of reconnaissance is to do discovery on the target system, meaning to gain knowledge about  
 303 which components/software are installed. This kind of information could for instance be obtained from someone on  
 304 the inside or using more technical scanning techniques (querying external interfaces or analysing data packages). In  
 305 this example both of these options have a similar cost estimate of \$100, but since we are more unsure about how easily  
 306 an insider would give up the information, the confidence value is set to 0.5. Since both values are the same, the cost  
 307 of the discovery/inventory resource amounts to \$100, while the confidence becomes 0.7 (average).

308 An attacker would also have an interest in obtaining documentation of the target system, and that could be done  
 309 legally at a relative low cost for this particular GNSS component. We can actually find and purchase the documen-  
 310 tation from the system provider Web-side, which means an accurate cost estimate with a high confidence. The other  
 311 alternative is to obtain the documentation in an illegal way, for instance by breaking into the system provider premises  
 312 or bribing an insider. Since it is the minimum cost that propagates up the tree, it does not matter so much which cost  
 313 we put into this alternative as long as it is higher than the one above. After a discussion with the system providers,  
 314 who know their premises and employees best, we assume a sum of at least \$10000, but with a low confidence.

315 Another typical part of reconnaissance is to obtain a target unit replica that the attacker could test and experiment  
 316 with. In some cases, the target component could simply be purchased directly from the supplier for a known cost, in  
 317 this example \$1000. It is often possible to obtain a unit from underground channels, black markets, or online auctions.  
 318 In the GNSS example we can quickly search sites such as ebay.com to get price listings of similar second-hand units.  
 319 Since it is more difficult to know the state of used components, possibly stolen from a ship recycling facility, we have  
 320 set the confidence to 0.3. If a physical unit is not needed, another alternative would be to obtain simulation software.  
 321 However, since we already know that the underground alternative is so cheap, we do not have to spend time on this  
 322 estimate. We can also add additional cost to the reconnaissance stage for expenses we cannot fit under the template  
 323 structure.

324 The weaponization stage represents the resources an attacker would have to invest in order to find exploitable  
 325 vulnerabilities in the target system and craft a malicious payload. The threat template contains some reference values  
 326 that can be of support when making these estimates. This includes typical prices for vulnerability data as announced  
 327 in darknet fora and marketplaces (see e.g. Meland et al. [66]), average size of malware (from Calleja et al. [67])  
 328 and average development costs per *source line of code* (SLOC). These numbers are used as a starting point when  
 329 discussing with system owners what kind of investment would be needed to make malware that could perform an  
 330 exploit. We also include reference values for outsourcing development based on hacker group ads as a basis for  
 331 discussion. Of course, the costs of weaponization are crude estimates, only meant to roughly indicate the magnitude  
 332 of attacker investment.

333 After the threat template calculates the resulting means value and confidence, we have to create weights for the  
 334 threat summary. For each threat actor we consider how likely it would be to obtain the required amount of resources.  
 335 A weight value of 1 indicates that it would be nearly impossible for the threat actor, while the other end of the scale  
 336 implies that the resource costs are insignificant.

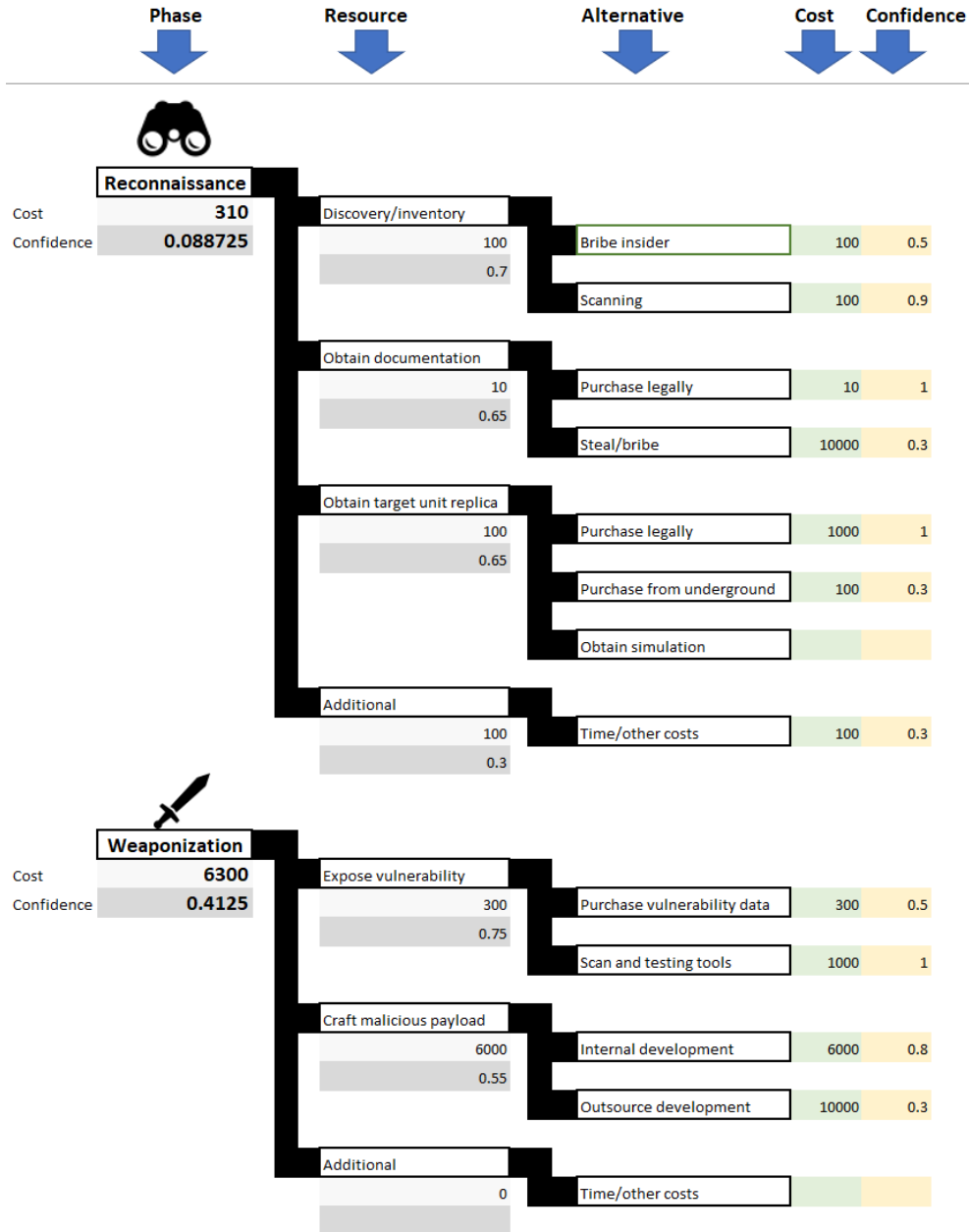


Figure 6. Tool screenshot of the means template, which takes attacker cost with confidence values as input to the various kill stages.

337 4.4. What are the motives and intent?

338 *Motivation* identifies the driver that causes the threat agent to commit harmful acts, and we employ the taxonomy  
 339 by Casey [68] in our motivation template to help us identify the nature of the expected harmful actions. This taxonomy  
 340 is shown in Figure 7, and as the motivations are independent of each other, we can assign any number to one or several  
 341 of the threat actors. A concept related to motive is *intent*, which in criminal law is concerned with the purposeful action  
 342 the threat actor is willing to carry out [69]. We have extended the objective actions presented by Casey [70] with what  
 343 we consider to be additional relevant intents (marked with \*).

344 Based on the motivation template we discuss and fill in values for each threat actor in the threat summary with



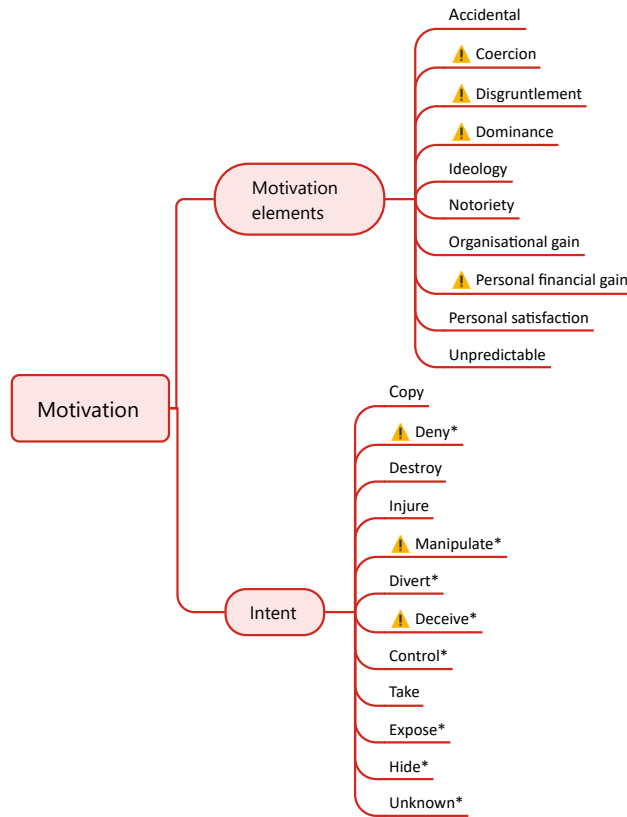


Figure 7. The template suggests possible motivation factors and intended actions that could be tied to the threat actors.

345 a justification of our selection. Just as with the other likelihood factors, we assign a weight between 0 and 10 by  
 346 considering what the actor will get out of it if the attack succeeds (*reward*). Similar to motive in the OWASPRR [52],  
 347 a weight close to 0 indicates that there is little or no reward, a value around 5 possible reward, and 10 a high reward.

348 *4.5. The overall threat and unwanted event estimation*

349 Having completed likelihood estimations for threat actors, their opportunity, means and motivation, we are now  
 350 ready to make a combined average weight as shown in Table 1. In this example there are many possible threat agents,  
 351 of whom cyber extortionist has the highest average weight (6.25), which we will use as the overall likelihood for this  
 352 threat. As pointed out by Williams in the OWASPRR [52], it is better to “err on the side of caution” and use the  
 353 worst-case threat agent and that likelihood value.

354 Our example threat ( $T_0$ ) is one of the possible threats that can cause an unwanted event, as seen in Figure 8.  
 355 The model in this figure is a bow-tie diagram [71, 72, 73], which is one possible way of graphically representing  
 356 multiple potential threats and consequences. It was applied in our case study since this notation is well-known from  
 357 risk management within the maritime industry.

358 In order to give an overall threat estimation that can be utilized in a risk assessment, we can for instance apply  
 359 the model for combining mutually independent threats as proposed by Bernsmed et al. [74]. It is straight forward  
 360 to normalise the likelihood values of the threats to probability values by dividing by 10. Given the assumption that  
 361 the threats can manifest themselves as cyber attacks independently, the probability of the unwanted event  $U$  can be  
 362 computed as:

$$p(U) = p(\text{at least one } T_i \text{ occurs}) = 1 - \prod_{i=1}^n (1 - p(T_i)) \tag{3}$$



Table 1. A simplified threat summary.

Threat actor	Weight	Opportunity	Weight	Means assessment	Weight	Motivation (intent)	Weight	Average weight
Officer (multiple types)	3	Anytime, anywhere	8	Lower means than the reference value, but still significant.	5	Coercion, personal financial gain, accidental (manipulate, deceive).	3	4.75
Sailor/rating	4	Anytime, anywhere	5	Significant sum for this kind of crew.	3	Coercion, personal financial gain, disgruntlement (manipulate, deceive).	5	4.25
Technical worker	3	At a dock, updating	7	Already has expertise and resources available, lower means than reference value.	5	Coercion, personal financial gain, accidental (manipulate).	3	4.5
Cyber extortionist	8	Remote access, external interface	4	Experience from similar attacks would lower required means.	5	Personal financial gain (deny).	8	6.25
Government cyber warrior	5	Remote access, external interface	4	Unlimited resources.	3	Dominance (deny, manipulate, deceive).	5	4.25

where  $p(T_i)$ ,  $i = 1 \dots n$ , is the probability of threat  $T_i$ .

According to Bernsmed et al. [74], Equation 3 is much more realistic than simplistic models where threats are considered mutually exclusive (i.e.  $p(U)$  will be computed as a sum of the individual threats). Allowing threats to manifest themselves within the same time interval corresponds more closely to the real world, where multiple attackers can work simultaneously to exploit different vulnerabilities.

In our case we end up with a probability for the unwanted event close to 0.96 when we apply Equation 3 for  $T_{0..5}$  with the example likelihood values from Figure 8. We would subsequently try to assess the risk by taking consequences ( $C_{1..3}$ ) and treatments into consideration as well. However, this kind of continued risk assessment has been outside the scope of this study and evaluation.

## 5. Evaluation results

Step 4 and 5 of Figure 1 were conducted in five separate workshop sessions assessing the threats  $T_{1..5}$  (see Section 3.1 with five groups of participants, G1-5). The configuration of these groups is shown in Table 2, showing the distribution of security experts and domain specialists among the participants. One security expert acted as an overall session facilitator and one domain specialist was responsible for taking observational notes and record statements during all the sessions, whereas the rest of the participants belonged to the owner (organisation) of the component that the given threat was targeting. The organisations had first-hand knowledge of their own components and operations, with prior experience from assessing risks towards these and similar systems using various techniques. Though the organisations originate from the same geographical area (Norway), they are all well-recognised in international shipping and provide systems and services to customers globally. The results included in this paper do not contain any information that promotes or discredits these. Furthermore, the participants had no commercial nor conflicting



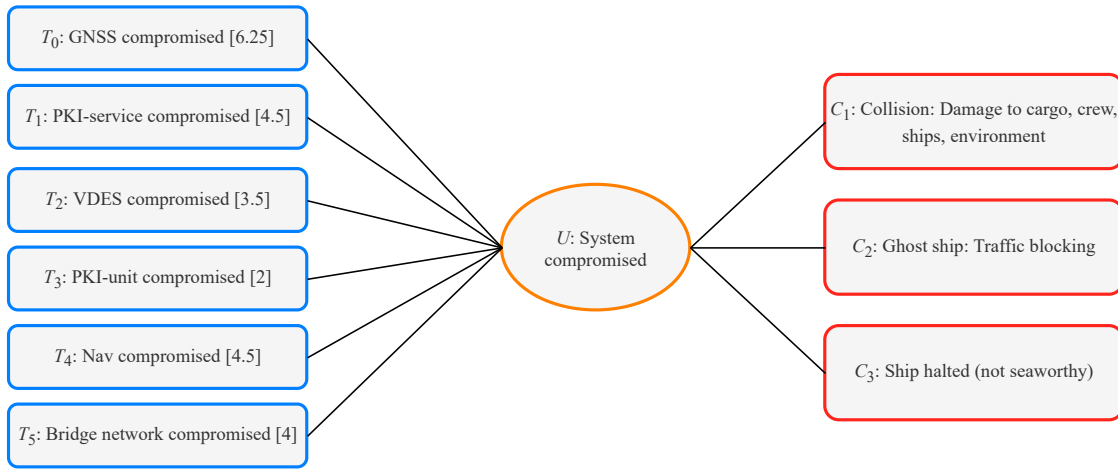


Figure 8. A bow-tie model showing different threats that can cause an unwanted event and subsequent consequences. Likelihood values shown in brackets.

383 interests related to the threat modelling approach. Four of the participants had experience from the validation of  
 384 the initial version of the approach, and all had a general awareness of it since it had been developed as part of the  
 385 CySiMS-SE [48] project that they had participated in.

Table 2. Participants in the evaluation.

Threat	Group	Organisation	Security experts	Domain specialists	Total participants
$T_1$	G1	Maritime authority	2	3	5
$T_2$	G2	System provider	1	1	3
$T_3$	G3	System provider	3	1	4
$T_4$	G4	System provider	1	2	3
$T_5$	G5	Maritime research	1	2	3

386 Each session was organised online using video conferencing, lasted between 60 and 90 minutes and was conducted  
 387 in Norwegian, as this was the native language of all participants. To ensure a proper mindset for the participants,  
 388 there was a general introduction to the session explaining the goals and restrictions of the evaluation. Afterwards,  
 389 a summary of the results was sent to all participants, so that they could comment, modify and finally approve these  
 390 contents.

391 *5.1. Feature-based evaluation results*

392 As already explained in Section 3, we applied a feature-based evaluation. The features we selected correspond to  
 393 the four likelihood factors for threat actors, opportunity, means and motivation, as well as finding the overall threat  
 394 estimation value based on these. The participants discussed how well the approach and templates supported the  
 395 determination of these estimation values, and agreed upon a score from a Likert scale between  $-1$  and  $5$  described in  
 396 Table 3. The resulting scores from each group for each feature are shown in Figure 9. In general, we obtained positive  
 397 scores for all features, with little variance for each group of participants, but more interesting are the comments and  
 398 suggestions we recorded from the discussions. The following sections give a summary of these comments and our  
 399 interpretation of their significance.

400 *5.1.1. Identify potential threat actors*

401 This feature received the highest average score (3.8), which indicates a very strong support of the approach. The  
 402 rather extensive list of potential threat actors found within the template was considered to be a very good starting point



Table 3. Likert scale definitions adapted from Kitchenham [37].

Generic scale point	Definition of Scale point	Scale Point Mapping
Makes things worse	Cause confusion. The way the feature is implemented makes it difficult to use and/or encouraged incorrect use of the feature.	-1
No support	Fails to recognise it. The feature is not supported.	0
Little support	The feature is supported indirectly, for example by the use of other tool features in non-standard combinations.	1
Some support	The feature appears explicitly in the feature list of the tools. However, some aspects of feature use are not catered for.	2
Strong support	The feature appears explicitly in the feature list of the tools. All aspects of the feature are covered but use of the feature depends on the expertise of the user.	3
Very strong support	The feature appears explicitly in the feature list of the tools. All aspects of the feature are covered and the tool provides tailored dialogue boxes to assist the user.	4
Full support	The feature appears explicitly in the feature list of the tools. All aspects of the feature are covered and the tool provides user scenarios to assist the user such as “Wizards”.	5

403 for the participants’ selections. One of the participants stated that “this is a systematic approach for assessing threat  
 404 actors. It cannot be trusted 100%, but it’s a good basis for further discussion.” Other statements were: “you still need  
 405 to think for yourself, but this support is appreciated”, “helps set the mindset for the threat picture” and “the template  
 406 saves us a lot of time”. A suggestion from one of the participants was that “the taxonomy could be linked to what the  
 407 maritime industry already considers to be the prevalent threat actors”.

408 Based on our observations, we believe that the level of exhaustiveness must be a compromise between complete-  
 409 ness and effectiveness for the assessment itself. It requires steady guidance from the facilitator to ensure that time is  
 410 not wasted on discussing minor or less relevant threat actors. For all groups, several threat actors that were similar in  
 411 nature were merged into fewer to avoid repetition and save time.

412 It was also observed that some participants found it difficult to discuss potential threat actors when the context of  
 413 the assessment was too vague, e.g., that the details of the ship, cargo and operations were not specific enough. This  
 414 context information could have been used to reduce the taxonomy to begin with, for instance by removing *passenger*  
 415 for cargo ships.

416 Furthermore, some participants found it somewhat difficult to discuss potential threat actors without relating these  
 417 to the foreseen barriers implemented to mitigate threat actors’ access to the asset(s), and the threat actors’ motivation  
 418 and intent to instigate an actual attack. These issues were more of a concern in later stages of the sessions related  
 419 to opportunity and motivation, which the facilitator explained to the participants. By shifting between or iterating  
 420 through the different templates we could in practice use, e.g., motivation as a screening criterion for the threat actors  
 421 as well.

422 Finally, the concept of weight size spurred some confusion among participants. The facilitator had to point out  
 423 that we were looking for relative and not precise numbers for the given threat actors. For the template, we may benefit  
 424 from creating a standardised presentation of the size parameter with concrete examples from the industry (for instance  
 425 the number of crew on-board certain ship types and/or ship sizes).

426 *5.1.2. Identify potential threat opportunities*

427 This feature had an average score of 3.2 indicating strong support from the approach. The statements from the  
 428 participants were among similar lines, for instance that “the template has suitable content”, “I could not think of  
 429 anything that was not already there” and “it kick-starts the reasoning process”. At the same time, it was expressed that  
 430 it provides “somewhat lower support (than threat actors), I’m not sure we have caught every aspect”.

431 We noted that all participants expressed a need to identify potential threat opportunities. Nevertheless, the concept  
 432 of *where* was considered less relevant than *when*, possibly because some participants related cyber threats to remote





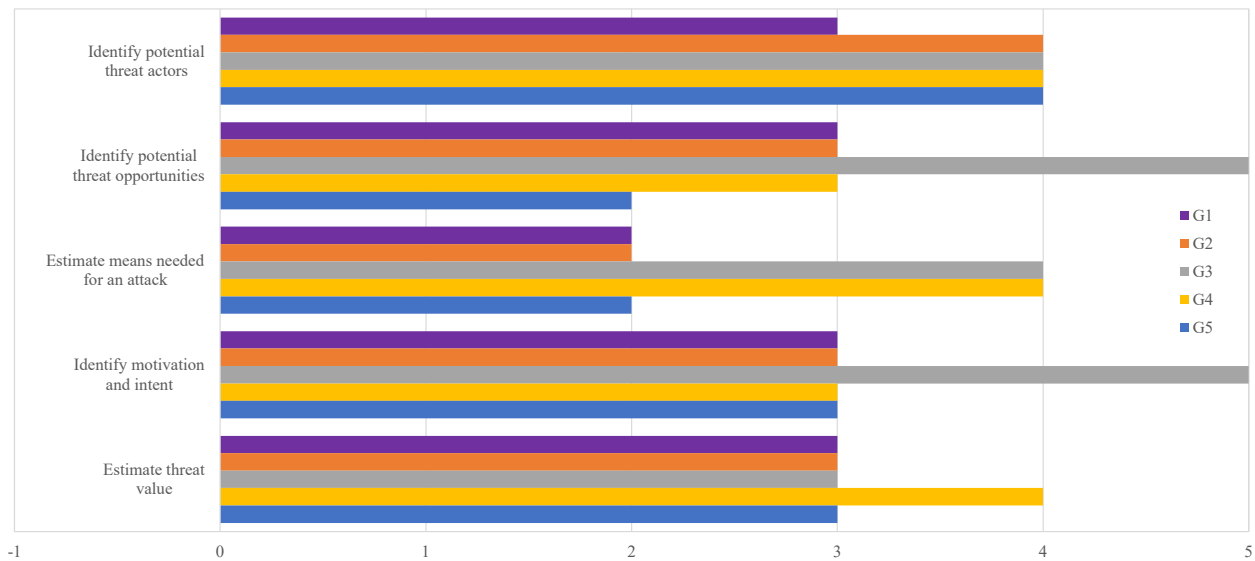


Figure 9. Scores from the feature-based evaluation.

433 access attacks only, hence considering physical attack points as less relevant. We do not think this was a major  
 434 issue for the assessment, but a lesson learned is that the approach would benefit from an improved explanation of the  
 435 importance and implications of the *where* concept.

436 The concept of *when* was considered highly relevant in some attack situations (such as disabling ship navigation  
 437 in ports or high traffic areas), but also less relevant for other types of attack (such as stealing or denying access to  
 438 information). As with threat actors, there is a need to ensure a proper compromise between being generic and specific  
 439 for our assessment.

440 The concept of *what* was not considered very relevant in this case study since we were assessing new designs,  
 441 though the storyless characteristics of the systems, such as number of components (complexity), network segregation  
 442 and external interface could have been highlighted more during the opportunity discussions.

### 443 5.1.3. Estimate means needed for an attack

444 This feature received the lowest average score (2.8), and it is also the activity within the approach that requires  
 445 most time and effort. The feedback from the participants indicated pros and cons for this part of the approach, such as  
 446 “the template saves us a lot of time coming up with estimates, but it is still a difficult task. The confidence parameter  
 447 is important”, “this is a cool way of calculating attack costs, which is not tied to a specific attack ... at the same  
 448 time we lost track of what we were really trying to achieve” and “it would have been difficult to estimate attack costs  
 449 without the template, we do not have a clear idea about these costs to begin with”. There were also suggestions  
 450 for improvements, though the participants acknowledged that this would require more effort, for instance “ideally we  
 451 should estimate costs for each of the selected threat actors, but that would be too time consuming”. Another participant  
 452 suggested to reduce effort at the cost of accuracy: “we could perhaps simplify the template by using scales rather than  
 453 explicit costs, the estimations will be rough anyway”.

454 The sheer size of the template puts substantial demands on the facilitator in terms of guiding the participants  
 455 through the different phases of planning and executing an attack. However, we observed good practices of reducing  
 456 the scope, such as disregarding the most “mission impossible” inspired ways of attacking. In addition, the option  
 457 of skipping or de-emphasizing some of the attack phases enabled a more practical approach that can be adapted to  
 458 the most likely attack scenarios. All the threats in our case study were related to malware infections, and the groups  
 459 focused mainly on the *reconnaissance*, *weaponization* and *delivery* phases of the cyber kill chain [64]. In these phases,  
 460 there are typical direct costs that the participants could relate to, while in the later phases the main means are more  
 461 about effort or indirect costs.

462 It was commented that good estimates require a combination of industry domain and ICT/security knowledge,  
463 which was regarded as well-balanced in our groups. However, it was an important task for the facilitator to keep the  
464 details of the discussion to a level that everyone could relate to. Also, searching for second-hand maritime technology  
465 on eBay.com and other market sites seemed to be a fun exercise to get price estimates, but could also steal quite some  
466 time and focus from the assessment. The use of USD as standard currency evoked some unnecessary confusion as  
467 this was a foreign currency for the participants. It may be beneficial to use the local currency or automatically convert  
468 currencies on the fly. This was no deal-breaker, but in some cases the trail of thought was broken and extra time had  
469 to be spent to align the amounts.

470 Finally, it became somewhat evident from the template that the relationship between blackmailing and bribing is  
471 something that must be considered depending on crew and location. Shipping is an international industry where crew  
472 originate from all over the world. In low-cost countries, a bribe may be cheaper than blackmailing, while in high-cost  
473 countries, the situation may be opposite. One could also relate this to cultural differences, but such a minefield may  
474 be better to avoid for the sake of the discussion.

#### 475 5.1.4. Identify motivation and intent

476 This feature had a high average score (3.4) between strong and very strong support. From three of the groups  
477 there was a general agreement that the taxonomy of motivation and intent seemed adequate, while one group stated  
478 that “maybe it is more complete than necessary”.

479 Participants saw this feature as very relevant and as useful documentation in addition to just determining a numerical  
480 weight value. Nevertheless, the role of motivation and intent, and especially their interrelationship, were observed  
481 to be somewhat confusing at times. One may argue that motive is more closely linked to the threat actors and should  
482 be part of their identification. Intent on the other hand, is more an aspect of the attack or its consequence, and was a  
483 subject that also came up when discussing means. The facilitator needs to guide these discussions and possibly shift  
484 between different parts of the template if new aspects are identified, e.g., an additional threat actor based on discussion  
485 around motivation. Also, the sheer number of motivational elements and intents require steady facilitation to ensure  
486 that focus is kept on the most relevant ones.

#### 487 5.1.5. Estimate threat value

488 The feature that summarised the results from the other estimates received an average score of 3.2, indicating a  
489 strong support. It derives weighted values for threat actors, opportunity, means and motivation seen in combination  
490 with each other, calculates an average weighted threat value and highlights the most likely threat actor. It was stated  
491 that it “provides good background documentation of the estimates and basis for decision-making”. Another participant  
492 pointed out that it “provides a good structure and ranking of threat actors, but could also lead to a false sense of  
493 completion. The approach is good as long as the implementation (of it) is done properly”. As each group only  
494 assessed one type of threat towards their component, they could not really see the greater threat picture. This became  
495 apparent by the statement: “we cannot really say what the threat value means without knowing the other threats”.

496 In general, all participants expressed positive remarks towards how the different stages in the approach resulted  
497 in an overview. It is imperative that we have identified which threats to include in the assessment in the first place.  
498 Even threats with a low score, e.g.,  $T_3$ , are still relevant and should by no means be disregarded. When we apply  
499 Equation 3, such threats contribute to raising the overall probability of the following unwanted event. This implies  
500 that with more threats, the more likely the unwanted event becomes. At the same time, assessing many threats is time  
501 consuming and we would like to include the ones that really makes an impact to the probability of the unwanted event,  
502 and subsequently a quantifiable risk value when we also take consequences into account.

#### 503 5.2. Evaluation of the approach as a whole

504 The last part of the evaluation treated questions from DESMET related to suitability for purpose, efficiency, draw-  
505 backs, and other advantages as mentioned in Section 3. Though these answers partially repeated or overlapped with  
506 the feature-based answers, the sections below summarise the participants opinions on the approach as a whole.

### 5.2.1. Suitability for purpose

Our impression is that the participants regarded the approach as a suitable tool for assessing threats. This was backed by the statements: “(the approach) achieves what it’s meant to achieve”, “it does what it’s supposed to do in a good way” and “this is a scientific approach that both reduces and shows uncertainty. It would have been more difficult to estimate threat likelihood without this kind of organisation”.

They also saw it as a useful addition to more classic (and more resource demanding) methods for estimating threat likelihood based on threat intelligence and historical data. Some participants even saw it as better than classic methods as the data availability, or the lack of thereof, is a barrier when trying to use statistical probability. It was stated that “risk assessments are notoriously difficult, and anything that helps is a step in the right direction. This approach utilises several (likelihood) factors, which gives more credibility to the result”.

It takes some time to become familiar with the approach and the threat template, even for the people involved in developing these. We believe this will improve with time and application, something that was expressed by one of the participants as well: “it’s a good tool, but we need more experience with it”.

### 5.2.2. Efficiency in terms of resource usage

The participants from all groups shared mostly positive responses related to the time invested in the assessments, such as “it is pretty effective ... not sure the results would have been different if we spent more time”, “I don’t think we would get better results if we spent a week on this”, “with other methods it would have been difficult to get just as good answers in shorter time” and “it is much more efficient to use the template than creating models from scratch”.

The participants seemed to think that the approach was relatively simple to use, and yet there is some flexibility on how much time and effort that could be spent for each likelihood factor. Less time usually means less details, so there is always a trade-off. It was stated that “it’s a good thing that we do not model specific attacks. That’s complicated and expensive to do, and this approach provides just as good prioritisation of potential threat events”.

Based on our observations, 60-minute sessions would probably be too short for the type of threats we assessed as part of our case study, while 90 minutes proved to be more suitable.

### 5.2.3. Drawbacks

Though the approach seemed to do the job it was designed for, there were also some weaknesses pointed out. For instance, there were statements related to presence of uncertainty, but without clear suggestions for improvements: “even with this approach there is still a good deal of gut feeling, which is hard to quantify. However, the same issue goes for all other methods as well”, “some of the likelihood factors are easier to assess than others. The approach has great potential, though we have to accept that there is still a lot of uncertainty. I’m not aware of other methods that are more practical” and “the baseline information within the template, how complete is that?”.

We also recorded more detailed comments on the contents of the threat assessment, such as “opportunities related to physical access to the system could have been better explained. Maintenance (crew) would often have full access, but that would be logged and misuse detected. The model did not represent this in a clear way”. It should be noted that taking risk modifiers into account were not really the goal of this assessment. At the same time, it may be unnatural to discuss threats without considering existing barriers in the system environment.

One minor remark that should be easy to fix was “the terminology should have been translated (to Norwegian) to avoid some confusion and ease the discussion”.

All in all, it seems like the main drawbacks are not unique to this approach, and it would benefit from being adjusted to the local context.

### 5.2.4. Other advantages

This discussion point revisited many points that had already been covered, such that the approach “gives a quantification of uncertainty, which is a great plus” and “provides an insight into the underlying details/factors”. A bonus effect that could be highlighted was that the participants thought the approach bridged the communication gap between the domain specialists and ICT security experts. It was stated that “in a way, the discussions are useful by themselves”, and that it is useful to get these groups talking together as early as possible in such a project.

## 6. Discussion

We have developed the threat likelihood approach and associated template as artefacts addressing our first research question; *how can we estimate threat likelihood for a new design?* It should not be seen as a total replacement for existing assessment practices, but as an additional, systematic aid when dealing with storyless systems, that may still be on the drawing board or have not been released into the wild yet. At such stages, there is little quantifiable data such as known vulnerabilities, expected attack frequencies, and malware infections, which are often required input to traditional threat or risk analysis. Instead, threat likelihood estimates are based subjective predictions from security experts and domain specialists, coupled with quantifiable conditions derived from the system environment.

We have also tried to address some of the challenges related to practical application of such techniques, as shown by Bagnato et al. [26] and Hong et al. [27]. First and foremost, the amount of work put into detailed analysis of all possible attack opportunities can quickly outgrow its usefulness. Therefore, we have sought to develop an approach that is efficient but still accurate enough for its purpose. The level of detail should be adjusted to the need of the estimation task. One might want to drill down thoroughly for certain threats, which requires more effort than giving a superficial estimate for threats that are already well-known. For similar threats, it might be sufficient to do a detailed analysis of one and use those results for the others. The approach is based on a number of existing techniques and concepts, such as capability-based risk management [22, 23, 24], resource-cost modelling [40], the OWASP Risk Rating Methodology [52] and means, motive, and opportunity from criminal law [53]. Hence, it should be seen more as an evolutionary than revolutionary approach, with flexibility to be combined with other methods and techniques as well.

Creating a template for a specific domain, in our case maritime communication, is another way of achieving more efficiency and accuracy, but this requires a substantial up-front investment that can only be justified if it can be re-used for a large enough set of assessments. For our case study, this has already proved to be worthwhile as we have been assessing several systems more than one time within the same domain. The template [75] has been made openly available under a CC BY 4.0 license and can be readily applied to similar projects, thus seeking to address the tool availability challenge mentioned by Hong et al. [27].

In order to address our second research question; *what are the perceived advantages and disadvantages of such an approach?*, we have performed qualitative evaluations with domain specialists and security experts from our case study. Section 5 has already provided the main findings from the evaluation of the features and overall approach. The feature of identifying potential threat actors along with their relative size parameter was very well received by the participants, while finding opportunities, motivation and overall threat value were also considered as strongly supported. Estimating the means needed for an attack was the most demanding task in terms of time usage and finding quantifiable values, and received a score somewhat lower than the others. Still, this was a clearly positive score and the statements from the participants indicate that they liked the method despite being unfamiliar with it.

The second part of the evaluation confirmed many of the positive remarks that already had been given for the features, and both suitability and efficiency were highly valued. We did not perform a direct benchmark comparison with any specific alternative methods, which would have required a different evaluation setup. However, the security experts and domain specialists were familiar with various types of assessments methods from before, so the statements related to time usage and drawbacks should be seen as a general comparison. It is noteworthy that beside providing threat likelihood with traceability, the approach also worked as a platform for discussion that the participants appreciated. This shows the importance of having some common ground where people with different expertise can interact.

Based on the evaluation, we believe that the approach and template can become even more appreciated with some slight adjustments and increased familiarity among its users.

### 6.1. Threats to validity

As argued by Cruzes and ben Othmane [76], there will always be a number of potential threats to validity related to science of security and empirical software engineering. However, there are ways of mitigating these threats and thus improving the quality of the research. Here, we highlight threats related to *credibility, transferability, dependability and confirmability*.

601 Making use of an established evaluation method increased credibility and ensured that we gathered both supporting  
602 and discrepant opinions and observations concerning the approach. As depicted in Figure 1, we had developed a self-  
603 conscious research design that followed our case study project. Such a prolonged research engagement allowed us to  
604 do early validation, try out alternative variations within the approach and gave the researchers an opportunity to build  
605 trust with the end-users. At the same time, the threat assessment was only one of the tasks performed within the overall  
606 project, and most of the attention targeted the specification, implementation, and testing of the communication system  
607 itself. This gave the case study a realistic context where the approach was used in practice for security decision-making  
608 related to ongoing development. The results gave the participants a direct benefit and was not seen as an irrelevant  
609 extra burden. We have tried to address the bias of convenience sampling by making sure that the participants had  
610 different backgrounds and belonged to different types of organisations (see Table 2), but we acknowledge that the  
611 population was rather small. This limitation was the main reason why we chose a qualitative case study evaluation to  
612 begin with.

613 Though the approach was applied within a maritime cyber threat context, there are reasons to believe that it may  
614 be transferable to other domains and projects as well. First, the approach is based on existing techniques and concepts  
615 that have to some extent already been applied and evaluated for other domains. These techniques also come with  
616 some of their inherent limitations. For instance, the cyber kill chain has been criticised for being too much focused on  
617 malware, not capturing other types of attack so well. Pols [77] has shown that to remedy this limitation, the literature  
618 suggests many variations of the kill chain, some with up to eighteen different phases. For our approach, there is  
619 flexibility on which and how many phases to include, but as already mentioned in Section 5.1.3, it was for the first  
620 three phases that the participants could most easily estimate concrete costs. Second, we have provided a narrative  
621 context description as part of Section 3 to make it easier for other researchers or practitioners to judge whether the  
622 approach would fit for application partly or as whole in other assessments. Third, many of the participants had solid  
623 backgrounds from other domains, and were thus able to give opinions on transferability and external validity.

624 Based on the consistency of the scores from the feature-based evaluation (see Figure 9), we argue for a certain  
625 extent of dependable results from the evaluation. It is more difficult to assess the dependability of the threat assessment  
626 itself, since the different groups had their own sub-component as the main scope. Since the actual results of these  
627 assessments are confidential, we are unable to show what the details were. However, we would like to state that  
628 for this similar type of threat (malware infection), all of the groups regarded the same types of threat actors as the  
629 most likely ones. As shown by Holm et al. [78], there can be high degrees of uncertainty in data quality when  
630 expert judgment is used. Their experiments showed a significant negative correlation and a strong positive correlation  
631 between experience and calibration, suggesting that additional years' experience can both decrease and increase the  
632 calibration. It was outside the scope of our assessments to use calibration as the groups had different scope. However,  
633 the same facilitator was used in all workshops, and it became evident that the more experience he gained, the more  
634 effective the facilitation of the sessions. This is by no means a unique observation, but a lesson learned is that it may  
635 be useful to conduct a couple of pre-tests before the actual sessions.

636 To maintain confirmability, that is to reflect the voice of the participants from the evaluation, we have included  
637 representative statements in section 5, as raw as possible. Though there is a translation bias from the Norwegian to  
638 the English language, we do not consider this to be of any significance. The recorded observational data and process  
639 notes have more of a subjective nature, but were shared with the participants after the sessions to allow for comments  
640 and show transparency.

641 Finally, we have to acknowledge that we are dealing with models about the future, where there can be rapid  
642 changes in the threat environment and unknown unknowns that no security expert or domain specialist can be expected  
643 to foresee. We find that the famous quote from Box and Draper [79] sums this up in an excellent way: “Essentially,  
644 all models are wrong, but some are useful. However, the approximate nature of the model must always be borne in  
645 mind”.

## 646 7. Conclusion and further work

647 The threat likelihood approach has been developed to support security decision-making for storyless systems. It  
648 combines a number of existing concepts and techniques from risk management literature, expert judgements, and  
649 domain specific information in a systematic way. The main goal has been to create something applicable for real-life  
650 projects, efficient in terms of resource usage, and adjusted to what is the best data available. Through a systematic

651 evaluation within a maritime case study, we have been able to assess the appropriateness of our contribution. The  
 652 features supporting identification and quantification of threat actors, means, opportunity and motivation were all  
 653 considered to provide some, strong, very strong or full support from representative groups in the cyber security and  
 654 maritime community. Just as important as the threat likelihood value itself, is the ability to provide traceability on  
 655 how the participants estimated it. Furthermore, in cases of underlying uncertainties, it was considered valuable to flag  
 656 indication of this.

657 As for further work, it remains to develop better evidence on the generalisation of the results, both in terms of  
 658 transferability to similar projects within the maritime context and also to different settings. This could be done using  
 659 a similar research method for direct comparison, or through triangulation, mixing in quantitative methods applied to a  
 660 larger set of projects and participants. The approach itself should be considered domain-independent, but the template  
 661 should be adjusted to other contexts, e.g., critical systems related to water supply, energy, hospitals, and aviation to  
 662 name a few. This requires a systematic gathering of relevant domain knowledge that is relevant and reusable for the  
 663 threat assessments.

#### 664 **CRedit authorship contribution statement**

665 **Per Håkon Meland:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation,  
 666 Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Funding acquisition. **Karin**  
 667 **Bernsmed:** Conceptualization, Methodology, Writing - Review & Editing, Funding acquisition. **Dag Atle Nesheim:**  
 668 Validation, Investigation, Resources, Writing - Review & Editing, Project administration, Funding acquisition. **Gut-**  
 669 **orm Sindre:** Writing - Review & Editing, Supervision.

#### 670 **Declaration of competing interest**

671 The authors declare that they have no known competing financial interests or personal relationships that could  
 672 have appeared to influence the work reported in this paper.

#### 673 **Acknowledgments**

674 This work has been partially supported by the projects “Cyber Security in Merchant Shipping - Service Evolution”,  
 675 funded by the Research Council of Norway with contract number 295969, and “CyberSec4Europe”, funded by the  
 676 European Union under the H2020 Programme Grant Agreement No. 830929.

#### 677 **References**

- 678 [1] E. G. Franco, M. Kuritzky, R. Lukacs, S. Zahidi, The global risks report 2021, 16th edition, Tech. rep., World Economic Forum (2021).  
 679 URL [http://www3.weforum.org/docs/WEF\\_Global\\_Risks\\_Report\\_2021.pdf](http://www3.weforum.org/docs/WEF_Global_Risks_Report_2021.pdf)
- 680 [2] ENISA, Enisa threat landscape 2020: Cyber attacks becoming more sophisticated, targeted, widespread and undetected (2020).  
 681 URL <https://www.enisa.europa.eu/news/enisa-news/enisa-threat-landscape-2020>
- 682 [3] T. Burt, Microsoft digital defense report, Tech. rep., Microsoft (9 2020).  
 683 URL <https://www.microsoft.com/en-us/security/business/security-intelligence-report>
- 684 [4] M. S. Jalali, M. Siegel, S. Madnick, Decision-making and biases in cybersecurity capability development: Evidence from a simulation game  
 685 experiment, The Journal of Strategic Information Systems 28 (1) (2019) 66–82. doi:<https://doi.org/10.1016/j.jsis.2018.09.003>.  
 686 URL <https://www.sciencedirect.com/science/article/pii/S0963868717304353>
- 687 [5] ISO, ISO/IEC 27000:2018 Information technology - Security techniques - Information security management systems - Overview and vocab-  
 688 ulary, Standard, International Organization for Standardization (2 2018).  
 689 URL <https://www.iso.org/standard/73906.html>
- 690 [6] B. Schneier, Threat modeling and risk assessment, in: E-privacy, Springer, 2000, pp. 214–229. doi:[https://doi.org/10.1007/978-3-322-89183-9\\_20](https://doi.org/10.1007/978-3-322-89183-9_20).
- 691 [7] Z. Braiterman, A. Shostack, J. Marcil, S. d. Vries, I. Michlin, K. Wuyts, R. Hurlbut, B. S. Schoenfeld, F. Scott, M. Coles, C. Romeo,  
 692 A. Miller, I. Tarandach, A. Douglén, M. French, Threat modeling manifesto (2020).  
 693 URL <https://www.threatmodelingmanifesto.org/>
- 694 [8] ISO, ISO/IEC 27005:2018 Information technology - Security techniques - Information security management systems - Information security  
 695 risk management, Standard, International Organization for Standardization (7 2018).  
 696 URL <https://www.iso.org/standard/75281.html>
- 697

- 698 [9] ISO, ISO 31000:2018 risk management guidelines, Standard, International Organization for Standardization (2018).  
699 URL <https://www.iso.org/iso-31000-risk-management.html>
- 700 [10] G. Stoneburner, A. Goguen, A. Feringa, Risk management guide for information technology systems, NIST special publication 800 (30).
- 701 [11] NIST, Cybersecurity framework version 1.0 (2014).  
702 URL <https://www.nist.gov/cyberframework>
- 703 [12] R. Böhme, S. Laube, M. Riek, A fundamental approach to cyber risk analysis, *Variance* 12 (2) (2019) 161–185.
- 704 [13] J. M. Ahrend, M. Jirotko, *Anticipation in Cyber-Security*, Springer International Publishing, Cham, 2017, pp. 1–28. doi:10.1007/978-3-  
705 319-31737-3\_26-1.  
706 URL [https://doi.org/10.1007/978-3-319-31737-3\\_26-1](https://doi.org/10.1007/978-3-319-31737-3_26-1)
- 707 [14] M. Almukaynizi, E. Marin, M. Shah, E. Nunes, G. I. Simari, P. Shakarian, A Logic Programming Approach to Predict Enterprise-Targeted  
708 Cyberattacks, Springer International Publishing, Cham, 2020, pp. 13–32. doi:[https://doi.org/10.1007/978-3-030-38788-4\\_2](https://doi.org/10.1007/978-3-030-38788-4_2).
- 709 [15] D. W. Hubbard, R. Seiersen, How to measure anything in cybersecurity risk, Wiley Online Library, 2016.
- 710 [16] P. Santini, G. Gottardi, M. Baldi, F. Chiaraluce, A data-driven approach to cyber risk assessment, *Security and Communication Networks*  
711 2019. doi:<https://doi.org/10.1155/2019/6716918>.
- 712 [17] P. Tubío Figueira, C. López Bravo, J. L. Rivas López, Improving information security risk analysis by including threat-occurrence predictive  
713 models, *Computers & Security* 88 (2020) 101609. doi:<https://doi.org/10.1016/j.cose.2019.101609>.  
714 URL <http://www.sciencedirect.com/science/article/pii/S0167404819301592>
- 715 [18] T. Kissoon, Optimum Spending on Cybersecurity Measures: Part II, *Journal of Information Security* 12 (1) (2021) 137 – 161. doi:<https://doi.org/10.4236/jis.2021.121007>.
- 716 [19] N. Al-Hadhrani, M. Collinson, N. Oren, Modelling security risk scenarios using subjective attack trees, *Risks and Security of Internet and*  
717 *Systems* 2021. doi:[https://doi.org/10.1007/978-3-030-68887-5\\_12](https://doi.org/10.1007/978-3-030-68887-5_12).
- 718 [20] A. F. Brantly, Risk and uncertainty can be analyzed in cyberspace, *Journal of Cybersecurity* 7 (1), tyab001. doi:<https://doi.org/10.1093/cybsec/tyab001>.
- 719 [21] A. Buldas, P. Laud, J. Priisalu, M. Saarepera, J. Willemsen, Rational choice of security measures via multi-parameter attack trees, in:  
720 *International Workshop on Critical Information Infrastructures Security*, Springer, 2006, pp. 235–248. doi:[https://doi.org/10.1007/11962977\\_19](https://doi.org/10.1007/11962977_19).
- 721 [22] C. Knez, T. Llansó, D. Pearson, T. Schonfeld, K. Sotzen, Lessons learned from applying cyber risk management and survivability concepts  
722 to a space mission, in: *2016 IEEE Aerospace Conference*, IEEE, 2016, pp. 1–8. doi:<https://doi.org/10.1109/AERO.2016.7500812>.
- 723 [23] T. Llansó, M. McNeil, D. Pearson, G. Moore, BluGen: An analytic framework for mission-cyber risk assessment and mitigation recommen-  
724 dation, in: *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017.
- 725 [24] M. McNeil, T. Llansó, D. Pearson, Application of capability-based cyber risk assessment methodology to a space system, in: *Proceedings*  
726 *of the 5th Annual Symposium and Bootcamp on Hot Topics in the Science of Security*, 2018, pp. 1–10. doi:<https://doi.org/10.1145/3190619.3190644>.
- 727 [25] M. H. ter Beek, A. Legay, A. L. Lafuente, A. Vandin, Quantitative security risk modeling and analysis with RisQFLan, arXiv preprint  
728 arXiv:2101.08677.
- 729 [26] A. Bagnato, B. Kordy, P. H. Meland, P. Schweitzer, Attribute decoration of attack–defense trees, *International Journal of Secure Software*  
730 *Engineering (IJSSSE)* 3 (2) (2012) 1–35. doi:<https://doi.org/10.4018/jssse.2012040101>.
- 731 [27] J. B. Hong, D. S. Kim, C.-J. Chung, D. Huang, A survey on the usability and practical applications of graphical security models, *Computer*  
732 *Science Review* 26 (2017) 1–16. doi:<https://doi.org/10.1016/j.cosrev.2017.09.001>.
- 733 [28] M.-E. Paté-Cornell, M. Kuypers, M. Smith, P. Keller, Cyber risk management for critical infrastructure: a risk analysis model and three case  
734 studies, *Risk Analysis* 38 (2) (2018) 226–241. doi:<https://doi.org/10.1111/risa.12844>.
- 735 [29] A. Buldas, O. Gadyatskaya, A. Lenin, S. Mauw, R. Trujillo-Rasua, Attribute evaluation on attack trees with incomplete information, *Com-*  
736 *puters & Security* 88 (2020) 101630. doi:<https://doi.org/10.1016/j.cose.2019.101630>.  
737 URL <https://www.sciencedirect.com/science/article/pii/S0167404819301774>
- 738 [30] I. Mraković, R. Vojinović, Maritime cyber security analysis—how to reduce threats?, *Transactions on maritime science* 8 (01) (2019) 132–139.  
739 doi:<https://doi.org/10.7225/toms.v08.n01.013>.
- 740 [31] K. Tam, K. Jones, Macra: a model-based framework for maritime cyber-risk assessment, *WMU Journal of Maritime Affairs* 18 (1) (2019)  
741 129–163. doi:<https://doi.org/10.1007/s13437-019-00162-2>.
- 742 [32] G. C. Kessler, J. P. Craiger, J. C. Haass, A taxonomy framework for maritime cybersecurity: A demonstration using the automatic  
743 identification system, *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation* 12 (3) (2018) 429–437.  
744 doi:<https://doi.org/10.12716/1001.12.03.01>.
- 745 [33] B. Svilicic, J. Kamahara, M. Rooks, Y. Yano, Maritime cyber risk management: An experimental ship assessment, *The Journal of Navigation*  
746 72 (5) (2019) 1108–1120. doi:<https://doi.org/10.1017/S0373463318001157>.
- 747 [34] B. You, Y. Zhang, L.-C. Cheng, Review on cyber security risk assessment and evaluation and their approaches on maritime transportation,  
748 in: *Proceedings of the 30th Annual Conference of International Chinese Transportation Professionals Association*, Houston, TX, USA, 2017,  
749 pp. 19–21.
- 750 [35] G. Goldkuhl, The research practice of practice research: theorizing and situational inquiry, *Systems, Signs & Actions* 5 (1) (2011) 7–29.  
751 URL <https://www.diva-portal.org/smash/get/diva2:480214/FULLTEXT01.pdf>
- 752 [36] B. A. Kitchenham, Evaluating software engineering methods and tool part 1: The evaluation context and evaluation methods, *ACM SIGSOFT*  
753 *Software Engineering Notes* 21 (1) (1996) 11–14. doi:<https://doi.org/10.1145/381790.381795>.
- 754 [37] B. Kitchenham, Desmet: A method for evaluating software engineering methods and tools, Tech. rep., University of Keele (8 1996).  
755 URL <http://www.ose1.co.uk/papers/desmet.pdf>
- 756 [38] M. V. Zelkowitz, D. Wallace, Validating the benefit of new software technology, *Software Quality Practitioner* 1 (1).  
757 URL <http://www.cs.umd.edu/~mvz/pub/sqp.pdf>
- 758 [39] C. Marshall, Tool support for systematic reviews in software engineering, Ph.D. thesis, University of Keele (6 2016).

- 763 URL <https://eprints.keele.ac.uk/2431/1/MarshallPhD2016.pdf>
- 764 [40] K. Haga, P. H. Meland, G. Sindre, Breaking the cyber kill chain by modelling resource costs, in: International Workshop on Graphical Models  
765 for Security, Springer, 2020, pp. 111–126. doi:[https://doi.org/10.1007/978-3-030-62230-5\\_6](https://doi.org/10.1007/978-3-030-62230-5_6).
- 766 [41] C. Robson, K. McCartan, Real world research, John Wiley & Sons, 2016.
- 767 [42] DHS, National maritime domain awareness plan for national strategy for maritime security, Tech. rep., Homeland Security Digital Library  
768 (12 2013).
- 769 URL [https://www.hsd1.org/c/national-maritime-domain-awareness-plan-for-the-national-strategy-for-](https://www.hsd1.org/c/national-maritime-domain-awareness-plan-for-the-national-strategy-for-maritime-security/)  
770 [maritime-security/](https://www.hsd1.org/c/national-maritime-domain-awareness-plan-for-the-national-strategy-for-maritime-security/)
- 771 [43] C. A. Kontovas, H. N. Psaraftis, Formal safety assessment: a critical review, Marine technology 46 (1) (2009) 45.
- 772 [44] D. Cimpean, J. Meire, V. Bouckaert, S. Vande Castele, A. Pelle, L. Hellebooge, Analysis of cyber security aspects in the maritime sector,  
773 Tech. rep., ENISA (2011).
- 774 URL [https://www.enisa.europa.eu/publications/cyber-security-aspects-in-the-maritime-sector-1/at\\_download/](https://www.enisa.europa.eu/publications/cyber-security-aspects-in-the-maritime-sector-1/at_download/fullReport)  
775 [fullReport](https://www.enisa.europa.eu/publications/cyber-security-aspects-in-the-maritime-sector-1/at_download/fullReport)
- 776 [45] M. Caprolu, R. Di Pietro, S. Raponi, S. Sciancalepore, P. Tedeschi, Vessels cybersecurity: Issues, challenges, and the road ahead, IEEE  
777 Communications Magazine 58 (6) (2020) 90–96. doi:<https://doi.org/10.1109/MCOM.001.1900632>.
- 778 [46] C. Chang, S. Wenming, Z. Wei, P. Changki, C. Kontovas, Evaluating cybersecurity risks in the maritime industry: a literature review, in:  
779 Proceedings of the International Association of Maritime Universities (IAMU) Conference, 2019.
- 780 [47] O. Jacq, X. Boudvin, D. Brosset, Y. Kermaec, J. Simonin, Detecting and hunting cyberthreats in a maritime environment: Specification and  
781 experimentation of a maritime cybersecurity operations centre, in: 2018 2nd Cyber Security in Networking Conference (CSNet), IEEE, 2018,  
782 pp. 1–8. doi:<https://doi.org/10.1109/CSNET.2018.8602669>.
- 783 [48] CySiMS, Cyber security in merchant shipping (2021).
- 784 URL <http://cysims.no/>
- 785 [49] IALA, VDES - VHF Data Exchange System (2020).
- 786 URL <https://www.iala-aism.org/technical/connectivity/vdes-vhf-data-exchange-system/>
- 787 [50] IEC, IEC 61162-450:2018 maritime navigation and radiocommunication equipment and systems - digital interfaces - part 450: Multiple  
788 talkers and multiple listeners - ethernet interconnection, Standard, International Electrotechnical Commission (2018).
- 789 URL <https://webstore.iec.ch/publication/28704s>
- 790 [51] P. H. Meland, K. Bernsmed, E. Wille, Ø. J. Rødseth, D. A. Nesheim, A retrospective analysis of maritime cyber security incidents, in:  
791 Proceedings of the 14th International Conference on Marine Navigation and Safety of Sea Transportation (TransNav 2021), 2021.
- 792 [52] J. Williams, OWASP risk rating methodology, [Online] (4 2020).
- 793 URL [https://owasp.org/www-community/OWASP\\_Risk\\_Rating\\_Methodology](https://owasp.org/www-community/OWASP_Risk_Rating_Methodology)
- 794 [53] S. G. Pendse, Ethical hazards: A motive, means, and opportunity approach to curbing corporate unethical behavior, Journal of Business Ethics  
795 107 (3) (2012) 265–279. doi:<https://doi.org/10.1007/s10551-011-1037-0>.
- 796 [54] E. Van Ruitenbeek, K. Keefe, W. H. Sanders, C. Muehrcke, Characterizing the behavior of cyber adversaries: The means, motive, and  
797 opportunity of cyberattacks, in: 40th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Supplemental (DSN  
798 2010), 2010, pp. 17–18.
- 799 URL [https://www.perform.illinois.edu/Papers/USAN\\_papers/10VAN01.pdf](https://www.perform.illinois.edu/Papers/USAN_papers/10VAN01.pdf)
- 800 [55] A. Warikoo, Proposed methodology for cyber criminal profiling, Information Security Journal: A Global Perspective 23 (4-6) (2014) 172–  
801 178. doi:<https://doi.org/10.1080/19393555.2014.931491>.
- 802 [56] D. L. Shinder, M. Cross, Scene of the Cybercrime, Elsevier, 2008. doi:<https://doi.org/10.1016/B978-1-59749-276-8.X0001-5>.
- 803 [57] Seafarer's professions and ranks (2020).
- 804 URL [https://en.wikipedia.org/wiki/Seafarer%27s\\_professions\\_and\\_ranks](https://en.wikipedia.org/wiki/Seafarer%27s_professions_and_ranks)
- 805 [58] International ship and port facility security code (2020).
- 806 URL [https://en.wikipedia.org/wiki/International\\_Ship\\_and\\_Port\\_Facility\\_Security\\_Code](https://en.wikipedia.org/wiki/International_Ship_and_Port_Facility_Security_Code)
- 807 [59] Marine surveyor (2020).
- 808 URL [https://en.wikipedia.org/wiki/Marine\\_surveyor](https://en.wikipedia.org/wiki/Marine_surveyor)
- 809 [60] D. Dubay, Why we will never see fully autonomous commercial ships (6 2019).
- 810 URL <https://www.maritime-executive.com/editorials/why-we-will-never-see-fully-autonomous-commercial-ships>
- 811 [61] What is a data controller or a data processor? (2020).
- 812 URL [https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/](https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/controller-processor/what-data-controller-or-data-processor_en)  
813 [obligations/controller-processor/what-data-controller-or-data-processor\\_en](https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/controller-processor/what-data-controller-or-data-processor_en)
- 814 [62] M. A. McKendall, J. A. Wagner III, Motive, opportunity, choice, and corporate illegality, Organization Science 8 (6) (1997) 624–647.  
815 doi:<https://doi.org/10.1287/orsc.8.6.624>.
- 816 [63] E. M. Hutchins, M. J. Cloppert, R. M. Amin, Intelligence-driven computer network defense informed by analysis of adversary campaigns  
817 and intrusion kill chains, Tech. rep., Lockheed Martin Corporation (2010).
- 818 URL [https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-](https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf)  
819 [Driven-Defense.pdf](https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf)
- 820 [64] E. M. Hutchins, The cyber kill chain, [Online] (2021).
- 821 URL <https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html>
- 822 [65] B. Schneier, Attack trees, Dr. Dobbs's journal 24 (12) (1999) 21–29.
- 823 [66] P. H. Meland, G. Sindre, Cyber attacks for sale, in: 2019 International Conference on Computational Science and Computational Intelligence  
824 (CSCI), IEEE, 2019, pp. 54–59. doi:<https://doi.org/10.1109/CSCI49370.2019.00016>.
- 825 [67] A. Calleja, J. Tapiador, J. Caballero, A look into 30 years of malware development from a software metrics perspective, in: F. Monrose,  
826 M. Dacier, G. Blanc, J. Garcia-Alfaro (Eds.), Research in Attacks, Intrusions, and Defenses, Springer International Publishing, Cham, 2016,  
827 pp. 325–345. doi:[https://doi.org/10.1007/978-3-319-45719-2\\_15](https://doi.org/10.1007/978-3-319-45719-2_15).



- 828 [68] T. Casey, Understanding cyber threat motivations to improve defense, Intel White Paper.  
829 URL [https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/understanding-cyberthreat-](https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/understanding-cyberthreat-motivations-to-improve-defense-paper.pdf)  
830 [motivations-to-improve-defense-paper.pdf](https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/understanding-cyberthreat-motivations-to-improve-defense-paper.pdf)
- 831 [69] D. A. Webster, Is there a difference between intent and motive?, [Online] (6 2019).  
832 URL [https://www.thewebsterlawoffice.com/blog/2019/june/is-there-a-difference-between-intent-and-motive-/](https://www.thewebsterlawoffice.com/blog/2019/june/is-there-a-difference-between-intent-and-motive/)
- 833 [70] T. Casey, Threat agent library helps identify information security risks, Intel White Paper 2.
- 834 [71] J. E. Cockshott, Probability bow-ties: a transparent risk management tool, *Process Safety and Environmental Protection* 83 (4) (2005) 307–  
835 316. doi:<https://doi.org/10.1205/psep.04380>.
- 836 [72] P. H. Meland, K. Bernsmed, C. Frøystad, J. Li, G. Sindre, An experimental evaluation of bow-tie analysis for security, *Information &*  
837 *Computer Security* 27 (4) (2019) 536–561. doi:<https://doi.org/10.1108/ICS-11-2018-0132>.
- 838 [73] J. Aust, D. Pons, A systematic methodology for developing bowtie in risk assessment: Application to borescope inspection, *Aerospace* 7 (7)  
839 (2020) 86. doi:<https://doi.org/10.3390/aerospace7070086>.
- 840 [74] K. Bernsmed, C. Frøystad, P. H. Meland, D. A. Nesheim, Ø. J. Rødseth, Visualizing cyber security risks with bow-tie diagrams, in: *International Workshop on Graphical Models for Security (GramSec)*, Springer, 2017, pp. 38–56. doi:[https://doi.org/10.1007/978-3-319-](https://doi.org/10.1007/978-3-319-74860-3_3)  
841 [74860-3\\_3](https://doi.org/10.1007/978-3-319-74860-3_3).
- 842 [75] P. H. Meland, K. Bernsmed, CySiMS threat likelihood approach template (6 2021). doi:<https://doi.org/10.5281/zenodo.4899525>.
- 843 [76] D. S. Cruzes, L. ben Othmane, Threats to validity in empirical software security research, in: *Empirical Research for Software Security*, CRC  
844 Press, 2017, pp. 275–300.
- 845 [77] P. Pols, The unified kill chain: Designing a unified kill chain for analyzing, comparing and defending against cyber attacks, Tech. rep., Cyber  
846 Security Academy (2017).  
847 URL <https://www.unifiedkillchain.com/assets/The-Unified-Kill-Chain.pdf>
- 848 [78] H. Holm, T. Sommestad, M. Ekstedt, N. Honeth, Indicators of expert judgement and their significance: an empirical investigation in the area  
849 of cyber security, *Expert Systems* 31 (4) (2014) 299–318. doi:<https://doi.org/10.1111/exsy.12039>.
- 850 [79] G. E. Box, N. R. Draper, *Empirical model-building and response surfaces*, Vol. 424, Wiley New York, 1987.
- 851

## POSTERS

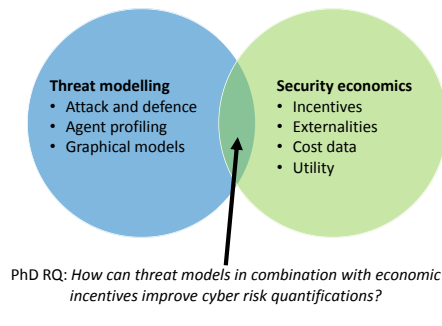
The following posters were presented and their extended abstracts included in the proceedings of conferences:

- Ø: P. H. Meland, 'Combining threat models with security economics,' in *The 11th Norwegian Information Security Conference (NISK)*, IEEE, 2018. [Online]. Available: <https://ojs.bibsys.no/index.php/NISK/article/view/570/486>
- Å: P. H. Meland, 'Resilient cyber security through cybercrime market analysis,' in *REA Symposium on Resilience Engineering Embracing Resilience*, 2019, ISBN: 978-91-88898-41-8. [Online]. Available: <https://open.lnu.se/index.php/rea/article/view/1975/1695>

# Combining threat models with security economics

Per Håkon Meland - per.hakon.meland@ntnu.no

## Research area



## Area of concern

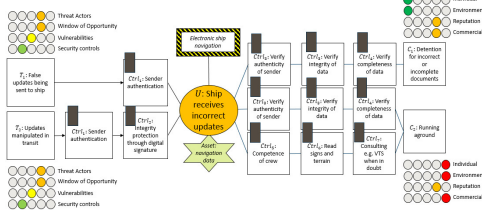


**Accidental threat:** A possibility of human error or omission, unintended equipment malfunction, or natural disaster (e.g., fire, flood, earthquake, windstorm, and other causes).

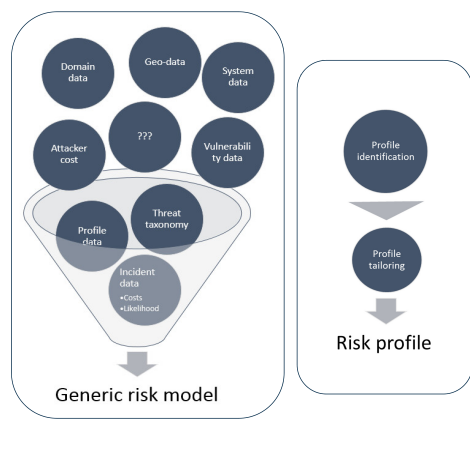
**Intentional threat:** A possibility of an attack by an intelligent entity (e.g., a criminal cracker or a criminal organization).

<https://tools.ietf.org/html/rfc4949>

## Graphical threat models



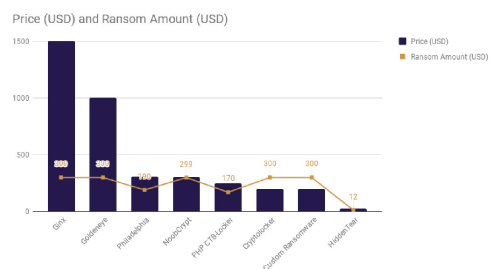
## Risk quantification models



## Incident data

Threat	Frequency	Cost	Frequency	Cost	Risk value
Data - Malicious Breach	0.217	8538707	0.217	8538707	1856717
Privacy - Unauthorized Contact or Disclosure	0.116	5191220	0.116	5191220	601702
Data - Physically Lost or Stolen	0.076	963992	0.000	963992	0
Data - Unintentional Disclosure	0.074	1547330	0.074	1547330	114929
Network/WebSite Disruption	0.032	1227197	1.000	100000	100000
Privacy - Unauthorized Data Collection	0.012	1770338	0.012	177033	21466
Identity - Fraudulent Use/Account Access	0.012	3167541	4.000	100000	400000
Phishing, Spoofing, Social Engineering	0.011	40435298	0.011	40435298	447715
Skimming, Physical Tampering	0.011	1973479	0.000	1973479	0
IT - Processing Errors	0.007	92043291	0.000	92043291	0
Undetermined/Other	0.003	0	0.000	0	0
Cyber Extortion	0.003	92615	0.003	92615	278
IT - Configuration/Implementation Errors	0.003	12427442	0.000	12427442	0
Industrial Controls & Operations	0.001	42655	0.000	4265	0
<b>Total risk value/expected loss per year</b>					<b>[2608007, 3542866]</b>

## Attacker cost ("economy of wickedness")



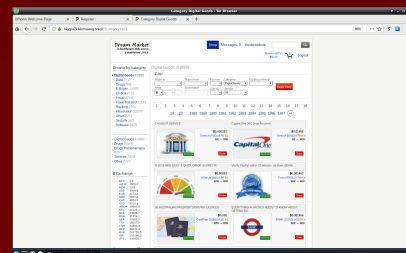
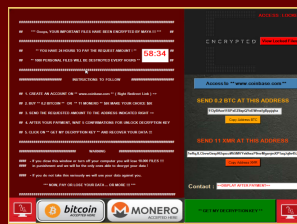
# Resilient Cyber Security through Cybercrime Market Analysis

**NTNU**  
Norwegian University of Science and Technology

**SINTEF**

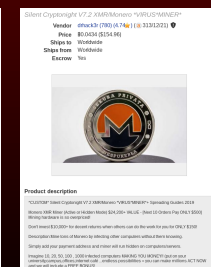
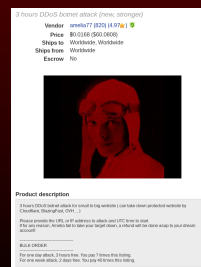
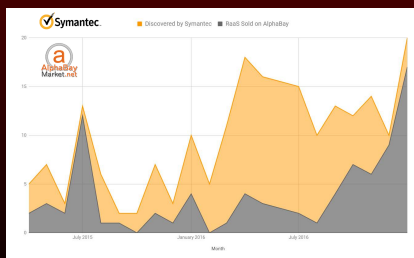
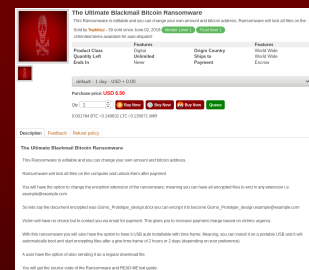
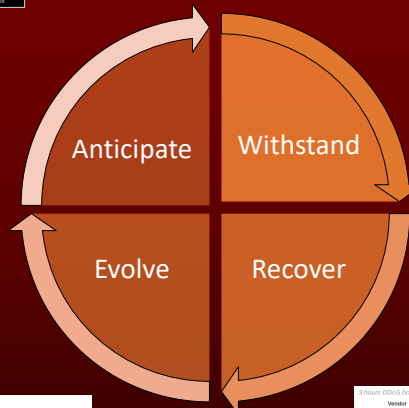
Per Håkon Meland – per.hakon.meland@ntnu.no

Observations from cybercrime markets can be used as sources of resilience, that is the ability to **anticipate** attacks, better **withstand** them once they occur, **recover** from the disturbances and **evolve** the protection mechanisms. Using information about the type and popularity of malicious digital goods, we can get indications about who the attackers are, when these arms are in their hands, what kinds of assets they are looking for and which vulnerabilities they are likely to exploit. Considering the economic incentives motivating the attackers, and not just the capabilities, can also give us an improved understanding of the risk likelihoods as historical incident data quickly becomes outdated, and in the worst case – misleading.



Item for Sale	Avg. Price	Avg. Price Change
Amazon	\$30.36	237 %
Best Buy	\$26.54	121 %
eBay	\$21.66	74 %
NBA	\$15.04	N/A
Fortnite	\$11.33	N/A
Uber	\$11.22	60 %
Netflix	\$10.73	29 %
Apple	\$11.36	-26 %
Facebook	\$9.12	75 %
Airbnb	\$7.61	-3 %

TOP10VPN Dark Web Market Price Index 2019







## AWARDS

The following awards were granted for the papers:

- A:** A. Bagnato, B. Kordy, P. H. Meland and P. Schweitzer, ‘Attribute decoration of attack–defense trees,’ *International Journal of Secure Software Engineering (IJSSE)*, vol. 3, no. 2, pp. 1–35, 2012. DOI: <https://doi.org/10.4018/jsse.2012040101>
- P:** P. H. Meland and F. Seehusen, ‘When to treat security risks with cyber insurance,’ in *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, IEEE, 2018, pp. 1–8, ISBN: 978-1-5386-4565-9. DOI: <https://doi.org/10.1109/CyberSA.2018.8551456>

---

# IGI Global's Fifth Annual Excellence in Research Journal Awards

*IGI Global is proud to present the 2012 Best Article Award to*

**Alessandra Bagnato** (TXT e-solutions, Italy)

**Barbara Kordy** (University of Luxembourg)

**Per Håkon Meland** (SINTEF ICT, Norway)

**Patrick Schweitzer** (University of Luxembourg)

*for the peer-reviewed paper*

"Attribute Decoration of Attack-Defense Trees"

*appearing in the*

International Journal of  
Secure Software Engineering, 3(2)



# CIRC 2018

*2018 International Workshop on Cyber Insurance and Risk Controls (CIRC) - A workshop organised as part of the International Conference On Cyber Situational Awareness, Data Analytics and Assessment (Cyber SA 2018), June 11-12, 2018, Glasgow, SCOTLAND, UK*

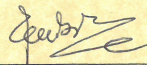
## Best Paper Award

PER HÅKON MELAND AND FREDRIK SEEHUSEN

for outstanding contribution on the paper titled:  
*"When to Treat Security Risks with Cyber Insurance"*

**C-MRiC.ORG**

 **IEEE**  
Advancing Technology  
for Humanity



Dr. Cyril Onwubiko  
Cyber Science 2018 Steering Chair



