Risk Analysis Using "Conflicting Incentives" as an Alternative Notion of Risk

Lisa Rajbhandari

Thesis submitted to Gjøvik University College
for the degree of Doctor of Philosophy in Information Security

HØGSKOLEN I GJØVIK

2013

Risk Analysis Using “Conflicting Incentives” as an Alternative Notion of Risk

Faculty of Computer Science and Media Technology
Gjøvik University College

Risk Analysis Using ”Conflicting Incentives” as an Alternative Notion of Risk / Lisa Rajbhandari
Doctoral Dissertations at Gjøvik University College 2-2013
ISBN: 978-82-93269-32-8
ISSN: 1893-1227

Dedicated to my family.

Declaration of Authorship

I, Lisa Rajbhandari, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed:

(Lisa Rajbhandari)

Date:

Summary

Risk analysis plays an important role in the protection of information systems. The initiatives by governments in many nations clearly show its significance in critical decision making in order to protect the information system. There is a considerable rise in the use of risk analysis methods by banks, hospitals, and many organizations and there is also a growing research interest in this field.

Classical methods for risk analysis usually rely on likelihood estimates that are sometimes difficult to verify. Typically, this is the case when the existing statistical data for the system being analyzed are irrelevant or insufficient (e.g. in the case of non-stationary systems) or one does not have a history for which reliable statistics are available (e.g. in the case of new and emerging systems). In addition, people are not well "calibrated" at estimating probabilities. In most of these classical methods, the events are not usually attributed to people. Moreover, most of these methods focus on risks in relation to threats, overlooking risks in relation to opportunity. Furthermore, the intrusive nature of the risk analysis process makes it hard for researchers or students to gain access to scenarios from operational organizations for evaluating or training on risk analysis methods.

This thesis contributes by developing a new approach for risk analysis: Conflicting Incentives Risk Analysis (CIRA). In CIRA, the stakeholders, their actions, and their perceived expected consequences are identified and used to characterize the risk situation. Risk is modeled in terms of conflicting incentives between the stakeholders in regards to the execution of actions. Thus, CIRA does not rely on the concept of incident likelihood, unlike most of the classical risk analysis methods. Moreover, human related risks are the focus in CIRA.

In order to reduce the sensitivity and confidentiality issues faced due to the intrusive nature of the risk analysis process, a Case Study Role Play (CSRP) approach is introduced. Using CSRP, the required data for a risk analysis method can be collected from the individuals playing the role of fictitious characters rather than from an operational setting. To further exemplify the feasibility of CIRA, a fictitious case study of an Identity Management System (IdMS) similar to the eGovernment IdMS of Norway is analyzed utilizing the CSRP approach.

This dissertation also contributes by presenting the theoretical concepts of risk acceptance and rejection, addressing both threat and opportunity risks in the context of CIRA. Furthermore, an initial insight into how CIRA can be extended to risk management is given by explaining the risk treatment (response) measures for threat (opportunity) risks.

Directions for future research in the area are given by highlighting some of the potential issues such as implementing, validating and improving the method with more case study research and the development of CIRA as a tool. Thus, in order to achieve a robust information security and privacy risk management method, both threat and opportunity risks should be considered, and the human factors need to be explicitly considered during the analysis. CIRA goes towards resolving these issues in the risk management domain.

Acknowledgments

This research has been carried out at the Norwegian Information Security laboratory (NISlab) at Gjøvik University College (GUC). The work reported in this thesis is part of the PETWeb II project sponsored by the Research Council of Norway (RCN).

I owe my deepest gratitude to my supervisor Einar Arthur Snekkenes for his guidance, support and encouragement throughout the research period. I am indebted for his time and guidance whenever I needed it.

I am also thankful to Nils Kalstad Svendsen for his valuable feedback on the thesis. Thanks to my friend Claudia Nickel for proofreading some of the chapters in the thesis and also for her valuable suggestions. I would also like to thank all the participants for agreeing to be a part of my research and for their time and input.

I am grateful to all my friends, colleagues and professors at GUC particularly those at NISlab. I would like to thank them for the wonderful time I had, the interesting talks during our lunch and coffee breaks and overall, for making it a nice and inspiring working environment which I am proud to be a part of. I am also thankful for the administrative, technical support and assistance from the library that I received at GUC. I would like to extend my gratitude to the project members of PETWeb II as well, specifically Lothar Fritsch.

I would also like to thank Steven Furnell, Karin Bernsmed and Stewart Kowalski for being part of the evaluation committee and Stephen Wolthusen for being the head of the committee.

Finally, I would like thank my family and friends, especially my parents for always believing in me. I owe a big thanks to my sister for proofreading the thesis and all the papers, which were sometimes on very short notice. And, I am very grateful to my husband for his encouragement during hard times. I could not have done it without your support. Thank you for being my inspiration!

Contents

List of Figures

List of Tables

List of Theorems

List of Definitions

Part I

Overview

Chapter 1

Introduction

This chapter provides the problem description and motivation for the thesis. It introduces the research questions addressed in this thesis, also depicting the relationship between research questions and published papers. Furthermore, the scope and structure of the dissertation are provided.

1.1 Problem Description and Motivation

Protecting information systems against security and privacy incidents may involve making decisions taking account of huge uncertainties and potential adverse consequences for the risk owner. The risk owner is the person whose perspective we consider when performing risk analysis, i.e., he is the stakeholder at risk. An information system is "a discrete set of information resources organized for the collection, processing, maintenance, use, sharing, dissemination, or disposition of information" including the environment (people, processes, technologies, facilities and cyberspace) in which it operates [20].

Today, there is an increasing number of security and privacy incidents, e.g. theft of intellectual property or financial information of organizations, theft of personal information of individuals, etc. Thus, the task of protecting information systems has become more critical than ever. Risk analysis helps to identify, estimate and evaluate risks, and to provide insight suitable for deciding if risk exposure needs to be changed. In some cases, a treatment or response action might be necessary. In other cases, higher risk exposure might be acceptable because it is more cost effective. Hence, information security and privacy risk analysis has become an essential part of many organizations.

Researchers have proposed a number of quantitative and qualitative risk analysis methods [14]. One of the dominant ways in classical risk analysis and management methods, and guidelines (hereafter called "classical methods") such as ISO/IEC 27005:2008 [12], NIST 800-30 [27] and CORAS [18] is expressing risk as a combination of likelihood and consequence. Usually, in these classical methods, the necessary data is gathered using expert elicitation activity and likelihood is interpreted as (subjective) probabilities or rate claims. However, these methods might be sensitive to estimate errors in the judgment of likelihood.

In most situations, it is difficult to obtain credible likelihood estimates due to the following issues:

- Firstly, people are not well calibrated at estimating probabilities [24]. Experts rely on heuristics in making judgments which might result in systematic biases and errors [24], [28]. As stated in [8], the challenges to Probabilistic Risk Analysis (PRA) are its reliance on subjective judgment and handling of human performance (including human errors, and management and organizational factors).

- Secondly, the statistical data might not be available. For instance in the case of new and emerging technologies or systems, reliable statistics might not be available.

- Thirdly, the existing statistical data might be insufficient [22] or irrelevant as the gathered sample may be too small and/or the system may be dynamic.

These classical methods break down when there is insufficient statistical data to validate the probability or rate claims. This is especially critical in the case of events with low likelihood and high consequence claims or judgments. The rationale behind this is that one requires strong evidence and patience in order to verify that an event is of low likelihood.

Usually, in most of the classical methods, the risk events are not attributed to people. It is often forgotten that people may be the cause of risk events either directly or indirectly. People affect information security risks by giving rise to security breaches or making decisions that are risky [15]. Moreover, misaligned or bad incentives of the individuals generally cause security failure, i.e. trigger risks [3]. Thus, the knowledge about the motives or interests of humans play a significant role in guiding the risk analysis process. The human factor has been overlooked in information security and its consideration is backed by researchers in [1], [4], [10], [15] and [17]. Other researchers are in more harmonization that the focus has mainly been placed on technological factors with less consideration on both human and organizational factors [7], [16], [29]. Murphy [19] writes: "failures of complex engineered systems are often the result of management and organizational factors that influence the decisions and actions of individuals, rather than pure technical problems or isolated instances of human error". He strengthens his point with reports of accidents (e.g. Chernobyl and Three Mile) that list the organizational and human factors as the root causes of the incidents. In the referenced literature, human factors are interpreted in different ways. By human factors, we mean the factors that motivate an individual to take or not to take action(s) to increase his perceived benefit. For instance, these may include money, social relationship, freedom, status or job satisfaction.

In addition, many of these classical methods (for e.g. ISO/IEC 27005:2008 [12], NIST 800-30 [27], CORAS [18], OCTAVE [2], RAMCAP [5] and ISRAM [14]) focus on risks in relation to threats overlooking risks in relation to opportunity. In these methods, the risk events are usually associated with having adverse or unfavorable effect. Hillson [11] states that most of the classical methods consider threats while the opportunities are ignored or addressed only reactively. Furthermore, in [21], Olsson puts forward the evidence that the existing risk management methods consider risk but ignore opportunity.

Risk analysis related activities may identify, process and document sensitive and confidential information regarding threats, vulnerabilities, assets and their valuation, security strategies, etc of an organization. Thus, these are of highly intrusive nature. The researchers or students will not usually be cleared for access to sensitive and confidential information, permitted to perform representative vulnerability discovery activities, or allowed to interview the stakeholders. These issues result in a lack of empirical research [15] and training on risk analysis methods. Because of the intrusive nature of information security research, Kotulic et al. [15] writes they faced difficultly in validating their conceptual model which was based on the study of security risk management at the firm level. Chang et al. [9] suggests using a cautious approach with rapport and trust when conducting empirical studies on information security management.

1.2 Research Questions

The objective of the thesis is to identify and address issues that are important for risk identification, estimation and evaluation so that the overall risk analysis method provides a credible picture of risks facing the risk owner. In addition, we restrict our attention to risks in relation to human behavior.

Main Research Question: What steps should a new risk analysis method involve that does not rely on the concept of incident likelihood ?

In order to accomplish the objective and the main research question, the following sub research questions (RQ) were formed. A brief motivation for each of these questions is given below. However, these are explained in further detail in Chapter 3.

RQ 1. To what extent can game theory be used for analyzing risks?

We started our research with the hypothesis that game theory is suitable for risk analysis. The incentives behind this are, using game theory, we can determine how the subjects select their strategies in situations of interdependence, and how they assess the values of the outcomes of incidents. Thus, this question investigates the suitability of game theory for risk analysis.

RQ 2. How can a risk analysis method be developed with an alternative notion of risk?

Our ultimate goal was to develop a risk analysis method that does not rely on the concept of incident likelihood because, in many cases, it is hard to obtain credible likelihood estimates. Thus, we were interested in investigating whether a new perspective of focusing on conflicting incentives of the stakeholders can provide an alternative notion of risk. This question investigates a new method for risk analysis that models risk in terms of conflicting incentives between the stakeholders.

RQ 3. To what extent is the developed method feasible for analyzing risk in a real life non-trivial setting?

After building the theoretical concept and framework for the method, it was important to explore its practicality. This question investigates the feasibility of the developed method for non-trivial scenarios.

RQ 4. How can we model opportunity risk in the developed method and how can the method be extended to risk management?

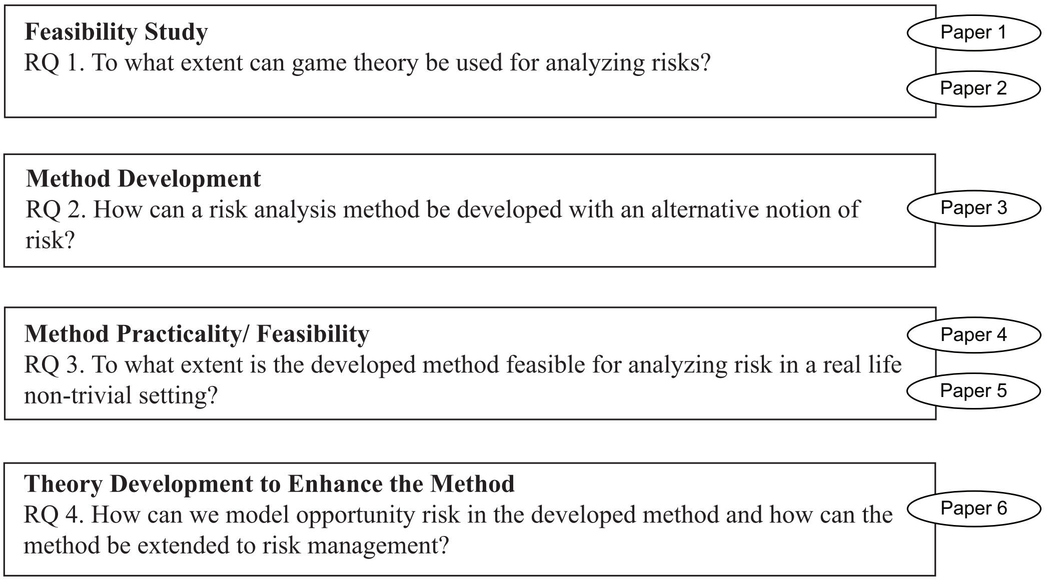
We discovered that one of the serious constraints in most of the risk management methods is the identification and management of opportunity risk. Thus, the given question investigates whether the concepts of risk acceptance and rejection for opportunity risk can be modeled in the developed method, and also looks into whether the method can be extended to risk management.

The above research questions are addressed by the following papers included in this thesis. Additionally, the formation of the research questions and their interdependency can be explained by the four research phases: feasibility study, method development, method practicality or feasibility and theory development to enhance the method. Figure 1.1 depicts the relationship between the research phases, the research questions and the published papers.

1. RAJBHANDARI, L., AND SNEKKENES, E. Using Game Theory to Analyze Risk to Privacy: An Initial Insight. In Privacy and Identity Management for Life, vol. 352 of IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg, 2011, pp. 41-51.

2. RAJBHANDARI, L., AND SNEKKENES, E. Mapping between Classical Risk Management and Game Theoretical Approaches. In Communications and Multimedia Security, vol. 7025 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2011, pp. 147-154.

3. RAJBHANDARI, L., AND SNEKKENES, E. Intended Actions: Risk Is Conflicting Incentives. In Information Security, vol. 7483 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012, pp. 370-386.



Feasibility Study
RQ 1. To what extent can game theory be used for analyzing risks?

Method Development
RQ 2. How can a risk analysis method be developed with an alternative notion of risk?

Method Practicality/ Feasibility
RQ 3. To what extent is the developed method feasible for analyzing risk in a real life non-trivial setting?

Theory Development to Enhance the Method
RQ 4. How can we model opportunity risk in the developed method and how can the method be extended to risk management?

Figure 1.1: Relationship between research phases, research questions and research papers. The numeration of the papers corresponds to the listing of the papers presented in Section 1.2.

4. RAJBHANDARI, L., AND SNEKKENES, E. Using the conflicting incentives risk analysis method. In Security and Privacy Protection in Information Processing Systems, vol. 405 of IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg, 2013, pp. 315-329.
5. RAJBHANDARI, L., AND SNEKKENES, E. Case Study Role Play for Risk Analysis Research and Training. In Proceedings of the 10th International Workshop on Security in Information Systems. SciTePress, 2013, pp. 12-23.
6. RAJBHANDARI, L., AND SNEKKENES, E. Risk Acceptance and Rejection for Threat and Opportunity Risks in Conflicting Incentives Risk Analysis. In Trust, Privacy, and Security in Digital Business, vol. 8058 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2013, pp. 124-136.

The additional research articles that are not included in the thesis but are relevant to the conducted area of research are given below:

- RAJBHANDARI, L. Consideration of Opportunity and Human Factors: Required Paradigm Shift for Information Security Risk Management. In European Intelligence and Security Informatics Conference. IEEE, 2013, pp. 147-150.
- LANGWEG, H., AND RAJBHANDARI, L. Flexible Regulation with Privacy Points. In Trust, Privacy and Security in Digital Business, vol. 7449 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012, pp. 161-166.

1.3 Scope of the Dissertation

The thesis focuses on analyzing risks in the context of information security and privacy management. Information security and privacy risk is concerned with the uncertainty inherent in protecting information systems that are critical to the individual and organization to fulfill their mission. The scenarios used to explore the feasibility of the proposed method

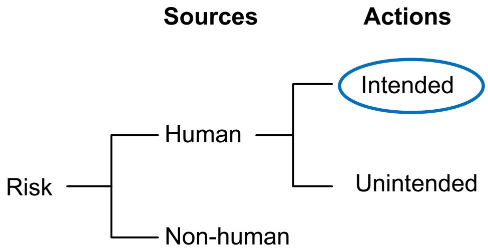


Figure 1.2: Sources of Risk.

in this thesis mainly focus on risks faced by an individual for e.g., risks faced by an end-user of an eGovernment service. Our hypothesis is that the method can be used to analyze organizational risks. However, this requires further investigation.

There are many definitions of security and privacy. Instead of trying to define or differentiate these terms, we look into the actions of a stakeholder(s) that may cause security and privacy risks to the risk owner. E.g. in [26], Solove has provided a taxonomy of privacy risks that is suitable for our work. The taxonomy includes information collection, information processing, information dissemination and invasion as the four main categories of activities of the entities (e.g. other individuals, organizations and the government) that cause privacy problems to a data subject. Each of these are further sub-categorized into activities as enclosed within brackets: information collection (surveillance, interrogation), information processing (aggregation, identification, insecurity, secondary use, exclusion), information dissemination (breach of confidentiality, disclosure, exposure, increased accessibility, blackmail, appropriation, distortion) and invasions (intrusion, decisional interference).

Researchers have put forward various definitions of risk as well. We agree with the view that risk is "relative to the observer" and hence, it is subjective [13]. It is important that we are aware of the presence of subjectivity in the risk analysis process [23]. In addition, for risk management to be considered rational, it should be based on strong argumentation and predetermined structures should be followed, and its steps and process elements should be traceable [6].

The working definition of risk in this thesis is the subjective concern that an individual has towards the outcome of incidents. This includes both the concern that something undesirable might happen and the concern that something desirable might not happen. The former is called threat risk and the latter opportunity risk as introduced in [25].

In the hierarchy given in Figure 1.2, the risk sources are classified into human and non-human. Human risk sources further include both the intended actions (e.g. social engineering attacks, espionage) and unintended actions (e.g. errors or mistakes) of human beings. The non-human risk sources consist of natural disasters (e.g. flood, earthquake) and environmental threats (e.g. power failure, system failure). Our scope is limited to risks caused by the intended (planned) actions of the human as depicted in Figure 1.2. Other categories of risk are out of scope of this thesis.

This thesis mainly focuses on risk analysis. We view risk analysis in a broad context that captures risk identification, estimation and evaluation. However, after evaluating risks, we also explain the risk treatment and response measures for threat and opportunity risks respectively. This extends the method to risk management as one of the activities in risk management is taking actions to treat or respond to those risks that are not within the risk acceptance criteria. The primary objective of risk management is to ensure effective preparedness by means of appropriate resource allocation among controls.

1.4 Structure of the Dissertation

This thesis consists of two parts: the overview in Part I and the research papers in Part II.

In Part I, related work is provided in Chapter 2 followed by the summary of the papers accomplished for this thesis in Chapter 3. In Chapter 4, we give the summary of the main contributions of this thesis. Chapter 5 introduces the potential topics for further research.

In Part II, Chapters 6-11 include the six research papers that constitute the main part of the thesis. An initial insight on the use of game theory for risk analysis is provided in Chapter 6. The mapping between classical risk management and game theoretical approaches is provided in Chapter 7. In Chapter 8, the Conflicting Incentives Risk Analysis (CIRA) method is introduced followed by its application in Chapter 9. Chapter 10 introduces the Case Study Role Play approach. Chapter 11 explains the risk acceptance and rejection for threat and opportunity risks in CIRA and also presents insight into the extension of the method to risk management.

1.5 Bibliography

[1] Adams, A., and Sasse, M. A. Users are not the enemy. *Commun. ACM 42*, 12 (1999), 40–46. doi:10.1145/322796.322806.

[2] Alberts, C., Dorofee, A., Stevens, J., and Woody, C. *Introduction to the OCTAVE Approach*. Carnegie Mellon University, 2003.

[3] Anderson, R., and Moore, T. Information Security Economics - and Beyond. In *In Proceedings of the 27th annual International Crytology Conference on Advances in Cryptology CRYPTO'07* (2007), Springer- Verlag, pp. 68–91. doi:10.1007/978-3-540-74143-5_5.

[4] Ashenden, D. Information Security management: A human challenge? . *Information Security Technical Report 13*, 4 (2008), 195 – 201. doi:10.1016/j.istr.2008.10.006.

[5] ASME Innovative Technologies Institute, LLC. *Risk Analysis and Management for Critical Asset Protection (RAMCAP): The Framework*, May 2006. Version 2.0.

[6] Aven, T. On the Meaning and Use of the Risk Appetite Concept. *Risk Analysis 33*, 3 (2013), 462–468. doi:10.1111/j.1539-6924.2012.01887.x.

[7] Beznosov, K., and Beznosova, O. On the imbalance of the security problem space and its expected consequences. *Information Management & Computer Security 15*, 5 (2007), 420 – 431.

[8] Bier, V. Challenges to the Acceptance of Probabilistic Risk Analysis. *Risk Analysis 19* (1999), 703–710.

[9] Chang, S. E., and Ho, C. B. Organizational factors to the effectiveness of implementing information security management. *Industrial Management & Data Systems 106*, 3 (2006), 345–361.

[10] Gonzalez, J. J., and Sawicka, A. A framework for human factors in information security. In *WSEAS International Conference on Information Security, Rio de Janeiro* (2002).

[11] Hillson, D. Extending the risk process to manage opportunities. *International Journal of Project Management 20*, 3 (2002), 235–240. doi:10.1016/S0263-7863(01)00074-6.

[12] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*, 1st ed. ISO/IEC, 2008.

[13] KAPLAN, S., AND GARRICK, B. J. On the Quantitative Definition of Risk. *Risk Analysis 1*, 1 (1981), 11–27.

[14] KARABACAK, B., AND SOGUKPINAR, I. ISRAM: information security risk analysis method. *Computers & Security 24*, 2 (2005), 147–159. doi:10.1016/j.cose.2004.07.004.

[15] KOTULIC, A., AND CLARK, J. Why there aren't more information security research studies. *Information & Management 41*, 5 (2004), 597–607.

[16] KRAEMER, S., CARAYON, P., AND CLEM, J. Human and organizational factors in computer and information security: Pathways to vulnerabilities. *Comput. Secur.* (2009). doi:10.1016/j.cose.2009.04.006.

[17] LACEY, D. *Managing the Human Factor in Information Security: How to win over staff and influence business managers*. John Wiley & Sons Ltd, 2009.

[18] LUND, M. S., SOLHAUG, B., AND STØLEN, K. A Guided Tour of the CORAS Method. In *Model-Driven Risk Analysis*. Springer Berlin Heidelberg, 2011, pp. 23–43.

[19] MURPHY, D. M., AND PATÉ-CORNELL, M. E. The SAM framework: Modeling the effects of management factors on human behavior in risk analysis. *Risk Analysis 16*, 4 (1996), 501–515.

[20] NIST. *NIST SP 800-39, Managing Information Security Risk - Organization, Mission, and Information System View*, 2011.

[21] OLSSON, R. In search of opportunity management: Is the risk management process enough? *International Journal of Project Management 25*, 8 (2007), 745–752. doi:10.1016/j.ijproman.2007.03.005.

[22] PATÉ-CORNELL, E. On black swans and perfect storms: Risk analysis and management when statistics are not enough. *Risk Analysis 32*, 11 (2012), 1823–1833.

[23] REDMILL, F. Risk Analysis- A Subjective Process. *Engineering Management Journal (IEEE) 12*, 2 (2002).

[24] SHANTEAU, J., AND STEWART, T. R. Why study expert decision making? Some historical perspectives and comments. *Organizational Behavior and Human Decision Processes 53*, 2 (1992), 95–106. doi:10.1016/0749-5978(92)90057-E.

[25] SNEKKENES, E. Position Paper: Privacy Risk Analysis Is about Understanding Conflicting Incentives. In *Policies and Research in Identity Management*, vol. 396 of *IFIP Advances in Information and Communication Technology*. Springer Berlin Heidelberg, 2013, pp. 100–103.

[26] SOLOVE, D. J. A Taxonomy of Privacy. *University of Pennsylvania Law Review 154*, 3 (January 2006), 477. GWU Law School Public Law Research Paper No. 129.

[27] STONEBURNER, G., GOGUEN, A., AND FERINGA, A. *NIST SP 800-30, Risk Management Guide for Information Technology*. NIST, 2002.

[28] TVERSKY, A., AND KAHNEMAN, D. Judgment under Uncertainty: Heuristics and Biases. *Science 185*, 4157 (1974), 1124–1131. doi:10.1126/science.185.4157.1124.

[29] WERLINGER, R., HAWKEY, K., AND BEZNOSOV, K. An integrated view of human, organizational, and technological challenges of IT security management. *Information Management & Computer Security 17*, 1 (2009), 4–19.

Chapter 2

Related Work

This chapter presents the previously published research work, which aided as the main inspiration for the thesis. To comprehend the concept of a field, it is important to have a clear understanding of the terminology. Thus, the chapter starts with an explanation of important terms: risk, uncertainty, threat and opportunity. Then, an overview of risk analysis and management is presented focusing on ISO/IEC 27005:2008 [33] to provide insight relevant for understanding the process of risk management. Several classical risk analysis and management methods are introduced, further explaining the issues around collecting sensitive and confidential information. Additionally, the existing comparison, classification scheme and taxonomy for the methods are presented. As the thesis focuses both on privacy and security issues, a section on privacy impact assessment is also included. Afterwards, relevant literature on the use of game theory for risk analysis and some of the existing works on motivational factors are briefly introduced. Note that parts of this chapter are also included in the publications in Part II and [45].

2.1 Risk, Uncertainty, Threat and Opportunity

In this section, the terms risk, uncertainty, threat and opportunity are briefly discussed to make the concept clearer.

The term risk usually relates to the outcome of events that may be hazardous or that may cause loss. We make decisions everyday taking into account the risk we might face, e.g. when crossing the road, baking a cake, etc. In this case, the decisions that we make are guided by our "fast, instinctive and intuitive reactions to danger" [50]. However, in most of the situations, it is important that we can identify, estimate and evaluate risks logically. Slovic [50] refers to the first case which involves the intuitive or experimental way of thinking as 'risk as feelings' and the second case which involves the analytical way of thinking as 'risk as analysis'. The second case covers the scientific approach to risk such as risk analysis methods. Slovic states that for rational decision making both the analytical and experimental ways of thinking are required.

The term risk is used widely and this often leads to confusion. Risk perspectives are related with "how risk as a concept is defined, interpreted and measured" [9]. In [8], Aven provides an extensive review of how the definition and concept of risk has evolved from the risk perspective relying on 'probabilities' to 'consequences and uncertainties'. The different risk perspectives [8], [9] are given below:

- Risk = Expected value (loss).
- Risk = Probability of an (undesirable) event.
- Risk = Objective uncertainty.
- Risk = Uncertainty.
- Risk = Potential or possibility of a loss.
- Risk = Probability and scenarios/consequences/severity of consequences.

- Risk = Event or consequence.
- Risk = Consequences/damage/severity of events + Uncertainty.
- Risk = Effect of uncertainty on objectives.

In risk analysis, uncertainties are divided into two types: aleatory and epistemic [44]. Aleatory uncertainty (also called stochastic or aleatory variability) represents the randomness in nature. Epistemic uncertainty (also called ambiguity or knowledge-based uncertainty) are those that arise due to limited or lack of knowledge about the fundamental event.

Probability is one of the ways to quantify uncertainties. According to Pate-Cornell [44], there are two schools of thought to understand probability which are the frequentist and the Bayesian. In the frequentist category (including classical statisticians), probability is defined as "a limiting frequency and applies only if one can identify a sample of independent, identically-distributed observations of the phenomenon of interest" [44]. She states in the Bayesian category, probability is defined as a degree of belief. Thus, it is supported by information that includes not only the statistical data and physical models but also subjective judgment of expert. One of the drawbacks of the frequentist school according to her is the definition of probability under it is appropriate only for aleatory uncertainties. However, the Bayesian school is suitable for the quantification of both aleatory and epistemic uncertainties.

The term opportunity is viewed in different ways by the risk analysts and researchers. Clearly there exists two perspectives on the concept of opportunity: firstly, opportunity is considered as the opposite of risk which is a more common way of thinking [43], [58] and secondly, risk is considered as the term that captures both opportunity and threat [28]. The former view is usually captured by the term uncertainty. Risk is defined as the uncertainty with negative consequences while opportunity is defined as the uncertainty with positive consequences. Despite the different opinions, most of the researchers agree that opportunity should be considered either by integrating it into risk management [28], by transforming risk management to uncertainty management [57] or by establishing a separate field referred to as opportunity management [43], [58].

This issue has been stressed mainly in the field of project management. Hillson [28] considers risk as an 'umbrella' term that captures both threat and opportunity. He emphasizes the importance of looking at both, threats (risk with negative consequences) and opportunities (risk with positive consequences), during the risk management process itself. In order to accomplish this, he states that the current risk management method can be made more comprehensive by adding new ways to effectively identify opportunities, using double probability-impact matrix for representing both risks and incorporating new strategies to respond to opportunities which are exploit, share, enhance and ignore. Ward et al. [57] argue that both, threats and opportunities, should be managed. To achieve this, they suggest to transform the current project risk management processes into project uncertainty management. Olsson [43] puts forward the evidence that the existing risk management methods consider risk but overlook opportunity. Furthermore, White [58] states that "the greatest enterprise risk may be in not pursuing enterprise opportunities". Thus, he points out that more concern should be given to opportunity management than risk management at the enterprise level.

2.2 Overview of Risk Analysis and Management

Risk management is usually differentiated from risk assessment or risk analysis; the latter two terms are also usually distinguished from each other. The challenges related to coming up with a common terminology for risk management are well reflected in literature [7], [9] and [54]. Most researchers agree that a consensus is yet to be reached on defining the

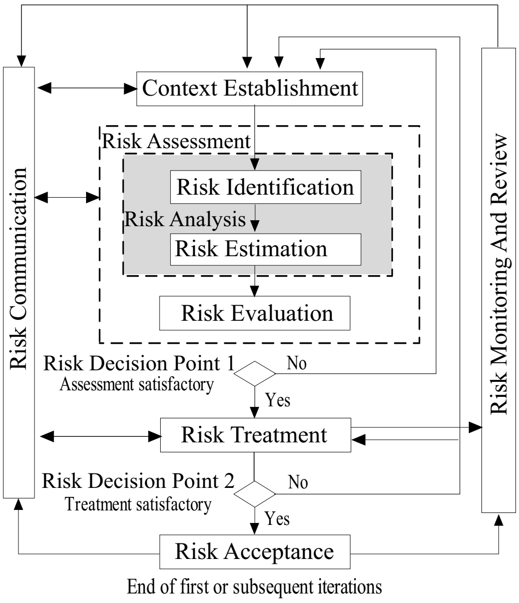


Figure 2.1: Information Security Risk Management Process (taken from [33]).

different terms for risk management. Leaving aside the issue with terminology, we provide an overview of risk analysis and management in this section.

Aven [7] writes: "The ability to define what may happen in the future, assess associated risks and uncertainties, and to choose among alternatives lies at the heart of the risk management system, which guides us over a vast range of decision-making, from allocating wealth to safeguarding public health,...".

According to ISO Guide 73 [34], risk management is the set of systematic activities used to direct and control an organization with regard to risk. Typically, risk management is used to represent the activities: context establishment, risk analysis, risk evaluation, risk treatment, monitoring and review, and communication and consultation. The steps of a risk management process differ widely, but to provide an insight into the risk analysis and management process, we concentrate on ISO/IEC 27005:2008 [33] as depicted in Figure 2.1.

The first step consists of context establishment which includes determining the objectives of the organization, specifying the basic criteria (e.g. setting risk evaluation criteria, risk acceptance criteria), outlining the scope and boundaries of information security risk management, among others.

In the standard, risk assessment consists of risk analysis (risk identification and risk estimation) and risk evaluation. Risk analysis is related to the activity of identifying and estimating risks. In the risk identification step, the assets and their owners are identified. It is followed by the identification of the threats to those identified assets, the existing and planned controls, the vulnerabilities that might be exploited and a record of incident scenarios with their impacts related to those identified assets. This provides a clear picture of the incident scenarios. Afterwards, the consequences and the likelihood of occurrence of those incidents (e.g. loss of confidentiality, integrity or availability of assets) are assessed (which may be expressed either in qualitative or quantitative form). It involves asking questions such as -"How bad can it get?, How likely is it to happen? Can something bad happen as often?" The likelihood and consequence of an incident is assessed taking into account the affected assets, threats, vulnerabilities, consequences and currently implemented or planned controls (if any). Then, the risk is estimated as the combination of the likelihood of an incident and its impact.

In the risk evaluation step, the estimated risks are prioritized according to the risk evaluation criteria and risk acceptance criteria. If the outcome of the risk assessment is satisfactory, the risk treatment options are selected. The selection is based on the result of risk assessment, plus considering the expected cost and benefit of implementing these options. The treatment options are risk reduction/mitigation, risk retention, risk avoidance and risk transfer. By selecting appropriate controls, the level of risk is reduced (by either reducing the likelihood and/or consequence) such that the residual risk is reevaluated as being acceptable. When the level of risk satisfies the risk acceptance criteria, the risks are retained. Risks are avoided in cases where the risks are very high or costs of implementing other risk treatment options surpass the benefits that may be achieved. Risk transfer means transferring or sharing risk with external parties e.g. through insurance. After a risk treatment plan is defined, the residual risks should be determined by reviewing the total exposure of all risks of interest. If the review is not satisfactory, the process is repeated again. In case of a successful review, the exposure to risks is assessed and the risks are accepted (providing justification for those that do not satisfy the risk acceptance criteria). These steps can be iterative until the results are satisfactory.

Risk communication should be carried out throughout the risk management process. The information obtained from the various risk management activities need to be regularly exchanged and shared among the decision maker and other stakeholders. Similarly, there should also be continuous monitoring, reviewing and improvement of the risk management process.

2.3 Classical Risk Analysis and Management Methods

In this section, we first provide an overview of some of the classical risk analysis and management methods and guidelines. Then, we explain the issues around the collection of sensitive and confidential information in the methods, and the existing comparison framework, classification scheme and taxonomy for the methods.

2.3.1 Introduction to Some of the Classical Risk Analysis and Management Methods

In order to determine the challenges in the risk analysis and management domain, it is important to have an understanding of the existing methods. There are many classical risk analysis and management methods and guidelines (hereafter called as "classical methods") but we focus on some of the methods relevant to this thesis. These are the ISO/IEC 27005:2008 standard [33], the ISO 31000 standard [32], NIST 800-30 [52], NIST 800-39 [40], CORAS [16], OCTAVE [3], ISRAM [35], Risk IT [30] framework, RAMCAP [5], CRAMM [49] and TVRA [25].

The ISO/IEC 27005:2008 [33] was developed by the International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) in 2008. Its updated version is ISO/IEC 27005:2011. ISO/IEC 27005:2008 provides guidance on the entire information security risk management process as explained above in Section 2.2. The risk management process includes context establishment, risk assessment (risk analysis and risk evaluation), risk treatment, risk acceptance, risk communication, and risk monitoring and review. Risk is defined as the "potential that a given threat will exploit vulnerabilities of an asset or group of assets and thereby cause harm to the organization" [33]. The origin of the threat is categorized into accidental, deliberate and environmental (natural). Risk can be estimated qualitatively (using descriptive scale e.g. low, medium, high) or quantitatively (using numerical scale) from different sources. The standard further presents the advantages and disadvantages of both the estimation approaches. Qualitative estimation is easier to communicate and understand but depends on "subjective choice of the scale" [33]. The benefit of the quantitative estimation according to the standard is when it is based on historical data, it can directly be associated with the information security objectives of the organization. However, quantitative estimation is not appropriate in the case of lack of reliable data on new systems or information security weaknesses.

The ISO 31000 [32] standard supersedes AS/NZS 4360:2004 [6]. The standard provides "principles and generic guidelines on risk management" [32]. The risk management process includes context establishment, risk assessment (risk identification, risk analysis and risk evaluation), risk treatment, monitoring and review, and communication and consultation. Risk is defined as the "effect of uncertainty on objectives", whether positive or negative [32]. Thus, the guideline can be used to determine risks having both positive and negative consequences.

National Institute of Standards and Technology (NIST) developed the NIST 800-30 [52] with the goal to help organizations improve the management of their IT related risks. Risk management consists of risk assessment, risk mitigation, and evaluation and assessment processes. Risk is estimated as "a function of the likelihood of a given threat-source's exercising a particular potential vulnerability, and the resulting impact of that adverse event on the organization" [52]. In NIST 800-30, risks from the given threat sources are considered: human (unintentional or deliberate actions), natural and environmental.

NIST 800-39 [40] supersedes NIST SP 800-30 and its revised version NIST 800-30 Rev. 1 [41] is a supporting document to NIST 800-39. NIST 800-39 integrates the NIST approach with risk management into a comprehensive Enterprise Risk Management (ERM) [42]. Risk management consists of four components which are (1) frame risk or establish the context, (2) assess risk, (3) respond to risk once it is estimated and (4) monitor risk (using organizational communications and feedback loop for ensuring continuous improvements). Risk is estimated based on "the degree of harm and likelihood of harm occurring" [40]. Further, a multi-tiered approach is used to integrate the risk management process in every part of the organization. This approach helps to address risk at the organizational level, mission or business process level and information system level.

CORAS [16], [15], [38] was developed under the Information Society Technologies (IST) program. It is a model based method that uses Unified Modeling Language (UML) [48] for security risk analysis. It is stated that its model based approach differentiates it from those that rely on text and table based documentation (e.g. CRAMM and OCTAVE) [16]. It is divided into eight steps [38]: preparation for the analysis, customer presentation of target, refining the target description using asset diagrams, approval of target description, risk identification using threat diagrams, risk estimation using threat diagrams, risk evaluation using risk diagrams and risk treatment using treatment diagrams. Risk is defined as "a characterization of the severity of an unwanted incident with respect to a single asset" [38]. The risk value is obtained from the likelihood and consequence of an unwanted incident. It categorizes threat into human threat (e.g. hacker) and non-human threat (e.g. system failure, software bug or natural threats). Human threat is further categorized into intentional threat and unintentional threat.

The Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE) [3], [4] approach is used for information security risk management. OCTAVE was developed by the Carnegie Mellon Software Engineering Institute. There are two categories: the OCTAVE method and OCTAVE-S method for large and small organizations respectively. The key differences of OCTAVE as compared to other methods are its focus on organizational evaluation, security practices, strategic issues and self direction rather than system evaluation, technology, tactical issues and expert requirement [4]. Risk is considered as the possibility of loss/ harm. The threat categories in OCTAVE [2] include: human using network access, human using physical access, system problems and other problems outside the control of an organization (e.g. floods, power outages). The method consists of three phases: build asset based threat profiles, identify infrastructure vulnerabilities, and develop security strategy and plans.

Information Security Risk Analysis Method (ISRAM) [35] was developed at the National Research Institute of Electronics and Cryptology and the Gebze Institute of Technology. It is a survey based model used to analyze risk in information security; two surveys are conducted for gathering probability and consequence. The method is based on risk being modeled as the combination of probability and consequence of a security breach. It consists of seven steps - the first four steps are the survey preparation phase, the fifth step consists of conducting the survey and the last two steps consist of obtaining and assessing the results. The benefit of ISRAM as compared to other methods is stated as its ease of use because no complex mathematical and statistical instruments are required.

ISACA developed the Risk IT [30] framework with the objective of helping organizations manage IT related risks. It is integrated with a business framework, the COBIT 5 [31]. The COBIT 5 framework is a new edition that was released in 2012 and provides the governance and management of enterprise IT. The Risk IT framework is divided into three domains, each with three processes: risk governance (establish and maintain a common risk view, integrate with ERM, make risk-aware business decisions), risk evaluation (collect data, analyse risk, maintain risk profile) and risk response (articulate risk, manage risk, react to events). Risk is estimated as the combination of frequency (rate by which an event occurs over a given period of time) and magnitude of IT risk scenarios. The framework looks at both IT risk and opportunity in an enterprise. The opportunity is concerned with the benefits that can be achieved (for e.g. identifying new business opportunities from using IT).

The Risk Analysis and Management for Critical Asset Protection (RAMCAP) framework [5] was developed by ASME Innovative Technologies Institute. Its new version is RAMCAP Plus. RAMCAP can be utilized for “identifying, prioritizing and coordinating protection of critical infrastructures” in various sectors from security threats [5]. The seven steps are asset characterization and screening, threat characterization, consequence analysis, vulnerability analysis, threat assessment, risk assessment and risk management. Risk is estimated as the combination of threat, vulnerability and consequence. Cox [21] has shown the limitations of estimating risk as the combination of threat, vulnerability and consequence.

CCTA Risk Analysis and Management Method (CRAMM) [49] is a qualitative risk assessment methodology used for risk analysis of information systems and networks. It consists of three stages: (1) asset identification and valuation, (2) threat and vulnerability assessment, and (3) countermeasure selection and recommendation. The steps are carried out with a dedicated automated tool.

Threat Vulnerability and Risk Analysis (TVRA) [25] method was developed by ETSI Technical Committee Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN) for the risk analysis of a telecommunication system. It consists of the following ten steps: identification of the target for evaluation, identification of objectives, identification of functional security requirements, systematic inventory of the assets, systematic identification of vulnerabilities, calculation of the likelihood of the attack and its impact, establishment of the risks, security countermeasure identification, countermeasure cost-benefit analysis and specification of detailed requirements. Risk to the system is estimated by the product of the likelihood of an attack and its impact on the system.

The review of the classical methods is of prime importance to this thesis as it helped to identify the weaknesses within the existing methods. However, the objective of the review is not to compare these methods in terms of their significance in practice, knowledge required to use the methods, etc.

The review shows that in most of the classical methods, risk is expressed as the combination of likelihood and consequence. In most situations, these methods may be sensitive to estimate errors in the judgment of likelihood. Moreover, it was determined that most of these methods focus on threat risks without considering opportunity risks and also, human factors are not explicitly considered during the analysis. This thesis investigates a method to address these shortcomings.

2.3.2 Collection of Sensitive and Confidential Information

The described classical methods involve collection and processing of sensitive and confidential data. In the ISO/IEC 27005:2008 [33] standard, for e.g., when the scope of risk management is established, information about the organization is collected so as to determine its operational setting or environment. The collected information includes the organization's strategic business objectives, strategies and policies, business processes, the organization's functions and structure, information assets, constraints affecting the organization, etc. Other information required for conducting risk analysis is also gathered throughout the process.

Risk management depends on the foundation of the best available information [32]. These information sources according to the ISO 31000 standard may include historical data, experience, stakeholder feedback, observations, forecasts and expert judgments. In Risk IT [30], the data collection process under the risk evaluation domain is dedicated to gathering data on the organization's operating environment and risk events in order "to enable effective IT related risk identification, analysis and reporting".

The collection of sensitive and confidential information by these classical methods may cause hindrance to get access to an operational setting when conducting research or training on risk analysis methods.

2.3.3 Comparison Framework, Classification Scheme and Taxonomy for the Methods

Comparisons between some of the above classical methods can be found in [14], [53] and [56]. In [14], the developed comparative framework for information security risk management methods is based on one of COBIT's Planning and Organisation Controls (which is Assess Risks). It helps to determine whether the methods are in line with the IT governance recommendations. Vorster et al. [56] introduced a framework to help the organization choose the most suitable method for its requirements. The framework is based on five criteria which are: if the risk analysis is done on a single asset or a group of assets, where in the methodology is the risk analysis carried out, the individuals involved in the risk analysis, the main formula used in the methodology and whether the results obtained are relative or absolute. There are studies that compare the risk management guidelines. For instance, in [46], Raz et al. provide a comparison of some of the risk management standards. The comparison is based on the scope (project or organization), the three main process steps (identification, analysis and treatment) and a special emphasis on the standards.

A classification scheme is developed by Campbell et al. [17] for the risk analysis and management methods. It relies on two orthogonal aspects, which are the level of detail and whether the approach is temporal, functional or comparative. The motive behind the classification scheme is to help the practitioners make the right choice by understanding "what to expect from a given method, how it relates to other methods and how best to use it [17]. In [24], European Network and Information Security Agency (ENISA) provides an inventory of risk management/assessment methods and tools and describes the characteristics of these methods and tools based on the selected attributes. Apart from activities such as risk identification and risk analysis, the attributes consist of language, price, licensing, etc.

A taxonomy for information security risk management methods is provided in [51], which is based on the identification of key building blocks of the methods and their sequencing. The building blocks include: information discovery and collection, processing of collected information, decision making, decision implementation and communication.

2.4 Privacy Impact Assessment

There are several methods that specifically look into privacy risks and are usually called Privacy Impact Assessment (PIA). For instance, there are Privacy Impact Guidelines of the Treasury Board of Canada Secretariat [55] and PIA of the Information Commissioner's Office, United Kingdom [29]. PIA is a "systematic process for evaluating the potential effects on privacy of a project, initiative, or proposed system or scheme" [59]. It helps to identify and manage privacy risks for an organization that deals with personal data of its stakeholders. However, these methods usually do not attribute the events to people. Wright [59] states that PIA should be integrated into risk management along with other strategic planning tools.

In [60], Wright et al. provide a comparative analysis of PIA policies and methods of six countries, Australia, Canada, Ireland, New Zealand, United Kingdom and United States of America, explaining the effectiveness and limitations of each. The PIA in the six countries were analyzed against various criteria, some of which are: is more than a compliance check, is a process, identifies privacy risk, encourages publication of the PIA report, etc. The objective of the analysis has been to determine the best elements from the existing PIAs so that these could be used in the development of a European PIA policy and methodology.

The Treasury Board of Canada Secretariat [55] provides the guidelines for PIA. It is stated that the PIA guidelines are built on the "universal privacy principles identified in the Canadian Standards Association's Model Code for the Protection of Personal Information in addition to federal privacy legislation and policies". The main goal of PIA is to ensure effective communication of privacy risks. The PIA process consists of four steps, which are project initiation, data flow analysis, privacy analysis and privacy impact analysis report. The process is similar to that of risk management. The privacy analysis step consists of answering the questionnaires provided in the document which help in identifying privacy risks. Afterwards, qualitative estimation (low, medium and high) is used to represent the level of risks.

The Information Commissioner's Office, United Kingdom [29] developed a PIA handbook (version 2.0) that provides the background information and guidance on the PIA process. It is intended to help organizations that are involved in projects that might have potential privacy impacts. The term "project" may refer to any activity or function the organization is assessing, system, database, program, etc. It consists of five phases: preliminary phase, preparation phase, consultation and analysis phase(s), documentation phase, and review and audit phase. However, it does not provide guidance on how risk analysis is to be carried out and what tool or method should be used.

The review of the two PIAs of Canada and United Kingdom show that the risk analysis step (specifically risk estimation) needs be improved. This thesis explores an alternative notion of risk to provide an approach that facilitates better risk analysis by improving the identification, estimation and evaluation of risk events.

2.5 Risk Analysis and Game Theory

As the game theoretic approach provides a way to analyze the situations of conflict between the players, it helps to understand the behavior of real world adversaries [26]. This section presents some of the work on risk analysis and game theory.

These include [27] for estimating the reliability of a system; [22] and [13] for adversarial risk analysis; [19] for cybersecurity risk assessment and [10] and [12] for counterterrorism. Hausken [27] merges Probabilistic Risk Analysis (PRA) and game theory to add the missing behavioral dimension to PRA. Kardes et al. [36] state that PRA does not consider the strategies of the adversary and thus, suggest using the game theoretic approach. The use of game theory for risk analysis can improve the existing (adversarial) risk analysis approaches by developing the risk models using the concept of risk analysis, then utilizing game theory for optimizing the decision of the defender in consideration to the attacker's best response [22]. The importance of game theory for risk analysis is also emphasized by Bier et al. [13]. They state using game theory for risk analysis results in the consideration of the actions of intelligent and adaptive adversaries. A quantitative cybersecurity risk assessment approach called Quantitative Evaluation of Risk for Investment Efficient Strategies (QuERIES) uses game theory for constructing and evaluating the attack/protect model [18], [19]. In [12], for the protection of complex systems against intelligent and adaptable adversaries, risk and reliability analysis is combined with game theory.

2.6 Motivational Theory

The study of motivation is the study of action [23]. This thesis emphasizes on understanding the incentives or interests of the stakeholders during the risk analysis process. Thus, this section briefly introduces some of the work on motivational theory.

According to Eccles et al. [23] recent studies on motivation research focus on the "relation of beliefs, values and goals with action". They state that by focusing on beliefs, values and goal constructs of an individual, we can understand the reasons individuals decide to engage or not in different activities and also how these constructs are related to their achievement behavior. They categorize the motivational theories into four categories. These include theories that are based on expectancy, theories based on the reasons for engagement, theories integrating expectancy and value constructs, and theories integrating motivation and cognition.

According to Ajzen [1], in the theory of planned behavior, motivational factors influence the behavior and are captured by intention. This indicates "how hard people are willing to try, of how much of an effort they are planning to exert in order to perform the behavior" [1]. In [37], Leonard et al. proposed a taxonomy of motivation sources given as: intrinsic process, instrumental motivation, external self-concept, internal self-concept and goal internalization.

Chulef et al. [20] write "the goals an individual has - and the interactions among them - play a crucial role in understanding and predicting the behavior in which individuals engage". They provide a hierarchical taxonomy of human goals which is built on the constructs used in the motivational literature. Moreover, the taxonomy is empirically generated from a diverse sample of subjects rather than being based on the theoretical classification generated by the researchers. Researchers have also developed various taxonomies, e.g. taxonomy of top managers' goal [11] and taxonomy of motives of human motives (as the basis for the motives of intelligent agents) [47] (which builds on and addresses the limitations in [20]). The System-Action-Management (SAM) framework [39] is an example of the framework that models the influence of management factors on human behavior during risk analysis. Apart from including the risk analysis model at the system level, the framework integrates the decisions and actions of humans that affect the physical system, and then associates management factors to those decisions and actions. Motivational theory is an area of our research which has been briefly touched in the publications in Part II. Thus, we leave the detailed investigation on motivational theory for future work.

2.7 Bibliography

[1] AJZEN, I. The Theory of planned behaviour. *Organizational Behaviour and Human Decision Processes 50* (1991), 179–211.

[2] ALBERTS, C., AND DOROFEE, A. OCTAVE SM Threat Profiles. *Software Engineering Institute, Carnegie Mellon University* (2001).

[3] ALBERTS, C., AND DOROFEE, A. *Managing information security risks, The OCTAVE approach*. Addison Wesley, 2002. ISBN 0-321-11886-3.

[4] ALBERTS, C., DOROFEE, A., STEVENS, J., AND WOODY, C. *Introduction to the OCTAVE Approach*. Carnegie Mellon University, 2003.

[5] ASME Innovative Technologies Institute, LLC. *Risk Analysis and Management for Critical Asset Protection (RAMCAP): The Framework*, May 2006. Version 2.0.

[6] AS/NZS 4360. *Risk management*. AS/NZS, 2004.

[7] Aven, T. *Foundations of Risk Analysis: A Knowledge and Decision-Oriented Perspective*. John Wiley & Sons, Ltd, Chichester, UK, 2003. ISBN: 0-471-49548-4.

[8] Aven, T. The risk concept - historical and recent development trends. *Reliability Engineering & System Safety 99*, 0 (2012), 33 – 44. doi:10.1016/j.ress.2011.11.006.

[9] Aven, T. On the Meaning and Use of the Risk Appetite Concept. *Risk Analysis 33*, 3 (2013), 462–468. doi:10.1111/j.1539-6924.2012.01887.x.

[10] Banks, D. L., and Anderson, S. Combining Game Theory and Risk Analysis in Counterterrorism: A Smallpox Example. In *Statistical Methods in Counterterrorism*, A. Wilson, G. Wilson, and D. Olwell, Eds. Springer New York, 2006, pp. 9–22. doi:10.1007/0-387-35209-0_2.

[11] Bateman, T. S., O'Neill, H., and Kenworthy-U'Ren, A. A hierarchical taxonomy of top managers' goals. *Journal of Applied Psychology 87*, 6 (2002), 1134–1148. doi:10.1037/0021-9010.87.6.1134.

[12] Bier, V. Game-Theoretic and Reliability Methods in Counterterrorism and Security. In *Statistical Methods in Counterterrorism*, A. Wilson, G. Wilson, and D. Olwell, Eds. Springer New York, 2006.

[13] Bier, V. M., Cox, Jr., L. A., and Azaiez, M. N. Why Both Game Theory and Reliability Theory are Important in Defending Infrastructure against Intelligent Attacks. In *Game Theoretic Risk Analysis of Security Threats*, vol. 128 of *International Series in Operations Research & Management Science*. Springer US, 2009, ch. 1, pp. 1–11. doi:10.1007/978-0-387-87767-9_1.

[14] Bornman, G., and Labuschagne, L. A comparative framework for evaluating information security risk management methods. In *Information Security South Africa Conference* (2004).

[15] Braber, F., Brndeland, G., Dahl, H. E. I., Engan, I., Hogganvik, I., Lund, M. S., Solhaug, B., Stølen, K., and Vraalsen, F. *The CORAS Model-based Method for Security Risk Analysis*. SINTEF, Oslo, 2006.

[16] Braber, F., Hogganvik, I., Lund, M. S., Stølen, K., and Vraalsen, F. Model-based security analysis in seven steps — a guided tour to the CORAS method. *BT Technology Journal 25*, 1 (2007), 101–117. doi:10.1007/s10550-007-0013-9.

[17] Campbell, P. L., and Stamp, J. E. *A Classification Scheme for Risk Assessment Methods*. Sandia National Laboratories, August 2004. Sandia Report.

[18] Carin, L., Cybenko, G., and Hughes, J. Quantitative Evaluation of Risk for Investment Efficient Strategies in Cybersecurity: The QuERIES Methodology, September 2007. Approved for Public Release: AFRL/WS-07-2145.

[19] Carin, L., Cybenko, G., and Hughes, J. Cybersecurity Strategies: The QuERIES Methodology. *Computer 41* (2008), 20–26.

[20] Chulef, A., Read, S., and Walsh, D. A Hierarchical Taxonomy of Human Goals. *Motivation and Emotion 25*, 3 (2001), 191–232(42).

[21] Cox, Jr., L. A. Some limitations of "Risk = Threat x Vulnerability x Consequence" for risk analysis of terrorist attacks. *Risk Analysis 28*, 6 (2008), 1749–61.

[22] Cox, Jr., L. A. Game Theory and Risk Analysis. *Risk Analysis 29* (2009), 1062–1068. doi:10.1111/j.1539-6924.2009.01247.x.

[23] Eccles, J. S., and Wigfield, A. Motivational beliefs, values, and goals. *Annual review of psychology 53*, 1 (2002), 109–132.

[24] ENISA. Risk Management: Implementation principles and Inventories for Risk Management/Risk Assessment methods and tools. Tech. rep., ENISA, 2006.

[25] ETSI TS 102 165-1 V4.2.3 (2011-03). *Method and proforma for Threat, Risk, Vulnerability Analysis*. ESTI, 2011.

[26] Fricker, Jr, R. D. Game theory in an age of terrorism: How can statisticians contribute? In *Statistical Methods in Counterterrorism*, A. Wilson, G. Wilson, and D. Olwell, Eds. Springer New York, 2006, pp. 3–7.

[27] Hausken, K. Probabilistic Risk Analysis and Game Theory. *Risk Analysis 22*, 1 (2002), 17–27. doi:10.1111/0272-4332.t01-1-00002.

[28] Hillson, D. Extending the risk process to manage opportunities. *International Journal of Project Management 20*, 3 (2002), 235–240. doi:10.1016/S0263-7863(01)00074-6.

[29] Information Commissioner's Office (ICO). Privacy Impact Assessment Handbook, 2009. Version 2.0.

[30] ISACA. *The Risk IT Framework*, 2009.

[31] ISACA. *COBIT 5: A Business Framework for the Governance and Management of Enterprise IT*. IT Governance Institute, 2012.

[32] ISO 31000. *Risk Management – Principles and Guidelines*. ISO, 2009.

[33] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*, 1st ed. ISO/IEC, 2008.

[34] ISO/IEC Guide 73. *Risk management - Vocabulary* . ISO, 2009.

[35] Karabacak, B., and Sogukpinar, I. ISRAM: information security risk analysis method. *Computers & Security 24*, 2 (2005), 147–159. doi:10.1016/j.cose.2004.07.004.

[36] Kardes, E., and Hall, R. Survey of Literature on Strategic Decision Making in the Presence of Adversaries. CREATE Report, 2005.

[37] Leonard, N., Beauvais, L., and Scholl, R. Work motivation: The incorporation of self-concept-based processes. *Human Relations 52*, 8 (1999), 969–998.

[38] Lund, M. S., Solhaug, B., and Stølen, K. A Guided Tour of the CORAS Method. In *Model-Driven Risk Analysis*. Springer Berlin Heidelberg, 2011, pp. 23–43.

[39] Murphy, D. M., and Paté-Cornell, M. E. The SAM framework: Modeling the effects of management factors on human behavior in risk analysis. *Risk Analysis 16*, 4 (1996), 501–515.

[40] NIST. *NIST SP 800-39, Managing Information Security Risk - Organization, Mission, and Information System View*, 2011.

[41] NIST and U.S. Department of Commerce. *NIST SP 800-30 Revision 1, Guide for Conducting Risk Assessments*, September 2012.

[42] Nocco, B. W., and Stulz, R. M. Enterprise risk management: Theory and practice. *Journal of Applied Corporate Finance 18*, 4 (2006), 8–20.

[43] Olsson, R. In search of opportunity management: Is the risk management process enough? *International Journal of Project Management 25*, 8 (2007), 745–752. doi:10.1016/j.ijproman.2007.03.005.

[44] Paté-Cornell, E. Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering & System Safety 54*, 2-3 (1996), 95 – 111. Treatment of Aleatory and Epistemic Uncertainty. doi:10.1016/S0951-8320(96)00067-1.

[45] Rajbhandari, L. Consideration of Opportunity and Human Factors: Required Paradigm Shift for Information Security Risk Management. In *European Intelligence and Security Informatics Conference* (2013), IEEE, pp. 147–150. doi:10.1109/EISIC.2013.32.

[46] Raz, T., and Hillson, D. A Comparative Review of Risk Management Standards. *Risk Management: An International Journal 7*, 4 (2005), 53–66.

[47] Read, S. J., Talevich, J., Walsh, D. A., Chopra, G., and Iyer, R. A comprehensive taxonomy of human motives: a principled basis for the motives of intelligent agents. In *Intelligent Virtual Agents* (2010), Springer, pp. 35–41.

[48] Siegel, J. Introduction To OMG's Unified Modeling Language (UML). http://www.omg.org. [Online accessed: 8-2013].

[49] Siemens Enterprise. CRAMM. http://www.cramm.com/overview/howitworks.htm, 2011. [Online accessed: 04-2013].

[50] Slovic, P., Finucane, M., Peters, E., and MacGregor, D. G. Risk As Analysis and Risk As Feelings: Some Thoughts About Affect, Reason, Risk, and Rationality. *Risk Analysis 24*, 2 (2004), 311–322.

[51] Snekkenes, E. An Information Security Risk Management Research Menu. *Norsk informasjonssikkerhetskonferanse (NISK) 2012* (2012).

[52] Stoneburner, G., Goguen, A., and Feringa, A. *NIST SP 800-30, Risk Management Guide for Information Technology*. NIST, 2002.

[53] Syalim, A., Hori, Y., and Sakurai, K. Comparison of Risk Analysis Methods: Mehari, Magerit, NIST800-30 and Microsoft's Security Management Guide. In *Availability, Reliability and Security, 2009. ARES '09. International Conference on* (2009), pp. 726–731. doi:10.1109/ARES.2009.75.

[54] Thompson, K. M., Deisler, P. F., and Schwing, R. C. Interdisciplinary Vision: The First 25 Years of the Society for Risk Analysis (SRA), 1980-2005. *Risk Analysis 25*, 6 (2005), 1333–1386. doi:10.1111/j.1539-6924.2005.00702.x.

[55] Treasury Board of Canada Secretariat. Privacy Impact Assessment Guidelines: A Framework to Manage Privacy Risks Guidelines. http://www.tbs-sct.gc.ca, April 2012. [Online accessed: 1-2013].

[56] Vorster, A., and Labuschagne, L. A framework for comparing different information security risk analysis methodologies. In *Proceedings of the 2005 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries* (2005), SAICSIT '05, South African Institute for Computer Scientists and Information Technologists, pp. 95–103.

[57] Ward, S., and Chapman, C. Transforming project risk management into project uncertainty management. *International Journal of Project Management 21*, 2 (2003), 97–105. doi:10.1016/S0263-7863(01)00080-1.

[58] White, B. E. Enterprise Opportunity and Risk. *INCOSE Symposium, Orlando, FL* (2006).

[59] Wright, D. Should privacy impact assessments be mandatory? *Commun. ACM 54*, 8 (2011), 121–131. doi:10.1145/1978542.1978568.

[60] Wright, D., Finn, R., and Rodrigues, R. A Comparative Analysis of Privacy Impact Assessment in Six Countries. *Journal of Contemporary European Research 9*, 1 (2013), 160–180.

Chapter 3

Summary of Papers

This chapter presents a summary of the papers accomplished for the thesis. The chapter follows the structure of Figure 1.1. The research questions are assigned to the following research phases: feasibility study, method development, method practicality or feasibility, and theory development to enhance the method. For each of the research questions (RQ), the corresponding research papers are introduced.

3.1 Feasibility Study

Reliable data collection is an important part of effective risk analysis and management. Our main focus was to understand and identify the appropriate way of gathering data from subjects so that they can easily provide the needed information. To investigate this, we started with the hypothesis that game theory is suitable for risk analysis. The rationale behind this are, using game theory, we can determine how the subjects select their strategies in situations of interdependence, and further collect representative data on how they assess the values of the outcomes of incidents. This led us to the following research question.

RQ 1. To what extent can game theory be used for analyzing risks?

Papers [5] and [6] are relevant for this research question.

In paper [6], we identified the necessary steps for modeling a game to analyze risks faced by a risk taker using game theory. We discovered some of the primary issues related to the use of game theory in risk analysis which are as follows. Firstly, when using a game-theoretically inspired risk analysis process, the analysis can be based on preferences or values of benefit which the subjects can provide rather than relying on subjective probability. For example, instead of asking the user of a service 'How likely is it that you will benefit by providing your personal information to the service (assuming the service does not exploit your information) ?'. The risk analyst can set the question to the user as: 'How much would you benefit by providing your personal information to the service (assuming the service does not exploit your information) ?'. The subject can provide the answer in terms of monetary values or time. Then, with these obtained values of benefit from the subjects, the risk analyst can compute probabilities and expected outcomes to determine risk. In addition, the game-theoretically inspired risk analysis process was determined to be applicable in settings where no actuarial data is available. Thus, it was concluded that this may in turn increase the quality and appropriateness of the entire risk analysis process.

In order to provide a more comprehensive answer to this research question, we investigated to what extent game theory can be mapped and utilized for risk management. So far, no prior works provide the mapping between game-theoretically inspired risk management steps and classical risk management steps. In paper [5], we first identified features and critical elements for each of the methods (classical risk management and game theoretical approach). Subsequently, we identified steps where a correspondence is missing. From the existing classical risk management methods, we chose the ISO/IEC 27005:2008 standard [4] as it presents a clear explanation of the stages and terminologies of the risk management process. Besides, it is one of the widely used standards. We showed how all the steps of ISO/IEC 27005:2008 can be mapped to the steps of the game-theoretically

Table 3.1: Mapping between ISO/IEC 27005:2008 and Game Theoretic Approach

ISO/IEC 27005:2008 Process/ Terminology		Game Theoretic Step/ Terminology
Context establishment	Setting the basic criteria Defining the scope & boundaries Organization for information security risk management (ISRM)	Scenario investigation (scope definition & asset identification) Player identification (good & bad guys)
Risk identification	Identification of assets	Included in scenario investigation
	Identification of threats	Determine the strategies for the bad guys
	Identification of existing controls	Identify implemented controls i.e. 'do nothing' option for the good guys
	Identification of vulnerabilities	Options that can be exploited by threats. Included while determining the strategies for the bad guys.
	Identification of consequences	Identify how the players value multiple orthogonal aspects of outcomes. Identify the preferences.
Risk estimation	Assessment of consequences	Define scale & weight for comparing outcomes, & ranking preferences. Represent by payoff/ utility (assign values in each cell of the matrix).
	Assessment of incident likelihood	Computed probabilities for each of the strategies of both the players
	Level of risk estimation (list of risks with value levels assigned)	Expected outcome for each of the strategy of the bad guy is the risk for good guy & vice versa.
Risk evaluation	List of risks prioritized	Prioritize the expected outcome for both the players.
Risk treatment	Risk treatment options- risk reduction, retention, avoidance & transfer	Strategies (control measures) for the good guys; can be categorized into different options based on the computed probabilities.
	Residual risks	Expected outcome of the game
Risk acceptance	List of accepted risks based on the organization criteria	Strategies of the good guy (based on the organization criteria)
Risk Communication	Continual understanding of the organization's ISRM process & results.	Strategies of the good guy
Risk Monitoring & Review	Monitoring & review of risk factors Risk Management monitoring, reviewing & improving	Process is repeated as the players' options and their outcome valuation may change
Not Included		Information gained by the opponent
Not Included		Beliefs & incentives of the opponent
Not Included		Optimization of the strategies

inspired risk management process, whereas some of the game theoretical steps at best have a very limited existence in the standard as shown in Table 3.1. The game theoretic framework was appropriate for the entire risk management process and not just for risk analysis. The excluded game theoretical steps are: information gained by the adversary, beliefs and incentives of the adversary, and optimization of the strategies by the players. From a risk management perspective, these correspond to the lack of explicit consideration of the behavior of the adversary during the analysis.

3.2 Method Development

When modeling a game, taking a game theoretic perspective on risk analysis, it was tempting to start off by investigating issues such as 'is it a perfect or imperfect information game', 'is it a complete or incomplete information game', 'is it a repeated game' etc. Thus, even though we began with the hypothesis that game theory is suitable for risk analysis, we determined that some of the more fundamental areas should be investigated first like 'how

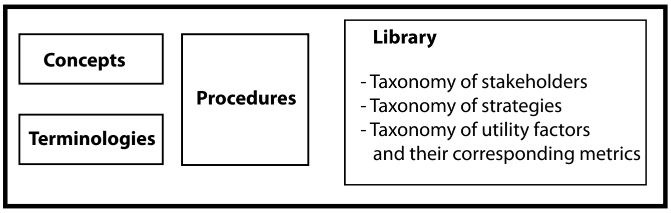

Concepts
Terminologies
Procedures
Library
- Taxonomy of stakeholders
- Taxonomy of strategies
- Taxonomy of utility factors and their corresponding metrics

Figure 3.1: The CIRA Method.

strong are the players' incentives to make the first move'. This led us to the following research question.

RQ 2. How can a risk analysis method be developed with an alternative notion of risk?

Paper [7] provides the answer to this research question.

Based on this new perspective of focusing on incentives of the stakeholders, the Conflicting Incentives Risk Analysis (CIRA) method was developed. The concepts, terminologies for understanding the method as well as a procedure for implementing it were established.

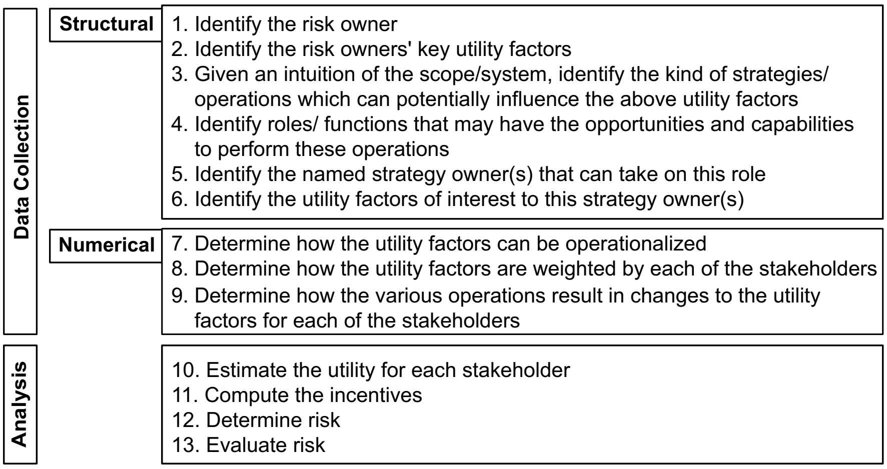
CIRA identifies stakeholders, their actions and perceived expected consequences that characterize the risk situation. We categorize the stakeholders into risk owner and strategy owner(s). The stakeholder, whose perspective is considered when performing the risk analysis, is a risk owner. On the other hand, a strategy owner is the stakeholder who is capable of triggering an action to increase his perceived benefit.

The key concept behind CIRA is that it replaces incident likelihood with incentives of the stakeholders. Risk is modeled in terms of conflicting incentives between the stakeholders in regards to the execution of actions. Moreover, the method focuses on human related risks. It also includes a library which is an extensible part of the method and its content is dynamic. As depicted in Figure 3.1, the library consists of the taxonomy of stakeholders, taxonomy of strategies, and taxonomy of utility factors and their corresponding metrics. The intention behind the library is to help the risk analyst by speeding up the data collection phase.

In CIRA, when the risk analyst interacts with the stakeholders, we assume that the stakeholders' preferences have been determined according to their perception of the available options and considerations about how the others decide to act. Game theory comes into play in our model, as it provides insight in understanding these strategic settings that influence the stakeholders' behavior. However, game theoretic modeling (i.e. constructing models of strategic settings) and computations are not used in our method.

We discovered that investigating the stakeholders' incentive to move first helps to understand the risks to the risk owner. We look into what motivates a stakeholder to take or not to take an action to increase his expected utility. Utility is the benefit as perceived by the corresponding stakeholder and it comprises of utility factors. Furthermore, each utility factor captures a specific aspect of utility e.g. wealth, privacy, reputation, legal compliance. Hence, utility can be estimated as the sum of weighted values for utility factors using aspects of Multi Attribute Utility Theory (MAUT) [1]. The utility factors and their values can be identified from past actions, surveys or interviews and research in psychology. Thus, it was determined that CIRA is applicable in situations where the utilities for the stakeholders can be estimated reasonably well by understanding the psychological perspective of the stakeholders.

The updated procedure of CIRA (as given in [8] and [9] from its older version in [7]) is depicted in Figure 3.2. After context establishment, the procedure consists of three phases:



1. Identify the risk owner
2. Identify the risk owners' key utility factors
3. Given an intuition of the scope/system, identify the kind of strategies/ operations which can potentially influence the above utility factors
4. Identify roles/ functions that may have the opportunities and capabilities to perform these operations
5. Identify the named strategy owner(s) that can take on this role
6. Identify the utility factors of interest to this strategy owner(s)
7. Determine how the utility factors can be operationalized
8. Determine how the utility factors are weighted by each of the stakeholders
9. Determine how the various operations result in changes to the utility factors for each of the stakeholders
10. Estimate the utility for each stakeholder
11. Compute the incentives
12. Determine risk
13. Evaluate risk

Figure 3.2: Procedure in CIRA.

the structural data collection phase (consisting of steps 1 to 6), the numerical data collection phase (consisting of steps 7 to 9) and the analysis phase (consisting of steps 9 to 13).

3.3 Method Practicality or Feasibility

Apart from building the theoretical concept and framework for the method, we believe practicality is equally important. After developing CIRA, it was essential to investigate the practicality of the method which led us to the following research question.

RQ 3. To what extent is the developed method feasible for analyzing risk in a real life non-trivial setting?

Papers [9] and [10] provide insight to this research question.

The initial idea was to use a Norwegian eGovernment service called MinID [3] as a case study for implementing CIRA. We determined that in a typical real life risk analysis when applying CIRA, it results in sensitive output as shown in Figure 10.1(a). Note that CIRA$_1$, CIRA$_2$ and CIRA$_3$ correspond to the structural data collection phase, numerical data collection phase and analysis phase of the CIRA method as depicted in Figure 3.2. We realized that this intrusive nature of the risk analysis process makes it difficult for practitioners (e.g. researchers, students) to gain access to scenarios from operational organizations for evaluating or training on risk analysis methods. Thus, the idea of using the Case Study Role Play (CSRP) approach was initiated. CSRP is an approach that is obtained from the integration of case study [13] , persona [2] and role play [12]. CSRP helps to establish a platform for doing risk management related research and training in a reasonable realistic environment as compared to an operational setting by ensuring that the time and resources needed to set up the required environment is low and can be anticipated. This platform removes some of the obstacle for the practitioners in regards to confidentiality and sensitivity issues with respect to the application of the method, red tape and the need of acquiring permissions in carrying out their tasks, etc. Thus, using CSRP, the practitioners can gain new knowledge about risk analysis methods.

The CSRP preparation phase consists of (1) determining the objective of the activity i.e. whether it is to mimic an operational setting for gaining knowledge about feasibility of a risk analysis method or to provide a real life like training platform to learn about the method, (2) selecting the organization appropriate for the identified activity, (3) familiarizing with the chosen method and (4) finally, designing and building the organization to

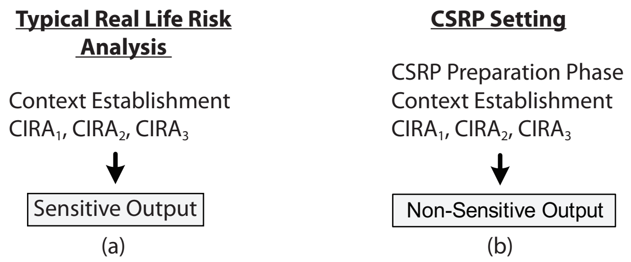


Figure 3.3: Steps and output of CIRA in (a) Typical real life risk analysis and (b) CSRP setting.

carry out the risk analysis activities. Other details on CSRP is explained in paper [10]. It was determined that when using CSRP with CIRA, the information extracted was non-sensitive as illustrated in Figure 3.3(b).

The CIRA method is illustrated for a hypothetical but realistic case study of a Norwegian eGovernment service using CSRP in paper [9]. The application of CSRP proved beneficial as it helped to analyze the risks faced by an end-user when using the eGovernment service without having to worry about the intrusive nature of the study and the hassle of getting access to an operational setting. The results from the case study helped us to explore the feasibility of CIRA in non-trivial settings.

3.4 Theory Development to Enhance the Method

Previously, CIRA had been investigated in analyzing threat risks [7], [9]. Threat risks are the risks facing the risk owner caused by the intentional execution of strategies by the strategy owner which results in loss for the risk owner and/or gain for himself. In the incentive graph as shown in Figure 3.4, these risk events fall below the X-axis in the boundary of the fourth quadrant. Note the details on the incentive graph is provided in paper [8].

Opportunity risks occur when there are actions that one can reasonably expect the strategy owner should take but for which he would have to take a loss in utility (or not have a sufficiently significant gain) while the risk owner would have the possibility of a gain [11]. Such situations arise where the risk events lie above the X-axis and fall in the boundary of the second quadrant.

One of the other areas was to investigate the treatment (response) measures for threat (opportunity) risks in CIRA. This led us to the following research questions.

RQ 4. How can we model opportunity risk in the developed method and how can the method be extended to risk management?

Paper [8] provides the answer to this research question.

We figured out that one of the serious constraints in most of the risk management methods is the identification and management of opportunity risk in the context of information security management. We modeled the theoretical concept of risk acceptance and rejection for opportunity risk in the context of CIRA. The risk owner faces an opportunity risk when there is a loss in utility or not a sufficiently significant gain for the strategy owner by triggering a strategy. This, thereby, leads to a potential failure on the strategy owners' side to trigger a strategy that the risk owner could reasonably expect that he should trigger. The definitions for risk acceptance and rejection for both threat and opportunity risks were formalized and some theorems establishing consequences of the definitions are introduced in [8].

In classical methods, risk is usually managed by reducing incident likelihood and consequence. It was determined that, in CIRA, risk treatment or response amounts to the

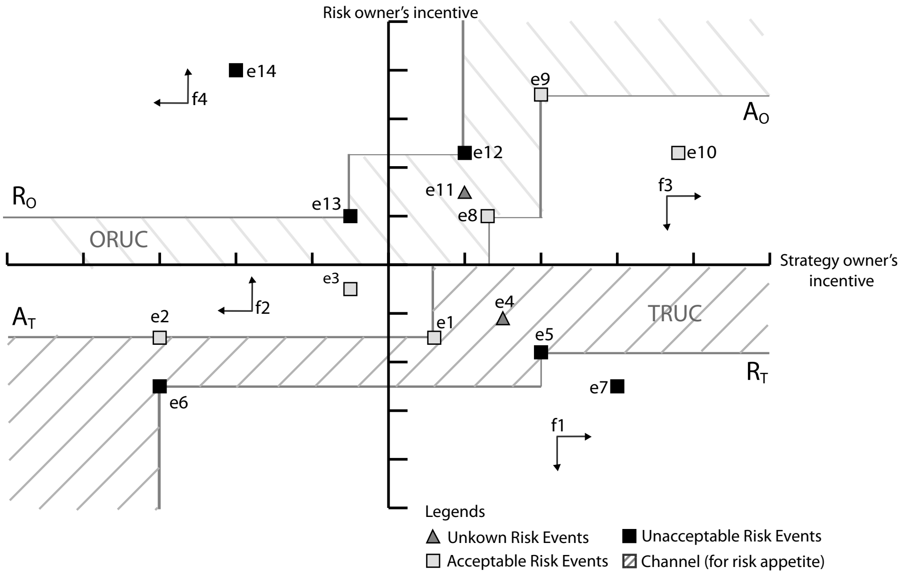


Figure 3.4: Risk Visualization in CIRA using the Incentive Graph.

modification of perceived utility caused by the strategies in question. In other words, a risk treatment or response measure aims to modify the weights that the stakeholders assign to the relevant utility factors or modify the incentives of the stakeholders. Thus, CIRA can be extended to a comprehensive Conflicting Incentives Risk Analysis and Management (CIRAM) method with the potential to enhance the overall risk management process.

3.5 Bibliography

[1] Clemen, R. T. *Making Hard Decision: An Introduction to Decision Analysis*, 2nd ed. Duxbury, 1996.

[2] Cooper, A. *The Inmates Are Running the Asylum*. Macmillan Publishing Co., Inc., Indianapolis, IN, USA, 1999.

[3] Difi (Direktoratet for forvaltning og IKT). MinID. http://minid.difi.no/minid/minid.php?lang=en. [Online accessed: 06-2012].

[4] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*, 1st ed. ISO/IEC, 2008.

[5] Rajbhandari, L., and Snekkenes, E. Mapping between Classical Risk Management and Game Theoretical Approaches. In *Communications and Multimedia Security*, vol. 7025 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2011, pp. 147–154. doi:10.1007/978-3-642-24712-5_12.

[6] Rajbhandari, L., and Snekkenes, E. Using Game Theory to Analyze Risk to Privacy: An Initial Insight. In *Privacy and Identity Management for Life*, vol. 352 of *IFIP Advances in Information and Communication Technology*. Springer Berlin Heidelberg, 2011, pp. 41–51. doi:10.1007/978-3-642-20769-3_4.

[7] Rajbhandari, L., and Snekkenes, E. Intended Actions: Risk Is Conflicting Incentives. In *Information Security*, vol. 7483 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 370–386.

[8] Rajbhandari, L., and Snekkenes, E. Risk Acceptance and Rejection for Threat and Opportunity Risks in Conflicting Incentives Risk Analysis. In *Trust, Privacy, and Security in Digital Business*, vol. 8058 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2013, pp. 124–136. doi:10.1007/978-3-642-40343-9_11.

[9] Rajbhandari, L., and Snekkenes, E. Using the Conflicting Incentives Risk Analysis Method. In *Security and Privacy Protection in Information Processing Systems*, vol. 405 of *IFIP Advances in Information and Communication Technology*. Springer Berlin Heidelberg, 2013, pp. 315–329. doi:10.1007/978-3-642-39218-4_24.

[10] Rajbhandari, L., and Snekkenes, E. A. Case Study Role Play for Risk Analysis Research and Training. In *Proceedings of the 10th International Workshop on Security in Information Systems*. SciTePress, 2013, pp. 12–23. doi:10.5220/0004599500120023.

[11] Snekkenes, E. Position Paper: Privacy Risk Analysis Is about Understanding Conflicting Incentives. In *Policies and Research in Identity Management*, vol. 396 of *IFIP Advances in Information and Communication Technology*. Springer Berlin Heidelberg, 2013, pp. 100–103.

[12] Yardley-Matwiejczuk, K. M. *Role play: theory and practice*. Sage Publications Limited, 1997.

[13] Yin, R. K. *Case Study Research: Design and Methods*, 4th ed., vol. 5 of *Applied Social Research Method Series*. Sage, 2009.

Chapter 4

Summary of Thesis Contributions

In this chapter, we summarize the main contributions of the thesis.

The main contribution of this thesis lies in the identification of limitations of ISO/IEC 27005:2008 [2], the development of the Conflicting Incentives Risk Analysis (CIRA) method, the establishment of the Case Study Role Play (CSRP) approach for risk analysis research and training, and conceptual development of CIRA in terms of including the theoretical concepts of risk acceptance and rejection for threat and opportunity risks, and extending it to risk management.

4.1 Limitations of ISO/IEC 27005:2008

One of the research contributions of the thesis is in the identification of the limitations of the ISO/IEC 27005:2008 [2] standard.

In ISO/IEC 27005:2008, risk is determined as the combination of likelihood and consequence. However, the challenge lies in obtaining credible likelihood estimates in relation to issues with human error and biases, unavailability of reliable statistical data for new and emerging systems, and the existing statistical data being irrelevant or insufficient as mentioned in Section 1.1.

From paper [3], by mapping game theoretical steps to risk management process in ISO/IEC 27005:2008, it was determined that game theoretically inspired risk management process can be integrated into ISO/IEC 27005:2008. However, we figured out that some of the game theoretical steps such as information gained by the adversary, beliefs and incentives of the adversary and optimization of the strategies by the players have limited existence in the standard. These correspond to the lack of explicit consideration of behavior of the adversary during the risk management process as pointed out in Section 3.1. Even though the standard looks at the motivations of the adversary to identify potential threat scenarios, it does not explicitly consider the human factors during the analysis phase. In addition, the standard overlooks opportunity risks for an organization.

4.2 A New Approach for Risk Analysis

This thesis contributes by developing a new approach for risk analysis, the CIRA method [4]. In CIRA, stakeholders, their actions and perceived expected consequences are identified which characterize the risk situation. Risk is defined as the subjective concern that an individual can feel towards the outcome of events. The principal concept behind CIRA, is instead of relying on the concept of incident likelihood, like in most of the traditional methods, the approach is based on the incentives of the stakeholders. The procedure provides a systematic way to conduct risk analysis.

Some of the consequences of using CIRA rather than the classical methods are:

1. In classical methods, risk is usually determined as the combination of likelihood and consequence; resulting in the unit of Ut^{-1}, where U represents utility and t represents time. On the other hand, in CIRA, risk is modeled in terms of conflicting incentives between the stakeholders and its unit is U^2.

2. Most of the classical methods focus on risks in relation to threats, failing to consider risks in relation to opportunity. However, both threat and opportunity risks are considered in CIRA (which is further explained in Section. 4.4).

3. In CIRA, the events are attributed to people. The method looks into who are involved, what are their incentives and what are their interests. Thus, it is applicable in settings where people are involved. However, in the classical approach, the events may or may not be attributed to people.

4. Usually, expert elicitation is carried out to collect data in the classical approach. In CIRA, as the perceived incentives of the stakeholder are considered, each relevant stakeholder is considered as an "expert". More precisely, each stakeholder is an expert at judging the value of the changes that a particular action has on himself.

5. The questions to ask the stakeholders when using the classical method and CIRA will be different. When using a classical method, the questions would be- "How bad can it get?, How likely is it to happen? Can something bad happen as often?". On the other hand, when using CIRA, the questions would be- "What utility factors are important to the stakeholder?, How important are these factors?".

6. In classical methods, a risk map (as in Risk IT [1]) or risk-level matrix (as in NIST 800-30 [7]) is often used to visually model or establish and communicate risk. In CIRA, the incentive graph is used to model and communicate risk.

7. In classical methods, the risk acceptance and rejection criteria are usually predetermined at the beginning of the risk analysis process. For instance, in the ISO/IEC 27005:2008 [2] standard, the risk acceptance criteria are determined before the threat and vulnerability discovery activity is carried out. This raises some concern on the credibility of the criteria as a decision maker will not have complete insight at the early stage of the risk analysis process. As emphasized in paper [5], it is questionable whether it is a good strategy to fix the risk acceptance criteria as an expected value before determining the relevant decision input parameters. In CIRA, risk acceptance and rejection criteria are determined at the end of the analysis phase after the decision input parameters are determined.

4.3 Case Study Role Play Approach

This thesis contributes by developing a CSRP approach [6] that helps to establish a platform for doing risk analysis related research and training in a reasonable realistic setting. Risk analysis and management related activities involve the collection and processing of confidential and sensitive information. This intrusive nature of these activities make it hard for the individuals to get access to an operational setting. They often need to acquire permissions to conduct the process, gain trust from the stakeholders and face delay due to red tape.

The main goal of CSRP is to provide a platform for conducting method specific research which help to validate a risk analysis method, determine its performance or provide training to students on a certain risk analysis method. The use of CSRP does not necessarily eliminate all the problems necessary for implementing a risk analysis method. However, it provides a partial solution by providing a platform for improving or gaining new knowledge about risk analysis methods. It also provides additional benefit by ensuring the time and resources needed to set up the required environment is low and predictable.

As mentioned before, CSRP may not be particularly applicable for discovering new unknown vulnerabilities or for determining actual risks in an operational system, unless the established platform or setting closely represents the real organization.

4.4 Introducing Risk Acceptance and Rejection for Threat and Opportunity Risks

This thesis contributes by presenting the theoretical concept of risk acceptance and rejection for both threat and opportunity risks in context of CIRA [5]. Threat risk is the concern that something undesirable might happen. On the other hand, opportunity risk is the concern that something desirable might not happen. In addition, we determined that identification and management of opportunity risk in the context of information security management is an area which is often overlooked.

Some of the other contributions of the thesis include a new way to visualize or model and communicate risk by using the incentive graph, introducing the risk treatment and response measures, for threat risks and opportunity risks respectively, hence, extending CIRA to risk management.

4.5 Bibliography

[1] ISACA. *The Risk IT Framework*, 2009.

[2] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*, 1st ed. ISO/IEC, 2008.

[3] Rajbhandari, L., and Snekkenes, E. Mapping between Classical Risk Management and Game Theoretical Approaches. In *Communications and Multimedia Security*, vol. 7025 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2011, pp. 147–154. doi:10.1007/978-3-642-24712-5_12.

[4] Rajbhandari, L., and Snekkenes, E. Intended Actions: Risk Is Conflicting Incentives. In *Information Security*, vol. 7483 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 370–386.

[5] Rajbhandari, L., and Snekkenes, E. Risk Acceptance and Rejection for Threat and Opportunity Risks in Conflicting Incentives Risk Analysis. In *Trust, Privacy, and Security in Digital Business*, vol. 8058 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2013, pp. 124–136. doi:10.1007/978-3-642-40343-9_11.

[6] Rajbhandari, L., and Snekkenes, E. A. Case Study Role Play for Risk Analysis Research and Training. In *Proceedings of the 10th International Workshop on Security in Information Systems*. SciTePress, 2013, pp. 12–23. doi:10.5220/0004599500120023.

[7] Stoneburner, G., Goguen, A., and Feringa, A. *NIST SP 800-30, Risk Management Guide for Information Technology*. NIST, 2002.

Chapter 5

Future Work

This chapter describes the potential issues for further research.

The Conflicting Incentives Risk Analysis (CIRA) method is in its early development phase so there is a large area for future work. We provide insight into some of the significant areas that can be investigated. In addition, the Case Study Role Play (CSRP) approach can also be further explored. This dissertation serves as a basis for discussions regarding these possibilities.

5.1 Conflicting Incentives Risk Analysis

This section provides some insight into additional development of CIRA.

5.1.1 Case Study

This thesis has evaluated the feasibility of CIRA by applying it to a non-trivial operational setting in [2]. The results show that it is feasible to analyze risk in non-trivial operational settings using CIRA. However, in order to explore, validate and improve CIRA, the method needs to be implemented with more case studies. The need for further work on this issue is evident and absolutely essential for the general use of the method.

5.1.2 Tool Development

For automation and support, we need to develop CIRA as a tool. By automating the steps in CIRA, it will be easy to implement the method. In addition, it may also assist in identifying areas where the method needs to be improved. The tool can significantly reduce the effort required by the risk analyst to conduct the analysis by saving up the time required for data collection, documentation and generating analysis reports. This, in turn, speeds up the decision making process and reduces cost.

5.1.3 Extension of Library

For the extension of the library, we need to collect definitions of utility factors and validate them. Moreover, we can also create a standard process for formulating metrics for the utility factors. Metrics have always been a challenge in information security. During our study, it was difficult to formulate the metrics for some of the utility factors as stated in [2]. Hence, it is essential to develop a taxonomy of these factors along with their operationalization. Similarly, the taxonomies of stakeholders and strategies need to be identified and/or developed. It is expected that the library will facilitate and speed up the data collection phase.

5.1.4 Comparison of CIRA with Classical Risk Analysis Methods and Privacy Impact Assessments

One of the questions raised during our work was how CIRA compares to the existing methods. Comparing CIRA with one or more of the classical risk analysis methods and Privacy

Impact Assessments (PIAs) will contribute towards understanding its effectiveness. The similarity and difference between them would have to be acknowledged in such works. Also, the effort must be made to ensure that the results of the analysis based on different methods are comparable so that it is useful in selection and decision situations. Furthermore, explaining how the classical risk analysis and PIA concepts manifest themselves in our method provides an interesting line of research.

5.1.5 Conceptual Development

The addressing of the above mentioned issues will help in the further conceptual development of CIRA. However, some of the significant works that would facilitate to create a robust foundation for CIRA are discussed in this section.

It is evident that uncertainties may be inevitably introduced in the method. More work is needed in capturing the uncertainties in relation to estimates. Some ways of dealing with this issue have been discussed in [1] and [2]. The approach could be to explicitly capture the uncertainty using techniques such as interval arithmetic, p-boxes or second order (subjective) probabilities instead of point values.

In CIRA, utility is assumed to be linear. The issues relating to the non-linearity of utility factor valuation needs to be investigated. Furthermore, work is needed to solidify the theoretical concepts of opportunity risk in the context of CIRA and the extension of CIRA to risk management.

Besides, an important part is the detail investigation on motivational theory as previously stated in Section 2.6 so as to better understand the incentives of the stakeholders.

5.2 Utilization of Case Study Role Play Approach

This section emphasizes further exploration of CSRP.

CSRP can facilitate the researchers with evaluating the feasibility of CIRA without having the need to get access to an operational setting as explained in [2] and [3]. However, more work is needed in exploring the CSRP setting with CIRA and also with other risk analysis methods. A possible usage of CSRP as stated in [3] can be to assess the performance of risk analysis methods by collecting performance data (e.g. risk analyst's time sheet, resource requirements, etc) of various methods. Then, those collected performance data can be compared allowing the best risk analysis method for a given setting to be selected. Further, its usage in an educational setting for e.g. to train students to use a risk analysis method can also be explored.

5.3 Bibliography

[1] RAJBHANDARI, L., AND SNEKKENES, E. Intended Actions: Risk Is Conflicting Incentives. In *Information Security*, vol. 7483 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 370–386.

[2] RAJBHANDARI, L., AND SNEKKENES, E. Using the Conflicting Incentives Risk Analysis Method. In *Security and Privacy Protection in Information Processing Systems*, vol. 405 of *IFIP Advances in Information and Communication Technology*. Springer Berlin Heidelberg, 2013, pp. 315–329. doi:10.1007/978-3-642-39218-4_24.

[3] RAJBHANDARI, L., AND SNEKKENES, E. A. Case Study Role Play for Risk Analysis Research and Training. In *Proceedings of the 10th International Workshop on Security in Information Systems*. SciTePress, 2013, pp. 12–23. doi:10.5220/0004599500120023.

Part II

Scientific Publications

Chapter 6

Using Game Theory to Analyze Risk to Privacy: An Initial Insight[1]

Abstract

Today, with the advancement of information technology, there is a growing risk to privacy as identity information is being used widely. This paper discusses some of the key issues related to the use of game theory in privacy risk analysis. Using game theory, risk analysis can be based on preferences or values of benefit which the subjects can provide rather than subjective probability. In addition, it can also be used in settings where no actuarial data is available. This may increase the quality and appropriateness of the overall risk analysis process. A simple privacy scenario between a user and an online bookstore is presented to provide an initial understanding of the concept.

Keywords: game theory, privacy, risk analysis

6.1 Introduction

Every individual has a right to the privacy of their personal information. People are dependent on information technology in their daily lives, which includes the risk of their personal information being misused, stolen or lost. The personal information of an individual might be collected and stored by government agencies, businesses and other individuals. These organizations and individuals might have the incentive to misuse such gained information, at least from the perspective of the individual.

In [1], Anderson stated that individuals produce as well as consume commodity information. There are growing problems of identity theft, tracking of an individual, personal information being used as a commodity and so on. Thus, there is a necessity to protect the privacy of information, perform risk analysis and evaluation for proper protection of the entire identity management infrastructure. According to the guidelines of ISO/IEC 27005, for information security risk assessment, risk identification, estimation and evaluation are necessary tasks [6].

In this paper, we suggest that instead of a classical risk analysis like Probabilistic Risk Analysis (PRA), we use a game theory based approach. The distinction between using PRA and Game theory for general risk analysis are shown in Table 6.1.

In PRA, the risk level is estimated by studying the likelihood and consequences of an event and assigning the probabilities in a quantitative or qualitative scale. Moreover, it can be considered as a single person game because the strategies followed by the opposing player or adversary are not considered. In [3], Bier has stated the challenges to PRA as subjective judgment and human error and performance.

With game theory we can consider settings where no actuarial data is available. Moreover, we do not have to rely on subjective probabilities. By obtaining the preferences or benefits from the subjects, we can compute the probabilities and outcomes to determine

[1] RAJBHANDARI, L., AND SNEKKENES, E. Using Game Theory to Analyze Risk to Privacy: An Initial Insight. In Privacy and Identity Management for Life, vol. 352 of IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg, 2011, pp. 41–51.

Table 6.1: Comparison of general Risk Analysis steps: Using PRA and Game theory

	Classical Risk Analysis	*Our proposal*
Risk Analysis	**PRA**	**Game theory**
Collect data	Ask for subjective probability or historical data	Ask for preferences or benefits
Compute risk	Compute risk (e.g. expected value)	Compute probability and expected outcome (e.g. mixed strategy Nash equilibrium)
Evaluate	Decide what to do	Decide what to do

the risk. We propose that it can be used in studying and evaluating the behavior of the players in privacy scenarios. It also allows for better audit as the outcomes can be verified at each incident.

In this paper, we will focus on two important issues - 'suitability of game theory for privacy risk analyses' and 'how the payoffs of the players are calculated'.

6.2 Overview of Game Theory

Game theory is a branch of applied mathematics proposed by John von Neumann and Oskar Morgenstern in paper [12]. It has been used in many fields [11] like economics, political science, evolutionary biology, information security and artificial intelligence. It is the study of the strategic interactions among rational players and their behavior [13], [10], which can be modeled in the form of a game.

For a game, the required four components are: the players, their strategies, payoffs and the information they have [9]. The players are the ones whose actions affect each other's payoffs. Whereas, a strategy is a plan of action that the player can take in response to the opponent's move. It is impossible to act against all the defensive attacks at all times [4]. Thus, it is important to find out the preferred strategies of the players.

The payoff of a particular player is affected by both the actions taken by him and the other player. The thing that matters is that the value of the payoff should be consistent throughout the game. According to Auda, besides ordering the preferences of the payoffs, the players can *'also express the ratio of the differences of the preferences'* on an interval scale called utility [2].

Players make decisions based on the gained information. According to the information they have, the game can be categorized into a perfect/imperfect game and a complete/incomplete game. A complete information game is one in which the players know about the strategies and payoffs of one another (vice versa for the incomplete information game). While, a game where at least a player has no knowledge about the previous actions, as a minimum of one other player is called the imperfect information game (vice versa for the perfect information game).

After determining the components, the game can be represented in the normal (strategic or matrix) form or extensive (tree) form. The normal form is usually used to represent static situations in which the players choose their actions simultaneously and independently. On the other hand, the extensive form is usually used to represent dynamic situations in which the players have some information about the choices of other players.

In a game, each player chooses among a set of strategies to maximize their utilities or payoffs based on the information they have. The choice of strategies can be determined by equilibrium solutions such as the pure strategy Nash equilibrium [8] and the mixed strategy Nash equilibrium, both named after John Forbes Nash.

A strategy profile is a Nash equilibrium if each player's chosen strategy is the best response to the strategies of the other [13] and the players have no incentive to unilaterally deviate. A mixed strategy is a probability distribution (mixing or randomization) of the players' pure strategies so that it makes the other player indifferent between their pure strategies [13], [5]. The equilibrium gives the outcome of the game [9].

6.3 Why Game Theory?

Today, whenever we have to provide our personal information for instance, while purchasing a ticket online, most of us wonder and are concerned about our information being collected. Some questions that usually pop into our minds are - 'Is our information being stored and if so, to what extent? Who gets access to the stored information? Are all the insiders having access to the stored information 'good'?' If we ask people how often they face risks by providing their personal information, like a credit card number, the answer would be in terms of probability which would be rather vague. However, if we ask them how much they would benefit by providing it, we can have an appropriate answer, for example, in terms of monetary values or time. Thus, with game theory, we can ask expressive questions that the people can answer. Based on these data, risk analysis can be carried out.

In addition, we can perform risk analysis more accurately if we place the situation in the form of a 'game'. If we consider a game of poker, the players are rational. They not only think about their action but also what the other players will do in return to their own particular move. Kardes and Hall state that Probabilistic Risk Analysis (PRA) does not consider the strategies of the adversary and thus, suggest using the game theoretic approach [7]. In the real world, we have to plan our moves considering the moves of the others, especially if the opponent is an adversary. By using game theory, we can find out how the players choose their strategies in different situations of interdependence. For instance, let us consider zero-sum and cooperative games. In a zero-sum game, gain to one player is a loss to another. Thus, each player takes individual actions to maximize their own outcome. In cooperative games, players cooperate or negotiate their actions for the benefit of each other.

Moreover, the adversary usually does not give up when his attempts have been defended; he rather uses different strategies. In a game theoretic setting, the benefits are based on outcomes and the incentives of the players are taken into account. Thus, game theory helps to explore the behavior of real world adversaries [4].

6.4 Scenario and Game Formulation

In this section, we will look at the scenario between a user and an online bookstore and the steps used to formulate the scenario in a game theoretic setting.

6.4.1 Scenario

The user subscribes to a service from an online bookstore. The online bookstore collects and stores additional private information such as book and magazine preferences. These preferences can then be used according to the privacy policy of the online bookstore to provide customized purchase recommendations.

When these recommendations are selected, they generate additional sales for the online bookstore. Also, these recommendations are beneficial to the user, as they save him valuable time. However, it is somewhat tempting for the online bookstore to breach the agreed privacy policy by providing these additional preferences of the user to third parties to be utilized for marketing. This third party marketing incurs additional costs for the user, mostly in terms of time wasting activities like advertisements. However, at the initial stage, the online bookstore cannot determine whether the given information is genuine or fake.

6.4.2 Game Formulation

For formulating the game, we take into account of the following assumptions- it is a complete information game but of imperfect information. The game is of complete information as we assume both the user and the online bookstore know about the strategies and outcomes of one other. It is of imperfect information as we stipulate that they have no information about the previous action taken by the other player when they have to make their decision. Moreover, it is a one shot game between the user and the online bookstore as a single interaction between them is considered and their actions are taken to be simultaneous and independent.

We now explain the strategies of the players and how the data are collected to estimate the payoffs. We then represent the game in the normal form.

6.4.2.1 Players.

It is a two player game, between the user and the online bookstore. We assume that both the players are intelligent and rational. They have the incentive to win or to optimize their payoffs.

6.4.2.2 Strategies.

We will use a set of simple strategies for this two player non-cooperative game between a user and an online bookstore. The user has the choice to either provide his genuine or fake personal information, knowing the possibility of his information being sold. The strategies of the user are given by {GiveGenuineInfo, GiveFakeInfo}.

The online bookstore either exploits the personal information of the user by selling it to third parties or does not exploit and uses it for its own internal purpose, given by {Exploit, NotExploit}.

6.4.2.3 Payoff.

For obtaining the payoffs of the players, we collect the data and then estimate it.

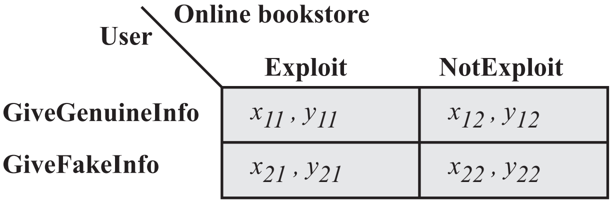
1. **Data collection:** We have assumed the values for the user and the online bookstore as shown in Table 6.2. However, it can be collected by conducting experiments and surveys. The values are in hours, reflecting saved or lost time. A positive value represents the hours saved while a negative value represents the hours lost. The profit corresponds to the hours of work saved. The variables 'a' to 'h' are used to represent the cells. The values of the user and the online bookstore are explained below.

   For the user : If the online bookstore uses the user's preferences and personal information for its own internal purpose according to the policy, we assume that the user saves an equivalent of an hour, if he had provided genuine personal information, and 0.1 hours if he had given fake information. However, if the online bookstore sells the information to third parties, the user wastes time dealing with the sale attempts. Thus, we assume that the user loses an hour if he had provided genuine personal information, and 0.01 hours if he had provided fake information.

   For the online bookstore : Similarly, when the bookstore uses the user's personal information for its own internal purpose according to the policy, we assume it saves an hour if the user had provided genuine information whereas, it loses 0.01 hours dealing with the fake information of the user. However, when it violates the privacy policy and sells the information to third parties, it saves 0.5 hours in case the user had provided genuine information and loses 0.2 hours in case of fake information. We stipulate that it is not possible to assess if private information is fake or not before

Table 6.2: Assumed saved or lost hours for the user and online bookstore.

	For user		For online bookstore	
Information provided by the user	**Genuine**	**Fake**	**Genuine**	**Fake**
The online bookstore usage of information for its internal purpose	(a) 1	(b) 0.1	(c) 1	(d) -0.01
The online bookstore usage of information by selling it to third parties	(e) -1	(f) -0.01	(g) 0.5	(h) -0.2



User \ Online bookstore	**Exploit**	**NotExploit**
GiveGenuineInfo	x_{11}, y_{11}	x_{12}, y_{12}
GiveFakeInfo	x_{21}, y_{21}	x_{22}, y_{22}

Figure 6.1: Normal form representation of the scenario.

the information sale is finalized.

2. **Estimation:** We can represent the game in a two players normal form as shown in Figure 6.1. The first value of each cell given by x_{ij} is the payoff of the user while the second value given by y_{ij} is the payoff of the online bookstore. Here, $i = 1$ to n, $j = 1$ to n and $n =$ number of players. Each value of the cell is explained and estimated below, along with stating how much each of the players influences the outcome.

   The payoffs are in utility, estimated using the assumed values of hours from Table 6.2. We have to keep in mind that when the online bookstore exploits the information, it uses the information for its own internal purpose as well as to gain profit by selling it to third parties.

   The first strategy profile **(GiveGenuineInfo, Exploit)** states that the user provides the personal information genuinely, while the online bookstore exploits it by selling it to third parties. Now, we will calculate the values of the payoff for this particular strategy profile-
   x_{11}: Even though the user will benefit from the service, he will have to waste time dealing with advertisements and sale attempts by third parties, incurred from the exploitation by the online bookstore. The user payoff is obtained by adding the cell values of online bookstore usage of information for its internal purpose (a) and usage by selling it to third parties (e). As mentioned earlier, when the bookstore exploits the information, it uses the data for its own purpose and also sells it to the third parties. Thus, the user's payoff is given by: $x_{11} = a + e = 0$.

   y_{11}: The online bookstore will be able to utilize and exploit the user's personal information both for legitimate and unauthorized usage. However, it will not lose time.

Thus, y_{11} is obtained by summing the cell values of online bookstore usage of information for its internal purpose (c) and usage by selling it to third parties (g) i.e. $y_{11} = c + g = 1.5$.

The payoffs for the strategy profile **(GiveGenuineInfo, NotExploit)** is estimated as given below-
x_{12}: As the user provides his genuine information, he will receive a customized service in accordance with the agreed privacy policy and save time by utilizing the recommendations. Thus, x_{12} equals the cell value 'a', which is obtained as the user provides genuine information and the online bookstore uses the information for its internal purpose i.e. $x_{12} = a = 1$.

y_{12}: The online bookstore will be able to utilize personal data to offer an improved service in accordance with the agreed privacy policy and will not lose time. Thus, y_{12} equals the cell value 'c', which is obtained as the user provides genuine information and the online bookstore uses the information for its internal purpose i.e. $y_{12} = c = 1$.

The payoffs for the strategy profile **(GiveFakeInfo, Exploit)** is estimated as given below-
x_{21}: The online bookstore will try to exploit the data, but later on, will discover that the data was incorrect. The user will only have some limited benefits from the service but will not lose time dealing with the sale attempts from third parties. Thus, x_{21} is obtained by summing the cell values of the online bookstore usage of information for its internal purpose (b) and usage by selling it to third parties (f) i.e. $x_{21} = b + f = 0.09$.

y_{21}: The fake data provided by the user may only be discovered by the online bookstore at a later stage, for example, at the time when the fake information is to be used to generate profit. The online bookstore will receive limited benefits from the interaction and lose time dealing with the fake data. Thus, the online bookstore's payoff is obtained by summing the cell values of the online bookstore usage of information for its internal purpose (d) and usage by selling it to third parties (h) i.e. $y_{21} = d + h = -0.21$.

The payoffs for the strategy profile **(GiveFakeInfo, NotExploit)** is estimated as given below-
x_{22}: The online bookstore will try to use this fake information given by the user to provide a customized service. However, the user will not receive any benefits and saves less time, as the improved service generated from fake data will be irrelevant. Thus, the user's payoff equals the cell value 'b', which is obtained as the user provides the fake information and the online bookstore uses the information for its internal purpose i.e. $x_{22} = b = 0.1$.

y_{22}: As the user provides fake information, the online bookstore will not be able to provide a customized service, resulting in reduced future sales. Moreover, it will lose time dealing with the fake data. Thus, the online bookstore's payoff equals the cell value 'd', which is obtained as the user provides the fake information and the online bookstore uses the information for its internal purpose i.e. $y_{22} = d = -0.01$.

The normal form representation with the estimated payoffs is given in Figure 6.2.

		Online bookstore	
	User	*q*	*1-q*
		Exploit	**NotExploit**
p	**GiveGenuineInfo**	0 , 1.5	1 , 1
1-p	**GiveFakeInfo**	0.09 , -0.21	0.1 , -0.01

Figure 6.2: Normal form representation of the scenario with estimated payoffs.

6.5 Game Solution

6.5.1 Pure/Mixed Strategy Nash Equilibrium

Using the above payoffs, we found that the game has no pure strategy Nash equilibrium as the players do not agree on a particular strategy profile. However, we can always find the mixed strategy Nash equilibrium.

For obtaining the mixed strategy Nash equilibrium, we will use the calculation as explained in [13](p. 123).We assume that the user plays the strategies GiveGenuineInfo and GiveFakeInfo with probabilities p and 1-p respectively, for $(0 \leq p \leq 1)$. After the calculation, we get $p = 0.29$. Thus, the user plays with the mixed strategy $(0.29, 0.71)$. This means that the user provides genuine information with a 0.29 probability and fake information with a 0.71 probability when playing this game.

Similarly, assume that the online bookstore plays the strategies Exploit and NotExploit with probabilities q and $1 - q$ respectively, for $(0 \leq q \leq 1)$. After the calculation, we obtain the mixed strategy as $(0.91, 0.09)$ for the strategy profile (Exploit, NotExploit). Hence, with this mixed strategy we can know the probabilities with which each of the players will choose a particular strategy.

6.5.2 Expected Outcome

We can represent the game in the normal form with the matrix A. Then, a_{ij} represents each cell of the matrix. In case of a two player game, the expected outcome of the game using mixed strategy to each player is given by

$$\sum_{i=1}^{k} \sum_{j=1}^{l} p_i \, q_j \, a_{ij} \quad . \tag{6.1}$$

where,
i- number of strategies of player 1 (user) $(1 \leq i \leq k)$,
j- number of strategies of player 2 (online bookstore) $(1 \leq j \leq l)$,
p_i- probabilities with which player 1 plays each of his strategies
$(0 \leq p_i \leq 1)$, $\sum p_i = 1$,
q_j- probabilities with which player 2 plays each of his strategies
$(0 \leq q_j \leq 1)$, $\sum q_j = 1$.

By using (6.1) and substituting the values of p and q, the expected outcome of the game for the user and the online bookstore is 0.09 and 0.28 respectively. We can conclude that by playing this game, the online bookstore benefits more than the user.

The overall values of the expected outcome that the players get by playing each of the strategies can also be estimated which are given in Figure 6.3. The benefit to each of the player can be based on these outcomes.

Expected outcome			0.25	0.03	Sum: 0.28
			q = 0.91	1-q = 0.09	
		Online bookstore User	Exploit	NotExploit	
0.03	p = 0.29	GiveGenuineInfo	0, 1.5	1, 1	
0.06	1-p = 0.71	GiveFakeInfo	0.09, -0.21	0.1, -0.01	

Sum: 0.09

Figure 6.3: Normal form representation along with the probabilities and expected outcomes.

6.6 Discussion

We formulated the scenario in the form of a strategic game. We used the concept of mixed strategy Nash equilibrium to compute the probabilities with which the players play each of their strategies, the expected outcome the players gain by playing each of the strategies and also the expected outcome of the game for each player.

Risk analysis can then be based on these computed probabilities and outcomes. However, the following two issues needs to be considered-

The first issue is the preference of the players. It is important to understand the uncertainties related to the preferences of the players in any game. The players might think differently, which may lead them to choosing a different strategy than the equilibrium. Some of the questions that need to be taken into account are -

1. Does the user know what the online bookstore prefers and how he orders the preferences and vice versa?

2. What are the consequences if the two players play 'different games' i.e. it differs in the perception of outcome?

The second is obtaining appropriate data by conducting experiments, interviews and surveys.

In addition, this scenario in a real world situation is usually of partial information. The user knows the exact value of his own 'saved/lost' time. The online bookstore knows the distribution of saved/lost time of the user from the population of all users. However, the online bookstore cannot guess the exact value because, at a given instant, it does not know with which user it is playing the game while the saved/lost time of the online bookstore is known by all users.

6.7 Conclusion

We can conclude that, with game theory, risk analysis can be based on the computed expected outcomes and probabilities rather than relying on subjective probability. For demonstrating this, we have considered a simple scenario between the online bookstore and its user. Moreover, we have explained how the data can be collected for estimating the payoffs.

The present study provides a starting point for further research. We will conduct a survey for gathering data as the next step. Further, the main objective of the research will be to incorporate the use of game theory in real world privacy scenarios besides the theoretical details.

Acknowledgment.

The work reported in this paper is part of the PETweb II project sponsored by The Research Council of Norway under grant 193030/S10.

6.8 Bibliography

[1] Anderson, H. The Privacy Gambit: Toward a Game Theoretic Approach to International Data Protection. *Vanderbilt Journal of Entertainment and Technology Law 9*, 1 (2006).

[2] Auda, D. Game Theory in Strategy Development of Reliability and Risk Management. In *Reliability and Maintainability Symposium* (Jan 2007), pp. 467–472. doi:10.1109/RAMS.2007.328082.

[3] Bier, V. Challenges to the Acceptance of Probabilistic Risk Analysis. *Risk Analysis 19* (1999), 703–710.

[4] Fricker, Jr, R. D. Game theory in an age of terrorism: How can statisticians contribute? In *Statistical Methods in Counterterrorism*, A. Wilson, G. Wilson, and D. Olwell, Eds. Springer New York, 2006, pp. 3–7.

[5] Fudenberg, D., and Tirole, J. *Game theory*. MIT Press, Cambridge, MA, 1991.

[6] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*, 1st ed. ISO/IEC, 2008.

[7] Kardes, E., and Hall, R. Survey of Literature on Strategic Decision Making in the Presence of Adversaries. CREATE Report, 2005.

[8] Nash, J. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences of the United States of America 36* (1950), 48–49.

[9] Rasmusen, E. *Games and Information: An Introduction to Game Theory*, 4th ed. Wiley-Blackwell, 2006. Indiana University.

[10] Ross, D. Game Theory. *The Stanford Encyclopedia of Philosophy* (2010). http://plato.stanford.edu/archives/fall2010/entries/game-theory/.

[11] Shoham, Y. Computer science and game theory. *Commun. ACM 51* (2008), 74–79. doi:http://doi.acm.org/10.1145/1378704.1378721.

[12] von Neumann, J., and Morgenstern, O. *Theory of Games and Economic Behavior*. Princeton University Press, 1944. Princeton, NJ.

[13] Watson, J. *Strategy : An Introduction to Game Theory*, 2nd ed. W. W. Norton & Company, 2008.

Chapter 7

Mapping between Classical Risk Management and Game Theoretical Approaches[1]

Abstract

In a typical classical risk assessment approach, the probabilities are usually guessed and not much guidance is provided on how to get the probabilities right. When coming up with probabilities, people are generally not well calibrated. History may not always be a very good teacher. Hence, in this paper, we explain how game theory can be integrated into classical risk management. Game theory puts emphasis on collecting representative data on how stakeholders assess the values of the outcomes of incidents rather than collecting the likelihood or probability of incident scenarios for future events that may not be stochastic. We describe how it can be mapped and utilized for risk management by relating a game theoretically inspired risk management process to ISO/IEC 27005. This shows how all the steps of classical risk management can be mapped to steps in the game theoretical model, however, some of the game theoretical steps at best have a very limited existence in ISO/IEC 27005.

Keywords: Game theory, Risk management, Equilibrium, Strategies

7.1 Introduction

There are many classical risk management approaches and standards [2], [19] like NIST 800-30 [17], RiskIT [8], ISO/IEC 27005 [9] and CORAS [13]. For this paper, we consider the ISO/IEC 27005 [9] standard as it provides a clear description of the stages and terminologies of the risk management process.

In a typical classical risk assessment approach, the probabilities are usually guessed and not much guidance is provided on how to get the probabilities right. When coming up with probabilities, people are generally not well calibrated. Besides, history may not always be a very good teacher. The hypothesis of the paper is: 'Gathering representative probabilities for future events that may not be stochastic, is difficult. We claim it is a lot easier to obtain representative data on how stakeholders assess the values of the outcomes of events/incidents.' In a game theoretic approach, probabilities are obtained from the actual computation and analysis. Moreover, the strategy (mitigation measure to reduce risk) can be determined with respect to the opponent's strategy. When the risks are estimated more accurately, the effectiveness of the overall risk management approach increases.

The main contribution of this paper is to show that game theory can be integrated into classical risk management. For this, we provide a clear structure of both the classical risk management and game theoretical approaches. The intention is to enable the readers to have a better understanding of both methods. We then describe how it can be mapped by relating a game theoretically inspired risk management process to ISO/IEC 27005. This

[1] RAJBHANDARI, L., AND SNEKKENES, E. Mapping between Classical Risk Management and Game Theoretical Approaches. In Communications and Multimedia Security, vol. 7025 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2011, pp. 147-154.

shows how all the steps of ISO/IEC 27005 can be mapped to the steps in the game theoretical model; although some of the game theoretical steps at best have a very limited existence in ISO/IEC 27005.

The remainder of this paper is structured as follows. In Section 7.2, we present the state of the art and a summary of contributions. In Section 7.3, we first compare the top level perspectives of classical risk management and game theory. We then provide a more detailed mapping between the two approaches, identifying issues where a correspondence is missing. In Section 7.4, we discuss our findings. Conclusion and future work are given in Section 7.5.

7.2 State of the Art

The classical risk management approaches takes the perspective of the single player (individual, system, etc.) for which the risk analysis is being carried out. For example, in Probabilistic Risk Analysis (PRA), people and their actions and reactions are not given much importance [7]. Thus, Hausken [7] puts forward the way of merging PRA and game theory taking into account that, in risk assessment, the actions of the people affect each other. In addition, most of the classical risk assessment approaches are inclined to be rather subjective as the value of the probabilities of threats are either assumed or based on historical data. Taleb [18] has provided examples of Black Swan incidents that cannot be predicted accurately based on historical data.

In game theory, the incentives of the players are taken into consideration which is important in understanding the underlying motives for their actions. Liu and Zang [12] put forward the incentive-based modeling approach in order to understand attacker intent, objectives and strategies. Anderson and Moore [1] also state the importance of incentives, as misaligned or bad incentives usually cause security failure.

Game theory helps to explore the behavior of real-world adversaries [6]. Cox has stated that, by using game theory, the adversarial risk analysis can be improved [5] as the actions of the attacker, which were regarded as random variables and judged from the defender's perspective, can be computed. QuERIES, a quantitative cybersecurity risk assessment approach, uses game theory for constructing and evaluating the attack/protect model [3], [4].

While there are many papers discussing the use of game theory for specific application areas [4], [11], [16], we are aware of no works that integrate a risk management framework such as ISO/IEC 27005 and game theory.

7.3 Mapping between ISO/IEC 27005 and Game Theoretic Approach

In this section, we first compare the top level perspectives of classical risk management and game theory. We then provide a more detailed mapping between the two approaches, identifying issues where a correspondence is missing.

7.3.1 A Top Level Comparison

As stated above, there are many classical risk management approaches. To apply these approaches a clear understanding of the terminology and the overall process flow is necessary. We consider the risk management steps of the ISO/IEC 27005 [9] standard which is depicted in Figure 7.1 (a). These steps can be iterated until the results are satisfactory. The input and output for each of these steps are given in ISO/IEC 27005 [9].

Game theory helps us to understand how the strategic interactions and interdependence among the rational players influence the outcomes they gain [20], [15]. The steps that we have identified are given in Figure 7.1 (b). For each of the steps, we provide a

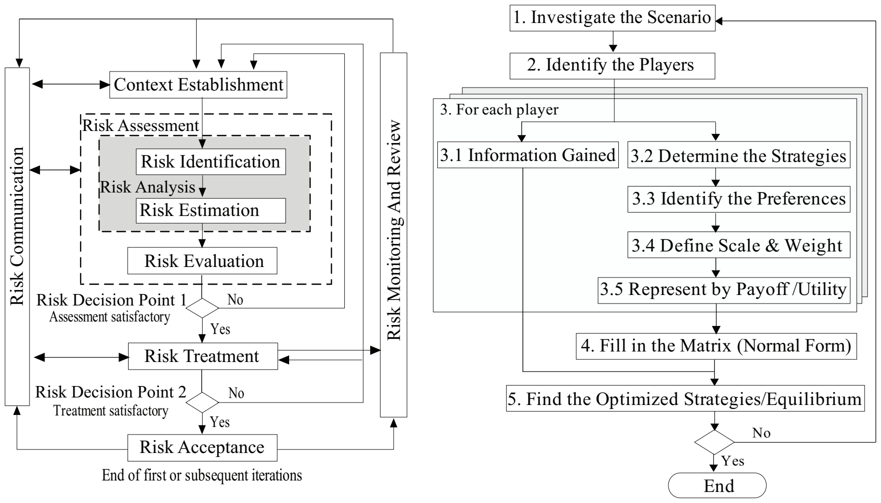
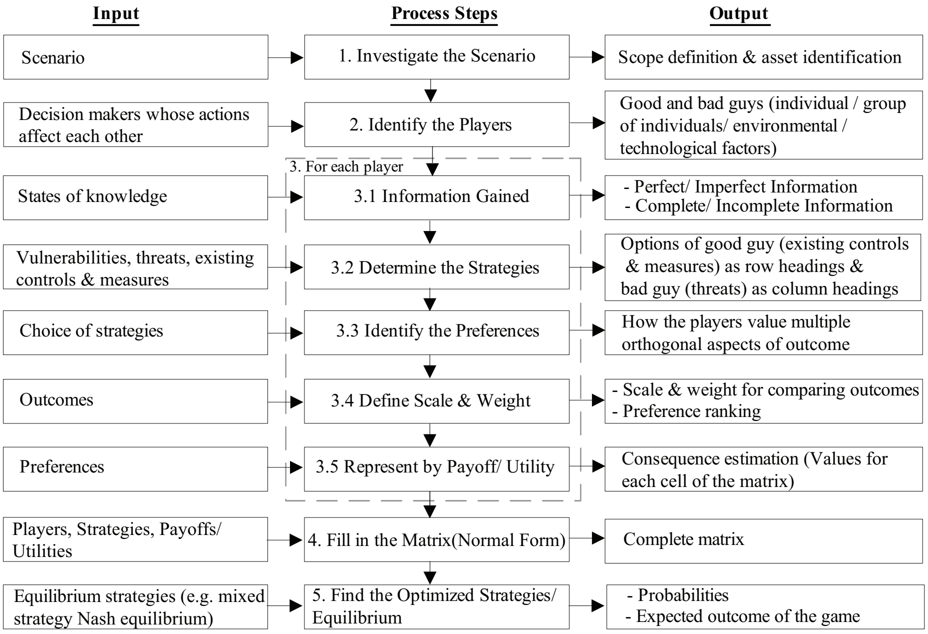


Figure 7.1: (a) Information Security Risk Management Process (taken from [9]) (b) Game Theoretical Steps



Input	Process Steps	Output
Scenario	1. Investigate the Scenario	Scope definition & asset identification
Decision makers whose actions affect each other	2. Identify the Players	Good and bad guys (individual / group of individuals/ environmental / technological factors)
	3. For each player	
States of knowledge	3.1 Information Gained	- Perfect/ Imperfect Information - Complete/ Incomplete Information
Vulnerabilities, threats, existing controls & measures	3.2 Determine the Strategies	Options of good guy (existing controls & measures) as row headings & bad guy (threats) as column headings
Choice of strategies	3.3 Identify the Preferences	How the players value multiple orthogonal aspects of outcome
Outcomes	3.4 Define Scale & Weight	- Scale & weight for comparing outcomes - Preference ranking
Preferences	3.5 Represent by Payoff/ Utility	Consequence estimation (Values for each cell of the matrix)
Players, Strategies, Payoffs/ Utilities	4. Fill in the Matrix(Normal Form)	Complete matrix
Equilibrium strategies (e.g. mixed strategy Nash equilibrium)	5. Find the Optimized Strategies/ Equilibrium	- Probabilities - Expected outcome of the game

Figure 7.2: Input and Output for Game Theoretical Steps

short description. In addition, Figure 7.2 depicts the input and output for each of the game theoretical steps.

1. The definition of scope of interest and assets that needs to be protected are identified by investigating the scenario.

2. Players whose actions affect each other are identified. The players are inherently good or bad, and who is 'good' or 'bad' depends on the perspective of the risk analyst. If the players show seemingly irrational behavior, this can be explained by at least two alternatives: (1) given the analyst (or objective) valuation of utility, it is

the players (irrational) reasoning that explains the irrational behavior; (2) the players have a different notion of utility than the risk analyst, but this notion of utility is (partially) unknown to the analyst. For the purpose of this paper, we choose the second alternative.

3. Once the players are identified, for each player we need to determine-

   3.1 Information they have when they make a decision.

   3.2 Strategies or options related to the actions of the players to overcome the threats or to gain opportunities.

   3.3 Preferences of the players, which can be obtained by asking how they value the outcomes, as the choice of each option results in an outcome. It is conceivable that players value multiple orthogonal aspects of outcome (e.g. cash, trust, reputation and legal compliance). Thus, in many cases, it may be desirable to model outcomes as vectors.

   3.4 Scale and weight should be defined so that the various outcomes can be compared. We can then rank the order of the preferences.

   3.5 These preferences can then be represented by numbers which are known as payoffs/ utilities. Higher payoffs represent more preferred outcomes. The values are assigned considering the players' motivation, capabilities (e.g. resources to implement or defend the attack) and experiences. The players in general have the incentive to maximize their payoff.

4. The scenario can then be formulated in the normal (strategic) form.

5. The optimum strategies for each player can be identified. The combination of optimum or best strategies chosen by the players is the equilibrium and this specifies the outcome of the game to the players [14].

The process is repeated as the players' options and their outcome valuation may change. Moreover, in the long run, the entire process should be repeated for effective risk management.

7.3.2 Mapping Individual Steps

Table 7.1 shows the result of the mapping between the risk management process of the ISO/IEC 27005 standard and the game theoretical steps. For each of the process of the ISO/IEC 27005 standard, the corresponding game theoretical steps are stated. The comparison is solely based on what is provided (process steps and terminologies) in the ISO/IEC 27005 standard. Both approaches are iterated until the result of the assessment is satisfactory.

The mapping shows that all the steps of ISO/IEC 27005 can be mapped to game theory. On the other hand, we have identified that some of the game theoretical steps like information gained, beliefs and incentives of the opposing players and optimization of the strategies by the players are not included in ISO/IEC 27005.

7.4 Discussion

In classical risk management, risk is calculated as a *'combination of the likelihood of an event and its consequence'* [10]. The limitations in this approach are: (1) Probability is difficult to assess as the underlying process may not be stochastic. Even if the process is stochastic, lack of historical data makes the parameters of the distribution difficult to estimate. Moreover, it is not appropriate and rather subjective to use the historical data in some of the

Table 7.1: Mapping between ISO/IEC 27005:2008 and Game Theoretic Approach

ISO/IEC 27005:2008 Process/ Terminology		**Game Theoretic Step/ Terminology**
Context establishment	Setting the basic criteria Defining the scope & boundaries Organization for information security risk management (ISRM)	Scenario investigation (scope definition & asset identification) Player identification (good & bad guys)
Risk identification	Identification of assets	Included in scenario investigation
	Identification of threats	Determine the strategies for the bad guys
	Identification of existing controls	Identify implemented controls i.e. 'do nothing' option for the good guys
	Identification of vulnerabilities	Options that can be exploited by threats. Included while determining the strategies for the bad guys.
	Identification of consequences	Identify how the players value multiple orthogonal aspects of outcomes. Identify the preferences.
Risk estimation	Assessment of consequences	Define scale & weight for comparing outcomes, & ranking preferences. Represent by payoff/ utility (assign values in each cell of the matrix).
	Assessment of incident likelihood	Computed probabilities for each of the strategies of both the players
	Level of risk estimation (list of risks with value levels assigned)	Expected outcome for each of the strategy of the bad guy is the risk for good guy & vice versa.
Risk evaluation	List of risks prioritized	Prioritize the expected outcome for both the players.
Risk treatment	Risk treatment options- risk reduction, retention, avoidance & transfer	Strategies (control measures) for the good guys; can be categorized into different options based on the computed probabilities.
	Residual risks	Expected outcome of the game
Risk acceptance	List of accepted risks based on the organization criteria	Strategies of the good guy (based on the organization criteria)
Risk Communication	Continual understanding of the organization's ISRM process & results.	Strategies of the good guy
Risk Monitoring & Review	Monitoring & review of risk factors Risk Management monitoring, reviewing & improving	Process is repeated as the players' options and their outcome valuation may change
Not Included		Information gained by the opponent
Not Included		Beliefs & incentives of the opponent
Not Included		Optimization of the strategies

situations, for example in estimating the risk of a terrorist attack, war or extreme events (Black Swan events). (2) Probability also depends largely on the risk analyst's perception or expert elicitation. People are generally not well calibrated. Thus, it is subjective in most of the cases. (3) The beliefs and incentives of the opponent are not considered. These limitations might result in inappropriate choices and decisions, which can be overcome by using game theory.

The benefits of using game theory for risk management are: (1) The quality of data collected is likely to be better as no actuarial data is needed. It focuses on incentives, capabilities and experiences of the players rather than asking an expert for historically based probabilities. (2) Expert judgment on collected data can be audited as we can determine and investigate how the players assess the values of the outcomes, what information is available to them, and whether they are utility optimizing or not taking into account the strategies of the opponent. (3) Probabilities are obtained from the actual computation and analysis. However, some of the limitations related to this approach are the players' limited knowledge about their own outcome(s) and the outcomes of others, and strategic uncertainty.

ISO/IEC 27005 takes the perspective of the organization for which the risk assessment is being carried out and thus, the information gained, beliefs and incentives of the adversaries and optimization of the strategies by the players are not included. Game theory is compatible with classical risk management and can be integrated into ISO/IEC 27005. This integration will provide the risk analyst additional guidance on what issues to address in his analysis and how more auditable probability estimates can be obtained. This integration also shows that game theoretic framework can be used for the entire risk management process and not just for risk analysis.

7.5 Conclusion and Future Work

Clear structure for both the classical risk management and game theoretical approaches have been presented. The mapping shows that game theoretically inspired risk management process can be integrated into ISO/IEC 27005. With game theory, we can obtain representative data on how stakeholders assess the value of outcomes of incidents rather than collecting the probability of incident scenarios for future events that may not be stochastic. Moreover, game theory is a rigorous method for computing probability and also the risk analyst can achieve additional guidance on how more auditable probability estimates can be obtained. However, some steps of game theory are not included in the current version of ISO/IEC 27005.

For future work, the above approach will be explored with a comprehensive case study and extended to the iterative aspect of risk management. Moreover, we will investigate the feasibility of adopting our ideas in the context of ISO 31000.

Acknowledgment: The work reported in this paper is part of the PETweb II project sponsored by The Research Council of Norway under grant 193030/S10. We would also like to thank the anonymous reviewers for their valuable comments and suggestions.

7.6 Bibliography

[1] Anderson, R., and Moore, T. Information Security Economics - and Beyond. In *In Proceedings of the 27th annual International Crytology Conference on Advances in Cryptology CRYPTO'07* (2007), Springer- Verlag, pp. 68–91. doi:10.1007/978-3-540-74143-5_5.

[2] Campbell, P. L., and Stamp, J. E. *A Classification Scheme for Risk Assessment Methods*. Sandia National Laboratories, August 2004. Sandia Report.

[3] Carin, L., Cybenko, G., and Hughes, J. Quantitative Evaluation of Risk for Investment Efficient Strategies in Cybersecurity: The QuERIES Methodology, September 2007. Approved for Public Release: AFRL/WS-07-2145.

[4] Carin, L., Cybenko, G., and Hughes, J. Cybersecurity Strategies: The QuERIES Methodology. *Computer 41* (2008), 20–26.

[5] Cox, Jr., L. A. Game Theory and Risk Analysis. *Risk Analysis 29* (2009), 1062–1068. doi:10.1111/j.1539-6924.2009.01247.x.

[6] Fricker, Jr, R. D. Game theory in an age of terrorism: How can statisticians contribute? In *Statistical Methods in Counterterrorism*, A. Wilson, G. Wilson, and D. Olwell, Eds. Springer New York, 2006, pp. 3–7.

[7] Hausken, K. Probabilistic Risk Analysis and Game Theory. *Risk Analysis 22*, 1 (2002), 17–27. doi:10.1111/0272-4332.t01-1-00002.

[8] ISACA. The Risk IT Framework. http://www.isaca.org, 2009.

[9] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*, 1st ed. ISO/IEC, 2008.

[10] ISO/IEC Guide 73. *Risk management - Vocabulary - Guidelines for use in standards*. ISO/IEC, 2002.

[11] Jormakka, J., and Mölsä, J. V. E. Modelling Information Warfare as a Game. *Journal of Information Warfare 4*, 2 (2005), 12–25.

[12] Liu, P., and Zang, W. Incentive-based modeling and inference of attacker intent, objectives, and strategies. In *Proceedings of the 10th ACM conference on Computer and communications security* (New York, NY, USA, 2003), CCS '03, ACM, pp. 179–189. doi:http://doi.acm.org/10.1145/948109.948135.

[13] Lund, M. S., Solhaug, B., and Stølen, K. A Guided Tour of the CORAS Method. In *Model-Driven Risk Analysis*. Springer Berlin Heidelberg, 2011, pp. 23–43.

[14] Rasmusen, E. *Games and Information: An Introduction to Game Theory*, 4th ed. Wiley-Blackwell, 2006. Indiana University.

[15] Ross, D. Game Theory. *The Stanford Encyclopedia of Philosophy* (2010). http://plato.stanford.edu/archives/fall2010/entries/game-theory/.

[16] Roy, S., Ellis, C., Shiva, S., Dasgupta, D., Shandilya, V., and Wu, Q. A Survey of Game Theory as Applied to Network Security. In *43rd Hawaii International Conference on System Sciences (HICSS)* (January 2010), pp. 1 –10. doi:10.1109/HICSS.2010.35.

[17] Stoneburner, G., Goguen, A., and Feringa, A. *NIST SP 800-30, Risk Management Guide for Information Technology*. NIST, 2002.

[18] Taleb, N. N. *The Black Swan: The Impact of the Highly Improbable*, 2nd ed. Random House Trade Paperbacks, May 2010.

[19] Vorster, A., and Labuschagne, L. A framework for comparing different information security risk analysis methodologies. In *Proceedings of the 2005 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries* (2005), SAICSIT '05, South African Institute for Computer Scientists and Information Technologists, pp. 95–103.

[20] Watson, J. *Strategy : An Introduction to Game Theory*, 2nd ed. W. W. Norton & Company, 2008.

Chapter 8

Intended Actions: Risk is Conflicting Incentives[1]

Abstract

Most methods for risk analysis take the view that risk is a combination of consequence and likelihood. Often, this is translated to an expert elicitation activity where likelihood is interpreted as (qualitative/ subjective) probabilities or rates. However, for cases where there is little data to validate probability or rate claims, this approach breaks down. In our Conflicting Incentives Risk Analysis (CIRA) method, we model risks in terms of conflicting incentives where risk analyst subjective probabilities are traded for stakeholder perceived incentives. The objective of CIRA is to provide an approach in which the input parameters can be audited more easily. The main contribution of this paper is to show how ideas from game theory, economics, psychology, and decision theory can be combined to yield a risk analysis process. In CIRA, risk magnitude is related to the magnitude of changes to perceived utility caused by potential state changes. This setting can be modeled by a one shot game where we investigate the degree of desirability the players perceive potential changes to have.

Keywords: Game theory, Risk analysis, risk, conflicting incentives, intended actions

8.1 Introduction

One of the key objectives of risk analysis is to provide insight suitable for deciding if risk exposure needs to be changed. That is, if a mitigation action is needed, or risk exposure may be increased. Most methods for risk analysis (including the ISO standard 27005 [20], NIST 800-30 [31], COBIT [21], CORAS [8]), take the view that risk is a combination of consequence and likelihood. Often, this is translated to an expert elicitation activity where likelihood is interpreted as (qualitative/ subjective) probabilities or rates. However, for cases where there is little data to validate probability or rate claims, this approach breaks down. Besides the use of subjective judgment, the other challenge to Probabilistic Risk Analysis (PRA) is handling of human performance and error [5]. Studies have shown that experts rely on heuristics in making judgments (decision) which might result in biases and errors [33, 29]. In addition, people are generally not well calibrated at estimating probabilities [29]. Taleb [32] has provided examples of incidents (called Black Swans) that cannot be accurately predicted based on the historical data. Thus, the questions is: What information relating to uncertainty can one reasonably expect to be able to collect reliably, and how should this information be framed so that it is ‘auditable’? Clearly, one approach could be to explicitly capture the uncertainty using techniques such as interval analysis, p-boxes or second order (subjective) probabilities. In this paper, we take a different view.

The objective of our Conflicting Incentives Risk Analysis (CIRA) method is to provide an approach where ‘probability affecting’ input parameters that can cause risks to go from ‘acceptable’ to ‘unacceptable’ can be audited more easily. In CIRA, we model risks in terms

[1] RAJBHANDARI, L., AND SNEKKENES, E. Intended Actions: Risk Is Conflicting Incentives. In Information Security, vol. 7483 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2012, pp. 370-386.

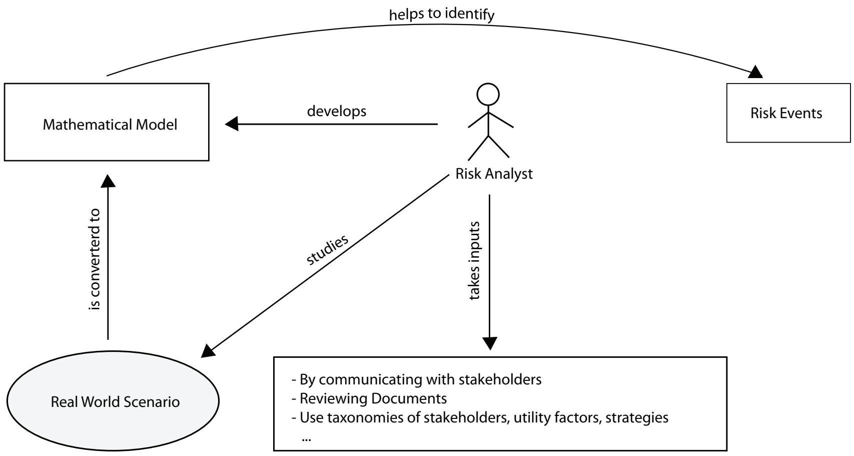


Figure 8.1: The Risk Analysis Process.

of conflicting incentives where risk analyst subjective probabilities are traded for stakeholder perceived incentives. These incentives are by necessity subjective. After all, it is these subjective perceptions that will influence the stakeholder actions. However, the risk analyst should aim to collect these incentives objectively without being influenced by his own subjective judgment. The key contribution of this paper is to show how ideas from game theory, economics, psychology, and decision theory can be combined to yield a risk analysis method.

When doing risk analysis, the risk analyst studies the real world scenario and makes the abstraction 'a mathematical model' as depicted in Figure 8.1. In reference to this model, the analyst then comes up with a 'hypothetical scenario' which he communicates with the stakeholders during workshops, interviews and/or surveys. He then collects data by gathering responses from the stakeholders, reviewing documents, using existing taxonomies of stakeholders, strategies, etc. Finally, he makes an estimation based on the model and identifies the risk events. However, the analyst needs to consider the validity of the model. That is, the abstraction should capture the key features of the real world such that deductions made from the abstraction provide a close approximation of what also will hold in the real world.

The CIRA method includes various concepts that are used to build the method, terminologies for understanding the method and procedures for implementing the method. The library includes the taxonomy of stakeholders, taxonomy of strategies, and taxonomy of utility factors and its corresponding metrics. Thus, the library can help the analyst to identify the relevant stakeholders, and for each stakeholder, identify their utility factors and its corresponding metrics, and strategies. The content of the library is dynamic, i.e. existing content can be updated and new content can be added. For instance, the choice of the metric for a utility factor such as privacy might differ depending on stakeholders and application. Thus, with the library, the risk analyst can be provided with a selection of possible definitions of metrics capturing privacy that he can select from or modify.

In CIRA, the risk owner is the stakeholder whose perspective we take when doing our analysis. We focus on risks at the managerial level rather than the technical level. We use aspects of Multi Criteria Decision Analysis (MCDA) and Multi-attribute Utility Theory (MAUT) [11, 34, 14] to estimate stakeholder utility. We model stakeholder incentives (i.e. changes in utilities) as the sum of weighted values for utility factors such as privacy, satisfaction etc. One important motivation for the work reported in this paper is that in classical risk analysis, risk having high consequence and low probability is problematic unless there is strong evidence that the likelihood is small. Verification of low probability is difficult and requires patience (i.e. waiting a long time for nothing to happen). On the other hand,

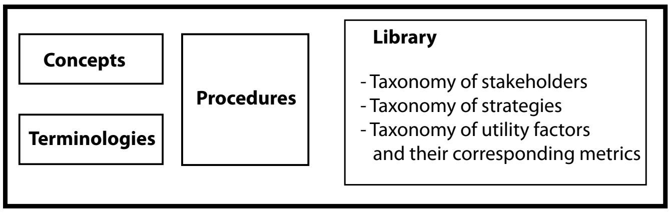

Concepts
Terminologies
Procedures
Library
- Taxonomy of stakeholders
- Taxonomy of strategies
- Taxonomy of utility factors
and their corresponding metrics

Figure 8.2: The CIRA Method.

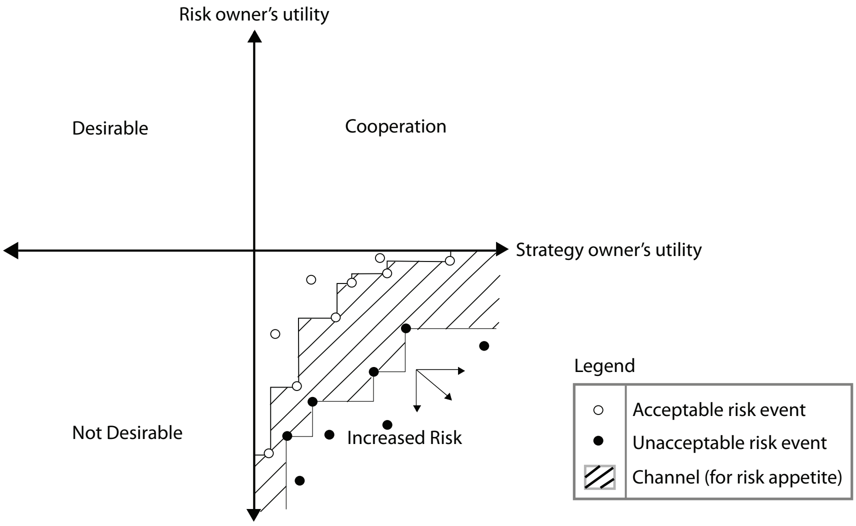


Figure 8.3: Incentive graph.

utility factors can be identified from past actions, surveys/ interviews and research in psychology. It is well known that many of the incidents are caused by psychological motives. Thus, we want to include psychological motives like revenge or desire for recognition in our methodology. However, documenting that a stakeholder perceives that he has little to gain by forcing an incident will often require less patience than verifying that the probability of the incident is small.

In our model, a strategy is an action that may modify the value of the utility factors. The strategy owner is the stakeholder that is in the position to trigger the execution of the strategy. Figure 8.3 depicts a scatter plot of the utilities of the outcome of some example strategies, where the strategy owner (risk owner) utility is plotted on the x (y) axis. Each quadrant represents various situations for the stakeholders: (1) 'cooperation', where both the stakeholders have the incentive to cooperate as they both gain positive utility, (2) 'desirable' situation for the risk owner as he gets positive utility and the converse for the strategy owner, (3) 'not desirable' situation for the stakeholders- as they both get negative utility, they both want to move out of this quadrant and (4) risky situation for the risk owner as he gets negative utility while the strategy owner gets positive utility.

In this paper, we focus on the fourth quadrant where each dot represents a particular *risk event*, caused by the 'intentional' execution of a strategy (intended action) by the strategy owner. Intended actions include all the adversarial activities, not just internal issues within the organization. Other risk events caused by unintended/ accidental behavior of the strategy owner, environmental risks, technical failures, etc will be investigated in future work. We argue that the underlying cause of risk is in fact the 'conflicting incentives' situation itself. In the case of intentional behavior of the strategy owner, we have a conflicting incentive situation when a strategy owner can cause a transition that will yield a gain for himself and a loss for the risk owner. This setting can be modeled by a one shot game where we investigate the degree of desirability the stakeholders perceive potential changes to have. The degree of desirability (undesirability) can be expressed in terms of positive (negative) change in utility. That is the greater the positive change in utility, the greater the degree of desirability. The idea being that risk is the combination of the degree of desirability that motivates the strategy owner to send the risk owner to an undesirable state and the magnitude of this undesirability. The risk for the risk owner increases when there is a decrease in his utility or an increase in the utility of the strategy owner or any combination of the former two. Other details regarding risk appetite will be explained later in Section 8.4.

The remainder of this paper is structured as follows. Related work is given in Section 8.2. In Section 8.3, we first give a brief overview of the steps of the CIRA method, then we explain the SNS scenario and finally our method by means of a running example of the interaction between two stakeholders in Social Networking Services (SNS): the data subject and the data controller. Even though CIRA can address multiple players game, for the purpose of explaining our approach we chose a simplified example including two players. In Section 8.4 and Section 8.5, we discuss our findings and future work respectively. Conclusion is given in Section 8.6.

8.2 Related Work

Researchers have put forward various definitions of risk. According to Slovic et al. [30], risk can be considered in three basic ways: risk as feelings (proposed by Lowenstein et al. [24]), risk as analysis and risk as politics. Risk as feeling consists of individuals' "fast, instinctive, and intuitive reactions to danger"; risk as analysis covers the scientific approach to risk such as risk assessment approaches; risk as politics comes into effect when there is a clash between the former two [30].

As mentioned earlier there are many classical risk management approaches such as the ISO standard 27005 [20], NIST 800-30 [31], COBIT [21] and CORAS [8]. For risk assessment, apart from estimating risk as the combination of consequence and likelihood as in the methods given above, risk is also determined as the combination of threat, vulnerability and consequence (e.g. in [3]). In [12], Cox has shown the limitations of estimating risk as the combination of threat, vulnerability and consequence.

Work on risk analysis and game theory includes: [18] for estimating reliability of a system, [13, 7] for adversarial risk analysis, [10] for cybersecurity risk assessment and [4, 6] for counterterrorism. Cox [13] states that the use of game theory for risk analysis can improve the existing (adversarial) risk analysis approaches. According to Cox, this can be achieved by developing the risk models using the concept of risk analysis, then utilizing game theory for optimizing the decision of the defender in consideration to the attacker's best response. Bier et al. [7] emphasize the importance of game theory for risk analysis as it considers the actions of intelligent and adaptive adversaries. In [6], risk and reliability analysis is combined with game theory for the protection of complex systems against intelligent and adaptable adversaries. As game theoretic approach provides a way to analyze the situations of conflict between the players, it helps to understand the behavior of real world adversaries [16]. Furthermore, in [27], it was shown that game theoretically inspired risk management process can be integrated into the ISO 27005 standard.

Further, research has been done on understanding the incentives of the stakeholders which can enhance the risk analysis process. Liu et al. [23] proposed the incentive based modeling approach in order to model and understand attacker intent, objectives and strategies. Anderson et al. [2] emphasize the importance of incentives, as misaligned or bad incentives usually cause security failure, i.e. gives rise to risk.

8.3 Conflicting Incentives Risk Analysis (CIRA)

In this section, we first give a brief overview of the steps involved in CIRA. We then explain the social networking services (SNS) scenario. Finally, we explain our method by means of a running example of SNS.

8.3.1 Summary of the CIRA Method

The overview of CIRA method is as follows:

1. Identify the stakeholders.
2. For each stakeholder,
   - 2.1. identify the utility factors. For each identified utility factor, determine the scale, measurement procedure, explain the underlying assumptions if any and assign weights as perceived by the respective stakeholder.
   - 2.2. identify the strategies to be considered.
3. Model a one shot game, then
   - 3.1. determine the final value of the utility factors after the strategies of the players are executed.
   - 3.2. estimate the utility for each of the strategies for each of the players.
4. Compute the incentives (i.e. changes in utilities) for each of the strategies for each of the players.
5. Determine risk by investigating each of the strategies with respect to the sign and magnitude of the changes determined in 4.

8.3.2 The Social Networking Services (SNS) Scenario

To explain our method, we model a setting between two stakeholders in SNS: the data subject (i.e. user) and the data controller (i.e. SNS provider). The data controller collects personal information of the data subject to provide the service. However, it is well known that this information can be used either to provide general advertisements (ads) in compliance with the privacy policy (not exploiting the personal information of the data subject) or by facilitating targeted ads (exploiting the personal information for commercial purposes). Even though most prominent SNS allow third parties to define the types of profiles they are targeting, to which the social network then delivers targeted ads. The SNS might change their policy to sell data to third parties.

8.3.3 Explaining the CIRA Method

We now explain our method by means of a running example of SNS.

1. Identify the Stakeholders.

The first step is to identify the stakeholders i.e. identify the risk owner along with the ‘n’ number of strategy owners. The risk owner is assumed to be player 0. In our example, the stakeholders are the data subject (risk owner) and the data controller (strategy owner).

2. Identify the Utility Factors and Strategies.

For each stakeholder, we need to iteratively identify the utility factors and the strategies.

2.1. Identify the Utility Factors. For each stakeholder, identify the associated collection of utility factors that provides the player with utility.

There are many definitions of assets (a term used in classical risk analysis usually meaning anything that has value to the organization) which lead to confusion [22]. Thus, in order to avoid this confusion and in the context of our method we use the term *utility factors*. All the assets can be cast to our notion of utility factors. Utility factors depend on the perspective of the player.

In our SNS example, we assume the utility factors for the data subject to be having privacy and to be satisfied (includes availability, support: responsiveness and effectiveness, service completeness) from using the service. Similarly, we assume that the data controller is only concerned about gaining marginal net profit and having a good reputation (includes experience of others and own experience). Now, for each of the above utility factors, we determine: the scale including the semantics of values, the measurement procedure (incorporating the idea) from behavioral economics [26, 9, 28] and the underlying assumptions if any. Then the weights are assigned to the utility factors as perceived by the respective player using aspects of Multi Criteria Decision Analysis (MCDA) techniques [14]. We assume that correlation among factors are captured in the assignment of weights.

Recall that the metrics used in our example are not prescribed, but belong to the extensible part of our method. The given metrics are just examples of what a metric might look like rather than metrics that have been validated for use in a specific real life risk analysis.

Privacy : First we need to decide if we need a projected or historic privacy metric. This decides if we will count actual incidents or make use of expected/projected number of incidents based on knowledge relating to data subject, data controller and third party behavior. We decide to construct a projected privacy metric.

Scale: Privacy (scale: %) is defined to be

$$1/(1+N) \tag{8.1}$$

where N is the number of times that a private data object is directly or indirectly utilized (e.g. read, copied or used as input to a decision process) in a way that is in conflict with a data usage purpose statement that is understood and accepted by the data subject. These numbers come from the analysis of the scenario, directly and indirectly caused by the events triggered by the various stakeholders. Note that we make the simplifying assumption that all data objects have the same sensitivity and that all breaches of use are equally serious. The privacy incident count is not to be restricted to incidents directly caused by the data controller, but also includes breaches initiated by a third party having obtained the sensitive data as a result of a privacy incident. If required, we can easily introduce separate categories relating usage breach and data object category. The above privacy metric is relative to a single subject. However, we may modify this metric to measure the commitment of the data controller to process data in a privacy friendly way by counting the incidents relating to all data subjects. Depending on the setting, it may be relevant to limit the counting of N to a specific time period (e.g. per month, per year). In the case that N cannot be uniquely defined, e.g. because we only get to know about a subset of the privacy incidents, we end up with a lower bound on N.

Measurement procedure:

1. Identify all data objects having a privacy requirement.
2. Identify all physical/logical locations where each of these data object instances are stored over the lifetime of the data (i.e. until all copies of the data have been deleted).
3. Identify what purpose is understood and accepted by the relevant subject for each of the data objects.
4. Over the lifetime of the data, count the number of projected privacy incidents.

In many cases it might be difficult to provide the number of data locations with any precision. However, in most cases, we will be able to provide upper and lower bounds. We can then use e.g. interval arithmetic [19] when computing the metric.

Satisfaction : There are many issues relating to customer satisfaction (see e.g. work on American Customer Satisfaction Index (ACSI) [15]).

Scale: For the purpose of this metric, we model satisfaction as expectation fulfillment relating to function. We stipulate that users have expectations relating to the following issues: (1) Service availability (scale: %), (2) Support if problems including: (2.1) responsiveness (scale: %) and (2.2) effectiveness (scale: %) and (3) service completeness (scale: %). Thus, the satisfaction metric consists of four 'sub metrics'. The relative significance that a user puts on each of these factors can be established using the MCDA explained below.

Measurement procedure: For each of the above service performance issues, we can establish objective values e.g. service availability by number of interactions with a response time of less than 1 second divided by the total number of interactions (this leaves out infrastructure issues). Support really relates to two separate issues: responsiveness (scale: %), and effectiveness (scale: %) i.e. the 'extent' to which the problem is solved (e.g. completely for now and all future similar problems, or just a partial solution to this particular instance of the problem). Responsiveness is given as

$$1/(1+t) \tag{8.2}$$

where t is the average time in minutes required to 'solve' a problem reported by the user. Service completeness relates to the number of features that the service actually delivers divided by the number of features that the user could reasonably expect (e.g. based on similar services provided by others, or suggested in marketing material from the service provider). Here we make the simplifying assumption that all users perceive each function to be equally important.

For satisfaction, other elements may be relevant such as satisfaction from socialization and usability. However, these are not included because we use a simplified model to explain the method. In addition, other strategies of the data controller such as increasing the 'lock-in-effect' may represent risk.

Profit : The unit for profit is currency units (Euros). The weight for profit will then specify how much utility each currency unit will give.

Reputation : We interpret reputation as the data subject's expectancy relating to future behavior of the data controller. The reader should note that 'expectancy relating to future behavior' refers to the subjective, psychological expectation of the data subject and not 'expected value' in a statistical sense. For the purpose of this case study, we restrict our attention to privacy. Thus, a more appropriate name for this metric would probably be 'Privacy Reputation'.

Scale: When it comes to expectations about reputation it may be strongly influenced by past behavior i.e. our experience [25]. Thus, we model reputation relating to two issues: experience of others (scale: %) and own experience (scale: %).

Measurement Procedure: Both of the above metrics can be established by doing a survey. Clearly, we may also construct more sophisticated models taking into account incident discovery rates, and how (negative) information is spread in a human communication network. For example, we may want to establish the following: "How many privacy incidents affecting your friends (yourself) would you be willing to accept before you would stop using the service?". Then, experience of others (own) can be computed using

$$1/(1+P) \tag{8.3}$$

where P is the maximum number of privacy incidents affecting your friends (yourself) that you are willing to accept. The assumption that users are more willing to accept privacy violations to others than themselves might not always hold. As mentioned earlier, we can easily address this by presenting the risk analyst with several reputations metrics so that he can choose and construct the final model using the most appropriate metric.

In our SNS example, we assume the data subject gives higher concern to his privacy than to the satisfaction he gets from using the service. On the other hand, the data controller values his profit more than reputation. In our example, we assume data subject privacy and satisfaction to be 0.6 and 0.4 respectively. Similarly, we assume the data controller assigns weights of 0.7 and 0.3 for his reputation and profit respectively. For the other elements comprising the utility factors, we make the assumption that the stakeholders perceive each of these to be equally important (see Table 8.1).

2.2. Identify the strategies. Determine the associated strategies of each player (except for risk owner) i.e. 'what can they do or consider doing?'. We assume the risk owner plays only the 'do nothing' option. In general, a strategy may be triggered by a planned behavior or stochastically. Note that in this paper we consider only planned or intentional execution of strategies by the strategy owner. Each strategy may modify the value of the utility factors belonging to own as well as other players' utility factors.

From the scenario, we know that the strategies of the data controller are 'do nothing' (DN'), 'exploit' (E) and 'not exploit' (NE). However, the data subject plays only the 'do nothing' (DN) option.

3. Model the Game.

Model a one shot game (static and simultaneous game) between the risk owner and strategy owners. Assume the system/ environment to be in a fixed initial state and all the players are utility optimizing. By utility optimizing, we mean that they are optimizing their behavior relative to the weighted sum of the elements in their utility factor vector.

Now, we need to first determine the final value of the utility factors, then estimate the utility for each of the strategies for each of the players.

3.1. Determine the Final Value of the Utility Factors. For each of the identified utility factors, determine the final value after the strategies of the players are executed (for the utility factors' valuation, we utilize the metrics explained above). We use the additive utility function of MAUT to estimate the utility. The additive utility function for a given player is defined to be the weighted average of its individual utility functions [11] given as:

$$U = \sum_{k=1}^{m} w_k \cdot u(a_k) \ . \tag{8.4}$$

where
m- number of utility factors of the player,
w_k is the assigned weight of utility factor a_k and $\sum_{k=1}^{m} w_k = 1$, and
$u(a_k)$ is the utility function for the utility factor 'a_k'.

For our SNS example, Table 8.1 depicts the initial value (IV) of the utility factors and also its final value, if the strategies of the data controller were to be executed. For the purpose of this example, we assume the values are obtained from interviews and surveys. Note the 'do nothing' option of the data controller and data subject does not incur any changes to utility factors.

Table 8.1: Final Values of the Utility Factors after the Strategy of the Data Controller is Executed (an example).

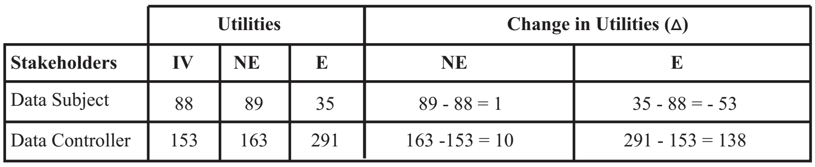
				Final Values	
Stakeholders	**Utility Factors**	**Weights**	**IV**	**NE**	**E**
Data Subject	Privacy (%)	0.60	100	100	9
	Satisfaction (%)	0.40	70	74	74
	Availability (%)	0.33	80	85	85
	Support (%)	0.33	53	56	56
	Responsiveness (%)	0.50	17	20	20
	Effectiveness (%)	0.50	90	92	92
	Service Completeness(%)	0.33	80	82	82
Data Controller	Profit (Euros)	0.70	200	200	400
	Privacy Reputation (%)	0.30	42	75	38
	Experience of others(%)	0.50	33	50	25
	Own experience(%)	0.50	50	100	50

For obtaining the values of *privacy* of the data subject, we instantiate (8.1) with the assumption N=0 per month at the initial state and when the data controller uses option NE. However, we assume N=10 per month when the data controller uses option E. This results in the values of privacy as 100% and 9% respectively.

Note that the values for satisfaction and reputation are obtained using the techniques borrowed from MCDA and MAUT. For *support* (an element of the satisfaction utility factor), the values for the *responsiveness* are obtained by instantiating (8.2) with $t = 5$ at the initial state and $t = 4$ when both the strategies of the data controller are executed. Thus, responsiveness increases from the IV of 17% to 20% for both the strategies. Besides, we assume the *effectiveness* increases from 90% to 92% when both the strategies of the data controller are executed. Now, we evaluate the values for support instantiating (8.4) with the obtained values of responsiveness and effectiveness: for the IV as $0.50 \cdot 17 + 0.50 \cdot 90 = 53\%$. Similarly, the final values for both the NE and E are evaluated as 56%. We make the following assumptions for the other elements of satisfaction: *availability* increases from 80% to 85% and *service completeness* increases from 80% to 82% after both the strategies of the data controller are executed. Thus, using (8.4) and the values determined for the other elements comprising our satisfaction utility factor, the obtained IV is 70% and the final values for both the strategies are evaluated as 74%.

We assume the data controller makes an additional profit of 200 when he uses the E option rather than the NE option. For *reputation*, for the *experience of others*, the values are obtained after we instantiate the number of privacy incidents in (8.3) with $P = 2$, $P = 1$ and $P = 3$ in the initial state, when the data controller uses NE and E as 33%, 50% and 25% respectively. On the other hand, for *own experience* i.e. personal experience of the data subject, the values are obtained as follows after we instantiate in (8.3), $P = 1$ in the initial state and when the data controller uses E and $P = 0$ when the data controller uses NE. Thus, the values obtained in the initial state, when the data controller uses NE and E are 50%, 100% and 50% respectively. Thus, using (8.4) from the values determined for the elements comprising reputation, reputation of the data controller increases from an IV of 42% to 75% when the data controller chooses the NE option. However, it decreases to 38% when the data controller selects the E option.

Usually, the individual utility functions (i.e. utility factors in our case) are assigned values in the interval of 0 (worst) to 1 (best) when using MAUT. For instance, in our case, we can easily compress the profit/ wealth to the interval 0 to 1. However, this would not be particularly helpful as most of the values will be clustered right at the end. Thus, it is more intuitive to utilize the given scales for the utility factors' valuation.



	Utilities			Change in Utilities (Δ)	
Stakeholders	**IV**	**NE**	**E**	**NE**	**E**
Data Subject	88	89	35	89 - 88 = 1	35 - 88 = - 53
Data Controller	153	163	291	163 -153 = 10	291 - 153 = 138

Figure 8.4: Matrix of Utilities and Change in Utilities w.r.t. Strategy of the Data Controller.

3.2. Estimate the Utility. We again use the techniques from MAUT to estimate the utility for each of the strategies for each of the players using (8.4). We make the simplifying assumption that utility is linear.

For our SNS example, we use (8.4) to compute the utilities for the data subject and the data controller with the values given in Table 8.1. In the initial state, the utilities are given as follows:
For the data subject: $0.60 \cdot 100 + 0.40 \cdot 70 = 88$
For the data controller: $0.70 \cdot 200 + 0.30 \cdot 42 = 153$
Similarly, when the data controller selects NE and E, the utilities are obtained as given in Figure 8.4.

4. Compute the Incentives.

We need to compute the incentives (i.e. changes in utilities) for each of the strategies for each of the players. The change in utility Δ is the difference between the utility of the player in the state resulting from strategy use and the initial state.

In the SNS example, from Figure 8.4, when the data controller uses the NE option, Δ for the data subject and data controller are 1 and 10 respectively. When the data controller uses the E option, the Δ for the data subject and data controller are -53 and 138 respectively.

5. Determine Risk.

This can be achieved by investigating each of the strategies with respect to sign and magnitude of the changes determined in 4. The idea being that risk is the combination of the strength of the force that motivates the strategy owner to send the risk owner to an undesirable state and the magnitude of this undesirability. Risk magnitude is related to the magnitude of changes to perceived utility caused by potential state changes.

In other words, we look into how strong the players' incentives are to make the first move. Investigating the players' motivation to move first helps to understand the risks faced by the risk owner. In our model, we make the simplifying assumption that all strategy owners will need the same time to act if they have the same magnitude of incentive. Furthermore, players will move ordered by decreasing incentives and all above a certain threshold will move.

In the SNS example, when the data controller uses the NE option, it results in a positive change in utility for both the players (falls in the first quadrant in the incentive graph). Thus, we know there is no risk. However, it is clear that the data controllers' degree of desirability to play the exploit option is high as it leads him to a better position with a gain of 138. In this case, 138 is the strength of the force that motivates the data controller to send the data subject to an undesirable state and -53 is the magnitude of this undesirability and the combination of these is the risk **(-53, 138)**.

8.4 Discussion

In a game theoretic interpretation, if you are in equilibrium, 'forces' will pull you towards your current state. That is, there is no incentive for the players to unilaterally change state, as the state where the risk owner is heading might be worse for both himself and the strategy owner. Thus, if the strength of the 'force' to change state is negative then we are in equilibrium. We may also define game theoretic risk to be a set related to the concepts of likelihood and consequence as:

1. Likelihood(L_i): How strong is strategy owner i's incentive to make the first move or the magnitude of incentive. The theory of planned behavior provides a link [1].
2. Consequence(C): The value of the game for the risk owner.

Furthermore, risk appetite can be determined by asking questions like: 'How strong a temptation is acceptable to give a strategy owner to cause the risk owner a given loss?' In this setting, risk appetite refers to the collection of (C, L_i) pairs (referred to as SCL) for which the risk owner accepts (and prefers) that the strategy owner i makes the first move. In other words, this collection helps to decide if a risk is acceptable or not. That is, for risk acceptance criteria (C, L_i), risk (C', L_i') is acceptable, iff $C' \geq C$ and $L_i' \leq L_i$. Note that a loss (gain) will have a negative (positive) sign and we need to specify many risk pairs in order to completely specify the risk acceptance criteria. We leave the definition of a total ordering of risk for further work. If the risk owner finds himself in a situation where the current (C, L_i) is not in SCL, he will need to consider other strategies than the 'do nothing' strategy. However, this takes us from risk analysis to risk management which is outside the scope of this paper.

Recall, as shown in Figure 8.3, that each dot represents a particular risk event. The quadrant is divided into three areas: the area for acceptable risk events (represented by white dots) and the area for unacceptable risk events (represented by black dots) separated by the shaded area. The shaded area represents the channel in which the analyst does not have enough information to know if a risk event is acceptable or not. The risk owner's risk appetite will be a graph somewhere in this channel which specifies a bound on the risk he currently is willing to take. However, the risk appetite will differ depending on how the event is triggered. For intended action, in many cases, it will be reasonable to make the assumption that the strategy owner will be rational in a behavioral economics sense. However, the risk appetite in the case of accidental execution may be different. Consequently, we may have multiple risk acceptance graphs.

From the strategy owner's perspective there is always the possibility that the risk owner will implement a control that reduces the incentive that the strategy owner may obtain. This concern is captured through the application of the theory of Discounted Utility (DU) ([17]). The first mover advantage is then equal to the saved opportunity cost that can be attributed to the discount factor. We make the simplifying assumption that all risk events have associated the same discount factor. However, if we have evidence suggesting that this is not the case, we can easily modify our calculations, e.g. by using interval arithmetic to capture the incertitude or by using specific discounting values if these are known. DU has received some criticism relating to the assumption that the discounting factor is assumed constant for all time intervals. However, we do not rely on this assumption as it can easily be seen that the greater the incentive, the greater the potential loss attributable to the discounting. Thus, from the assumption above, a rational strategy owner that is forced to make a choice between risk events will select the risk event with the highest value. This will (for a given strategy owner) result in risk events being sorted according to their relative urgency. However, in general we may not be able to predict the absolute time for when a risk event may occur. Thus, the relative timing of risk events triggered by different strategy owners may not necessarily be known. One way of dealing with this is to establish separate risk appetites for each of the strategy owners. This would then remove the need for assuming that all stakeholders have the same response time for a given incentive.

In addition, thinking of player strategies reveal risks in a way that typical occurrence of events in classical risk analysis does not. As we look into the strategies from the perspective of each stakeholder, we claim it provides a clearer view of the situation in terms of not only what the stakeholder can do but also what he might consider doing. Furthermore, it should be noted that quantification as such does not give objectivity. This is because there will be an element of subjectivity going into the choice and definition of the metrics capturing the utility factors. In most of the classical risk analysis approaches, the probability is obtained by asking questions such as "How likely it is that something will happen?". When there is insufficient historical data or when we have small probabilities, it is difficult to check if the probability is in fact correct. In CIRA, we ask questions such as "How much does a person perceive to benefit from a certain incident?". As objectivity is closely related to 'universal agreement', our transparent approach goes some way towards this goal by means of identifying the key issues where disagreement may exist. Once identified, one can work towards agreement and a common understanding.

8.5 Future Work

For future work, we will investigate to what extent CIRA scales to real world scenarios. We will consider risk from a single (i.e. 'first') adversarial move vs. risks from some number of adversarial moves vs. other scenarios. The idea being that we need to consider the risks posed by all the events up to the point where we have changed our exposure. Besides, we will compare CIRA with one or more of the classical risk analysis methods such as ISO 27005 which will contribute towards understanding its effectiveness.

The following issues require further work: (1) Expressing risk as a single value. (2) We will collect definitions of utility factors from the literature and put into the library. (3) Different people may value the same utility factor differently. Hence, we need to consider weights being captured as distributions. (4) In our SNS example we made the simplifying assumption that utility is linear. We need to investigate the issues relating to the nonlinearity of utility factor valuation. (5) The lack of precision or incertitude can be captured using uncertainty propagation techniques such as P-box, interval arithmetic or similar. This recognition is crucial in order not to enter into a game of self deception. (6) Finally, we will investigate how risk neutral, risk seeking and risk averse behavior can be framed in CIRA.

8.6 Conclusion

We have presented a method for risk analysis combining ideas from game theory, economics, psychology and decision theory where the input parameters can be audited and where risk is modeled in terms of conflicting incentives. Furthermore, investigating the players' incentive to move first helps to understand the risks faced by the risk owner. Our method trades subjective probabilities for stakeholder perceived incentives. By trading subjective probabilities for stakeholder incentives, the risk analyst can focus on utility factors incorporating idea from behavioral economics during the data collection phase of a risk analysis process. Thus, CIRA is applicable in situations where the utilities for the stakeholders can be estimated reasonably well.

Acknowledgement.

The work reported in this paper is part of the PETweb II project sponsored by The Research Council of Norway under grant 193030/S10. We would like to thank Dieter Gollmann and the anonymous reviewers for their valuable comments and suggestions.

8.7 Bibliography

[1] Ajzen, I. The Theory of planned behaviour. *Organizational Behaviour and Human Decision Processes 50* (1991), 179–211.

[2] ANDERSON, R., AND MOORE, T. Information Security Economics - and Beyond. In *In Proceedings of the 27th annual International Crytology Conference on Advances in Cryptology CRYPTO'07* (2007), Springer- Verlag, pp. 68–91. doi:10.1007/978-3-540-74143-5_5.

[3] ASME INNOVATIVE TECHNOLOGIES INSTITUTE, LLC. *Risk Analysis and Management for Critical Asset Protection (RAMCAP): The Framework*, May 2006. Version 2.0.

[4] BANKS, D. L., AND ANDERSON, S. Combining Game Theory and Risk Analysis in Counterterrorism: A Smallpox Example. In *Statistical Methods in Counterterrorism*, A. Wilson, G. Wilson, and D. Olwell, Eds. Springer New York, 2006, pp. 9–22. doi:10.1007/0-387-35209-0_2.

[5] BIER, V. Challenges to the Acceptance of Probabilistic Risk Analysis. *Risk Analysis 19* (1999), 703–710.

[6] BIER, V. Game-Theoretic and Reliability Methods in Counterterrorism and Security. In *Statistical Methods in Counterterrorism*, A. Wilson, G. Wilson, and D. Olwell, Eds. Springer New York, 2006.

[7] BIER, V. M., COX, JR., L. A., AND AZAIEZ, M. N. Why Both Game Theory and Reliability Theory are Important in Defending Infrastructure against Intelligent Attacks. In *Game Theoretic Risk Analysis of Security Threats*, vol. 128 of *International Series in Operations Research & Management Science*. Springer US, 2009, ch. 1, pp. 1–11. doi:10.1007/978-0-387-87767-9_1.

[8] BRABER, F., HOGGANVIK, I., LUND, M. S., STØLEN, K., AND VRAALSEN, F. Model-based security analysis in seven steps — a guided tour to the CORAS method. *BT Technology Journal 25*, 1 (2007), 101–117. doi:10.1007/s10550-007-0013-9.

[9] CAMERER, C. F., AND LOWENSTEIN, G. Behavioral Economics: Past, Present, Future. In *Advances in Behavioral Economics*, C. F. Camerer, G. Loewenstein, and M. Rabin, Eds. Princeton University Press, 2004, ch. 1, pp. 3–51.

[10] CARIN, L., CYBENKO, G., AND HUGHES, J. Cybersecurity Strategies: The QuERIES Methodology. *Computer 41* (2008), 20–26.

[11] CLEMEN, R. T. *Making Hard Decision: An Introduction to Decision Analysis*, 2nd ed. Duxbury, 1996.

[12] COX, JR., L. A. Some limitations of "Risk = Threat x Vulnerability x Consequence" for risk analysis of terrorist attacks. *Risk Analysis 28*, 6 (2008), 1749–61.

[13] COX, JR., L. A. Game Theory and Risk Analysis. *Risk Analysis 29* (2009), 1062–1068. doi:10.1111/j.1539-6924.2009.01247.x.

[14] DEPARTMENT FOR COMMUNITIES AND LOCAL GOVERNMENT. *Multi-criteria analysis: a manual*, 2009. Crown.

[15] FORNELL, C., JOHNSON, M. D., ANDERSON, E. W., CHA, J., AND BRYANT, B. E. The American Customer Satisfaction Index: Nature, Purpose, and Findings. *Journal of Marketing 60*, 4 (Oct 1996), 7–18.

[16] FRICKER, JR, R. D. Game theory in an age of terrorism: How can statisticians contribute? In *Statistical Methods in Counterterrorism*, A. Wilson, G. Wilson, and D. Olwell, Eds. Springer New York, 2006, pp. 3–7.

[17] GOLDIN, J. Making decisions about the future: the Discounted-utility Mode. *Mind Matters: The Wesleyan Journal of Psychology 2* (2007), 49–56.

[18] Hausken, K. Probabilistic Risk Analysis and Game Theory. *Risk Analysis 22*, 1 (2002), 17–27. doi:10.1111/0272-4332.t01-1-00002.

[19] Hayes, B. Computing Science: A Lucid Interval. *American Scientist 91*, 6 (2003), 484–488.

[20] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*, 1st ed. ISO/IEC, 2008.

[21] IT Governance Institute. *COBIT 4.1*, 2007.

[22] Kristandl, G., and Bontis, N. Constructing a definition for intangibles using the resource based view of the firm. *Management Decision 45*, 9 (2007), 1510–1524.

[23] Liu, P., and Zang, W. Incentive-based modeling and inference of attacker intent, objectives, and strategies. In *Proceedings of the 10th ACM conference on Computer and communications security* (New York, NY, USA, 2003), CCS '03, ACM, pp. 179–189. doi:http://doi.acm.org/10.1145/948109.948135.

[24] Loewenstein, G. F., Weber, E. U., Hsee, C. K., and Welch, N. Risk as feelings. *Psychological Bulletin 127*, 2 (2001), 267–286. doi:10.1037/0033-2909.127.2.267.

[25] Money, K., and Hillenbrand, C. Using Reputation measurement to create value: An analysis and integration of existing measures. *Journal of General Management 32*, 1 (2006).

[26] Mullainathan, S., and Thaler, R. H. Behavioral Economics. *NBER Working Paper 7948* (2000).

[27] Rajbhandari, L., and Snekkenes, E. Mapping between Classical Risk Management and Game Theoretical Approaches. In *Communications and Multimedia Security*, vol. 7025 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2011, pp. 147–154. doi:10.1007/978-3-642-24712-5_12.

[28] Sent, E. M. Behavioral Economics: How Psychology Made Its (Limited) Way Back Into Economics. *History of Political Economy 36*, 4 (2004), 735–760. doi:10.1215/00182702-36-4-735.

[29] Shanteau, J., and Stewart, T. R. Why study expert decision making? Some historical perspectives and comments. *Organizational Behavior and Human Decision Processes 53*, 2 (1992), 95–106. doi:10.1016/0749-5978(92)90057-E.

[30] Slovic, P., Finucane, M., Peters, E., and MacGregor, D. G. Risk As Analysis and Risk As Feelings: Some Thoughts About Affect, Reason, Risk, and Rationality. *Risk Analysis 24*, 2 (2004), 311–322.

[31] Stoneburner, G., Goguen, A., and Feringa, A. *NIST SP 800-30, Risk Management Guide for Information Technology*. NIST, 2002.

[32] Taleb, N. N. *The Black Swan: The Impact of the Highly Improbable*, 2nd ed. Random House Trade Paperbacks, May 2010.

[33] Tversky, A., and Kahneman, D. Judgment under Uncertainty: Heuristics and Biases. *Science 185*, 4157 (1974), 1124–1131. doi:10.1126/science.185.4157.1124.

[34] Wallenius, J., Dyer, J. S., Fishburn, P. C., Steuer, R. E., Zionts, S., and Deb, K. Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead. *Management Science 54*, 7 (2008), 1336–1349. INFORMS.

Chapter 9

Using the Conflicting Incentives Risk Analysis Method[1]

Abstract

Risk is usually expressed as a combination of likelihood and consequence but obtaining credible likelihood estimates is difficult. The Conflicting Incentives Risk Analysis (CIRA) method uses an alternative notion of risk. In CIRA, risk is modeled in terms of conflicting incentives between the risk owner and other stakeholders in regards to the execution of actions. However, very little has been published regarding how CIRA performs in non-trivial settings. This paper addresses this issue by applying CIRA to an Identity Management System (IdMS) similar to the eGovernment IdMS of Norway. To reduce sensitivity and confidentiality issues the study uses the Case Study Role Play (CSRP) method. In CSRP, data is collected from the individuals playing the role of fictitious characters rather than from an operational setting. The study highlights several risk issues and has helped in identifying areas where CIRA can be improved.

Keywords: Risk analysis, risk, privacy, conflicting incentives

9.1 Introduction

Risk is usually expressed as a combination of likelihood and consequence but obtaining credible likelihood estimates is difficult. Thus, there is a need to improve the predictability and the coverage of the risk identification process. This challenge is a consequence of limited availability of representative historic data relevant for new and emerging systems. Besides, people are not well calibrated at estimating probabilities [20]. Furthermore, to improve the efficiency of the identification process, there is a need to identify issues that are key to risk discovery, and avoid activities that shed little or no light on potential problem areas. The Conflicting Incentives Risk Analysis (CIRA) [19] method addresses these issues by using an alternative notion of risk. In CIRA, risk is modeled in terms of conflicting incentives between the risk owner and other stakeholders in regards to the execution of actions. However, little evidence exists to suggest that CIRA is feasible to analyze risk in non-trivial settings.

In this paper, we explore to what extent CIRA is feasible for analyzing risk in non-trivial settings. We look into the feasibility of CIRA for analyzing privacy risks in a case study of an identity management system. Privacy is "too complicated a concept to be boiled down to a single essence" [21]. We agree with the view of Solove [21] that it is important to understand the socially recognized activities that cause privacy problems to an individual in order to protect it. As the data collected using CIRA will be sensitive and confidential, data is collected through Case Study Role Play (CSRP). CSRP is developed from the integration of case study [26], persona [7] and role play [25]. Personas are "hypothetical archetypes of actual users" and embody their goals [7]. Each role as described in the persona is played

[1] RAJBHANDARI, L., AND SNEKKENES, E. Using the conflicting incentives risk analysis method. In Security and Privacy Protection in Information Processing Systems, vol. 405 of IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg, 2013, pp. 315-329.

by a real person. Using CSRP, data is collected from the individuals playing the role of fictitious characters rather than from an operational setting. In this paper, we have extended the previous work on CIRA by (1) improving the data collection and analysis phase, and (2) showing that it is feasible to use CIRA in non-trivial settings. Our work has contributed to the development of CIRA and helped to identify practical problems that can be addressed in future research.

The rest of the paper is organized as follows. Related work is given in Section 9.2 followed by a description of the case in Section 9.3. In Section 9.4, we present the analysis of the case. We further present and discuss the result of our analysis in Section 9.5. Section 9.6 concludes the paper.

9.2 Related Work

There are many classical risk management approaches and guidelines. Usually, in these approaches, risk is specified as a combination of likelihood and consequence. The ISO/IEC 27005 [14] standard (its new version ISO/IEC 27005:2011), the ISO 31000 [13] standard (that supersedes AS/NZS 4360:2004 [3]) and NIST 800-39 [17] provide the guidance on the entire risk management process. NIST 800-39 [17] supersedes NIST SP 800-30 [22]; its revised version NIST 800-30 Rev. 1 [18] is a supporting document to NIST 800-39. CORAS [16] is a model based method that uses Unified Modeling Language (UML) for security risk analysis. ISRAM [15] is a survey based model to analyze risk in information security; surveys are conducted for gathering probability and consequence. In Risk IT [11] framework (which is integrated into COBIT 5 [12]), risk is estimated as the combination of frequency (rate by which an event occurs over a given period of time) and magnitude of IT risk scenarios. In RAMCAP [2] (its updated version RAMCAP Plus), risk is estimated as the combination of threat, vulnerability and consequence. Cox has shown the limitations of estimating risk as the combination of threat, vulnerability and consequence [8].

There are several methods that specifically look into privacy risks, and are usually called Privacy Impact Assessment (PIA). For instance, there are Privacy Impact Guidelines of the Treasury Board of Canada Secretariat [23] and PIA of the Information Commissioner's Office, United Kingdom [10]. PIA is a "systematic process for evaluating the potential effects on privacy of a project, initiative, or proposed system or scheme" [24]. It helps to identify and manage privacy risks for an organization that deals with personal data of its stakeholders. However, these methods usually do not attribute the events to people. Wright [24] states that PIA should be integrated into risk management along with other strategic planning tools.

The CIRA Method [19] identifies stakeholders, their actions and consequences of actions in terms of perceived value changes to the utility factors that characterize the risk situation. The idea being that risk is the combination of the strength of the force that motivates the stakeholder that is in the position to trigger the action to send the risk owner to an undesirable state and the magnitude of this undesirability. Risk magnitude is related to the degree of change to perceived utility caused by potential state changes.

9.3 Case Description: NorgID Identity Management System

The case description is fictitious but the design of the system is inspired by MinID [9]. The Identity Management System (IdMS) helps to manage the partial identities of end-users. IdMS usually consists of three class of stakeholders: End-user, Identity Provider (IdP) and Service Provider (SP). IdP is the organization that issues the credentials/ electronic identity to the end-user. SP is the organization that provides services to end-user after verifying their identities.

A-SOLUTIONS is an organization with 20 employees that manages a federated IdMS. It developed an authentication system called NorgID and a portal (ID-Portal). Their goal is to

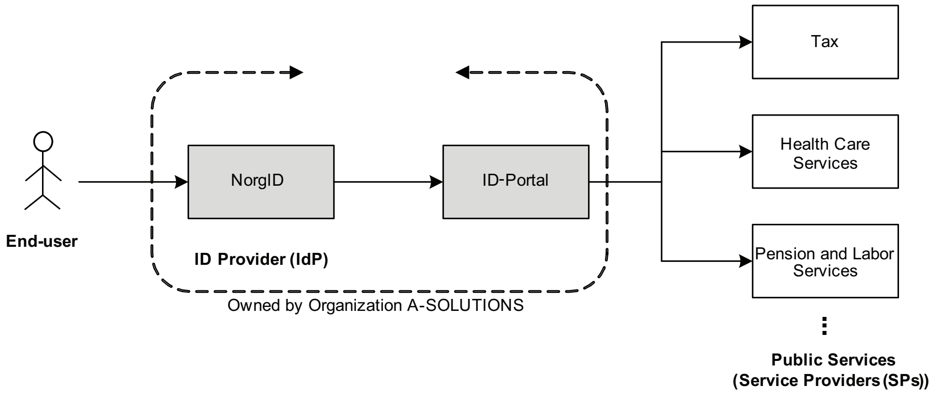

End-user
NorgID
ID-Portal
ID Provider (IdP)
Owned by Organization A-SOLUTIONS
Tax
Health Care Services
Pension and Labor Services
Public Services
(Service Providers (SPs))

Figure 9.1: NorgID Identity Management System.

provide secure access to digital public services. NorgID is one of the IdPs which provides authentication for logging on to a federation called 'ID-portal' as shown in Figure 9.1. It provides the end-user cross-domain Single Sign-On (SSO), i.e. the end-user needs to authenticate only once and can gain access to many services by using the portal such as tax, health care, pension, labor and other eGovernment services. The end-user can log on to the ID-portal using NorgID, by providing his personal ID, a password and a one-time PIN code. NorgID uses two databases: (a) for storing personal data about the users and (b) for storing logs containing usage of IdMS for each user (the details regarding the collected information are not mentioned in the privacy policy). The personal information collected includes his social security number, PIN-codes, password, email address, telephone number and address. NorgID has been quickly and widely adopted because of its easy access and features that have convinced enough people to use the application.

9.4 Analyzing Privacy Risks Using CIRA

In this section, we first provide the assumptions and considerations, along with the scoping for the risk analysis activity. We provide a brief summary of the method along with the steps for data collection (1-9) and analysis (10-13). We then implement the procedure on the given case of an IdMS. The analysis focuses on the risks faced by an end-user.

9.4.1 Assumptions and Considerations

For investigating the case, we used the Case Study Role Play (CSRP) method. We developed personas of the stakeholders based on empirical data collected for the representative stakeholders. However, for instance, in the case of a hacker, as the empirical data might not be easily elicitable, we used assumption persona [4]. According to Atzeni et al. [4], the assumptions may be derived from different sources of data for the type of individuals that are known to attack the systems. The scenarios were written to provide background information of the role to the participants. We assumed that the participants are honest when interacting with the risk analyst. During the data collection phase, the participants were presented with a set consisting of 3 relevant utility factors. We also asked the participants to provide other factors that they valued or gave them perceived benefit. However, for the simplification of the case we have not considered those factors.

9.4.2 Scoping

Scoping consists of the activities used to determine the boundary for the risk analysis activity. We (as the risk analyst) assumed that the system is in a certain initial state. Moreover,

Table 9.1: Procedure in CIRA with approximate time required for each step when implementing NorgID IdMS.

Steps	**Time (mins)**
1. Identify the risk owner (includes development of persona)	30
2. Identify the risk owners' key utility factors	30
3. Given an intuition of the scope/ system- identify the kind of strategies/ operations which can potentially influence the above utility factors	30
4. Identify roles/ functions that may have the opportunities and capabilities to perform these operations	60
5. Identify the named strategy owner(s) that can take on this role (includes development of persona)	90
6. Identify the utility factors of interest to this strategy owner(s)	90
7. Determine how the utility factors can be operationalized	240
8. Determine how the utility factors are weighted by each of the stakeholders	120
9. Determine how the various operations result in changes to the utility factors for each of the stakeholders	280
10. Estimate the utility for each stakeholder	20
11. Compute the incentives	15
12. Determine risk	15
13. Evaluate risk	210

we focused on privacy risk events that are caused by the intentional behavior of a stakeholder.

9.4.3 Summary of CIRA

CIRA identifies stakeholders, actions and perceived expected consequences that characterize the risk situation. In CIRA, a stakeholder is an individual (i.e. physical person) that has some interest in the outcome of actions that are taking place within the scope of significance. There are two classes of stakeholders: the strategy owner and the risk owner. Strategy owner is the stakeholder who is capable of triggering an action to increase his perceived benefit. Typically, each stakeholder has associated a collection of actions that he owns. The risk owner is the stakeholder whose perspective we consider when performing the risk analysis, i.e., he is the stakeholder at risk. By utility, we mean the benefit as perceived by the corresponding stakeholder. Utility comprises of utility factors. Chule et. al. [5] identify the utility factors relevant for our work. Each factor captures a specific aspect of utility e.g. prospect of wealth, reputation, social relationship. The procedure is as given in Table 9.1 along with the approximate time required for each of the steps when implementing the NorgID case study (the required time will be further explained in Section 9.5).

9.4.4 Implementing the CIRA Procedure

The application of CIRA to the NorgID IdMS is presented below.

1. Identify the risk owner.

At first we need to determine the risk owner. The user (Bob) is the risk owner. We assume he represents the general users of NorgID. The persona of Bob is given in Table 9.2.

2. Identify the risk owners' key utility factors.

This step consists of determining the key utility factors for the risk owner.

We presented Bob with three utility factors: privacy, satisfaction from the service and usability along with the explanation for each. We collected his opinion on whether he thought (as a user of NorgID), these factors are important and would give him perceived benefit.

3. Given an intuition of the scope/ system- identify the kind/ classes of operations/ strategies which can potentially influence the above utility factors.

For determining the strategies, we look into the taxonomy of activities that cause privacy problems as provided by Solove [21]. The strategies that we considered are:

- Secondary use of Bob's information (**SecUse**): It is related with using Bob's information for another purpose than that is mentioned in the policy without getting his consent.
- Breach of confidentiality of Bob's information: It is "breaking a promise to keep a persons information confidential" [21]. We consider two strategies that can lead to breach of confidentiality: Sharing credentials (**ShareCred**) and Stealing Information (**StealInfo**).

4. Identify the roles/ functions that may have the opportunities and capabilities to perform these operations.

There can be many strategy owners capable of executing these strategies. However, for this paper we consider only three stakeholders as the objective is to show the feasibility of the CIRA method. The stakeholders are CEO and System Administrator of A-SOLUTIONS, and a hacker capable of executing SecUse, ShareCred and StealInfo operations respectively.

5. Identify the named strategy owner(s) that can take on this role.

In this step, we pin point the strategy owner(s) that are in the position of executing the above strategies. We consider the stakeholders: John (CEO), Nora (System Admin) and X (Hacker). Their personas are provided in Table 9.2.

6. Identify the utility factors of interest to this strategy owner(s).

In CIRA, as we consider the perception of an individual, each relevant stakeholder is an expert. Like before, we provided a list of utility factors for John, Nora and Hacker X to choose from. For the hacker, we identified his utility factors from the existing literature [1]. The identified utility factors for John (CEO): privacy reputation, wealth for business continuity, compliance; for Nora (System Admin): availability, trust, free time and for X (Hacker): wealth, status, ego.

7. Determine how the utility factors can be operationalized.

For each identified utility factor, we determine the scale, measurement procedure, semantics of values and explain the underlying assumptions, if any. The brief explantion of the metrics presented in Table 9.3 and Table 9.4 are a flavor of the metric we used in the analysis for the stakeholders Bob (User) and John (CEO). It is to be noted that different flavors of the metric exist and can be used according to the context. Due to space constraint, we leave out the details of the metrics for the utility factors of Nora (System Admin) and X (Hacker).

Table 9.2: Personas of risk owner and strategy owners.

Role	Name	Description
End-user	Bob	30 years old, local school teacher, regular user of NorgID with general IT knowledge; aware of some privacy issues mainly due to the media coverage of data breaches (associated with services such as social networking and health care).
CEO	John	50 years old, ensures the overall development and relationship with its stakeholders; has motivation to increase the company's service delivery capacity.
System Admin	Nora	29 years old, known for her friendly behavior and highly trusts her co-workers; ensures both the NorgID and ID-Portal are functioning properly and secure; manages the access permission for internal staff to the server; in her absence, to assure that co-workers get proper system function, she usually lets them access servers and even shares important credentials to the server.
Hacker	X	28 years old, skilled in computing and interested in new challenges; to pursue his interest he left his job a year ago and now completely spends his time by gathering knowledge through first-hand experience; wants to earn money and also build status for himself in the so-called hackers' community.

Table 9.3: Metrics for the utility factors of the risk owner Bob (User).

Utility factor	Definition	Measurement Procedure
Privacy(%)	It refers to the extent to which you have control over your personal information. Defined by $1/(1+N)$ (9.1) where N- expected/ projected number of incidents per month.	N is obtained from the analysis of the scenario directly or indirectly caused by the events triggered by various stakeholders [19]. If $N = 0$, the value of privacy is 100%; if $N = 1$, the value of privacy decreases to 50% and so on. That is with increasing number of incidents, the value of privacy decreases.
Satisfaction(%)	It refers to the extent to which you perceive the continuance usage of the portal to access services based on your experience. Model as expectation fulfillment relating to function: service availability, support(reponsiveness (scale: %), effectiveness (scale: %)) and service completeness.	Service availability is the number of interactions with a response time of less than 1 second divided by the total number of interactions. Responsiveness is given as $1/(1+t)$ (9.2) where t is the average time in mins required to 'solve' a problem reported by the user. Effectiveness is the 'extent' to which the problem is solved. Service completeness relates to the number of features that the service actually delivers divided by the number of features that the user could reasonably expect (see [19]).
Usability(%)	It refers to the extent to which a user perceives the ease of interaction with the portal. Model as user's past experience with using the service.	The value can be obtained by doing the survey. A scale of 0 to 100% is used, a value of 0 denotes it takes more than 30 mins to get acquainted with the service; 25% denotes it can be done within 20-30 mins; 50% denotes it takes 10-20 mins; 75% it takes less than 10 mins; 100% denotes it takes less than 5 mins.

Table 9.4: Metrics for the utility factors of the strategy owner John (CEO).

Utility factor	Definition	Measurement Procedure
Privacy Reputation(%)	It refers to the reputation of the company with respect to privacy incidents (e.g. loss, misuse or breach of personal information). Model as user's expectation relating to future behavior of the company in terms of: experience of others and own experience; both defined by $1/(1+P)$ (9.3) where P is the number of privacy incidents.	P is obtained from the survey. If $P = 0$, the value of reputation is 100%; if $P = 1$, the value decreases to 50% and so on. That is with increasing number of incidents, the value of reputation decreases (see [19]).
Wealth(Million €)	The unit for wealth is currency units. The weight for wealth will then specify how much utility each currency unit will give.	It is obtained from the investigation of the entity by the risk analyst.
Compliance(%)	It refers to the extent to which you think the company would benefit by following the rules and regulations. This demonstrates the willingness of the company to take necessary steps to protect the personal information of its stakeholders. Model as percent of compliance with legislation (e.g. Data Protection Act, EU directive).	At first the risk analyst needs to gather the rules that needs to be followed by the company. A value of 0 means that no rules are followed; 25% means that 1/4 of those rules are followed; 50% means that half of those rules are followed; 75% means 3/4 of the rules are followed and 100% means all rules are followed.

Table 9.5: Utility factors for Bob (User).

Rank	Utility factors	Weights
1	Privacy	100
2	Satisfaction	80
3	Usability	70

8. Determine how the utility factors are weighted by each of the stakeholders.

We asked Bob to rank the utility factors based on its importance. Then, for collecting the weights for the utility factors the following question was asked- "Given that you have assigned a weight of 100 to utility factor #1, how much would you assign to utility factor #2, #3 and so on (on a scale of 0-99)?". Bob ranked and assigned weights of 100, 80, 70 to the utility factors privacy, satisfaction and usability respectively as given in Table 9.5.

Similarly, the weights of the utility factors according to their ranking for each of the strategy owners were also collected. John (CEO) assigned weights of 100, 80 and 50 to the utility factors compliance, privacy reputation and wealth respectively. Nora (System Admin) assigned weights of 100, 80 and 78 to the utility factors service availability, free time and trust respectively. X (Hacker) assigned weights of 100, 90 and 85 to the utility factors wealth, ego and status respectively.

9. Determine how the various operations result in changes to the utility factors for each of the stakeholders (start with risk owner).

We assume the system/ environment to be in a fixed initial state and all the players are utility optimizing. By utility optimizing, we mean that they are optimizing their behavior relative to the weighted sum of the elements in their utility factor vector. For each of the identified utility factors, we determine the initial and final values after the strategies of the players are executed (for the utility factors' valuation, we utilize the metrics explained above). We use the additive utility function of MAUT to estimate the utility. The additive utility function for a given player is defined to be the weighted average of its individual utility functions [6] given as:

Table 9.6: Final Values of the Utility Factors after the Strategy of the Strategy Owners are Executed.

				Final Values		
Stakeholders				**John**	**Nora**	**X-Hacker**
	Utility Factors	**Wts**	**IV**	**SecUse**	**ShareCred**	**StealInfo**
Bob(User)	Privacy(%)	0.40	100	8	17	5
	Satisfaction(%)	0.32	72	74	74	74
	Availability (%)	0.33	85	87	87	87
	Support (%)	0.33	52	55	55	55
	Responsiveness (%)	0.50	14	17	17	17
	Effectiveness (%)	0.50	90	92	92	92
	Service Completeness(%)	0.33	80	82	82	82
	Usability(%)	0.28	80	80	80	80
John(CEO)	Compliance(%)	0.43	80	60		
	Privacy Reputation(%)	0.35	67	15		
	Experience of others(%)	0.50	33	9		
	Own experience(%)	0.50	100	20		
	Wealth(Million €)	0.22	5	25		
Nora(Sys Adm)	Service Availability(%)	0.39	85		87	
	Free time(%)	0.31	0		30	
	Trust(%)	0.30	50		90	
X(Hacker)	Wealth(Thousand €)	0.36	0			50
	Ego(%)	0.33	40			95
	Status (%)	0.31	50			85

$$U = \sum_{k=1}^{m} w_k \cdot u(a_k) \tag{9.4}$$

where, m is the number of utility factors of the player, w_k is the assigned weight of utility factor a_k and $\sum_{k=1}^{m} w_k = 1$, and $u(a_k)$ is the utility function for the utility factor 'a_k'.

For our case study, Table 9.6 depicts the normalized weights (for the assigned weights in Step 8) for the utility factors, its initial value (IV) and its final values, if the strategies of the stakeholders were to be executed. For the other elements comprising the utility factors, we make the assumption that the stakeholders perceive each of these to be equally important. The values for the metrics are obtained either based on our investigation or by conducting interviews/surveys with the participants. Usually, the individual utility functions (i.e. utility factors in our case) are assigned values in the interval of 0 (worst) to 1 (best) when using MAUT. For instance, in our case, we can easily compress the wealth to the interval 0 to 1. However, this would not be particularly helpful as most of the values will be clustered right at the end. Thus, it is more intuitive to utilize the given scales for the utility factors' valuation. Moreover, the units of the weights are such that the utility is unit less. Next, the values for each of the stakeholders are determined.

For Bob (User). We determine the values of the first two utility factors for Bob from our investigation and the last one (usability) is based on the survey. To determine the value of privacy to the user, we investigated the number of privacy incidents at each state. Our findings are based on several studies on issues such as how secondary usage of data and breach of confidentiality will impact the end-user. Based on our study, $N = 0$ per month at the initial state. $N = 11$, $N = 5$ and $N = 20$ when John, Nora and Hacker X use their respective strategies. By instantiating (9.1) with the value of N, we obtain the IV of privacy as 100% and its final values as 8%, 17% and 5% respectively.

Note that the values for satisfaction are obtained using the techniques borrowed from MAUT and from our investigation. For support (an element of satisfaction), the values for the responsiveness are obtained after instantiating (9.2) with $t = 6$ at the initial state and $t = 5$ when the other strategies of the stakeholders are executed. Thus, responsiveness

Table 9.7: Matrix of Utilities and Change in Utilities w.r.t. Strategy of the Strategy Owners.

	Utilities				Changes in Utilities (Δ)		
Stakeholders	**IV**	**SecUse**	**ShareCred**	**StealInfo**	**SecUse**	**ShareCred**	**StealInfo**
Bob(User)	85	49	53	48	-36	-32	-37
John(CEO)	59	37			-22		
Nora(Sys Admin)	48		70			22	
X(Hacker)	29			76			47

increased from the IV of 14% to 17% for all three strategies. Besides, it was determined that effectiveness also increased from 90% to 92% when the three strategies of the stakeholders are executed. We then evaluate the values for support instantiating (9.4) with the obtained values of responsiveness and effectiveness: for the IV as 0.50*14+0.50*90 = 52%. Similarly, the final values for the three strategies are evaluated as 55%. The following values were determined for the other elements of satisfaction: availability increases from 85% to 87% and service completeness increases from 80% to 82% after the three strategies are executed. Thus, using (9.4) and the values determined for the other elements comprising our satisfaction utility factor, the obtained IV is 72% and the final values for the other strategies are evaluated as 74%. The value of usability as obtained from Bob was 80% for all cases.

Due to lack of space, we leave out the details of the computations of changes to the utility factors belonging to the other stakeholders. The results can be found in Table 9.6.

10. Estimate the utility.

We again use the techniques from MAUT to estimate the utility for each of the strategies for each player using (9.4). We make the simplifying assumption that utility is linear. For our case study, we use (9.4) to compute the utilities for the stakeholders with the values given in Table 9.6. In the initial state, the utilities are given as follows:
For Bob (User): $0.40 \cdot 100 + 0.32 \cdot 72 + 0.28 \cdot 80 = 85$
For John (CEO): $0.43 \cdot 80 + 0.35 \cdot 67 + 0.22 \cdot 5 = 59$
Similarly, for other stakeholders, the utilities are obtained as given in Table 9.7.

11. Compute the incentives.

We need to compute the incentives (i.e. changes in utilities) for each of the strategies for each player. The change in utility Δ is the difference between the utility of the player in the state resulting from strategy use and the initial state. In our case study, from Table 9.7, when John uses the SecUse option, Δ for Bob and himself are -36 and -22 respectively. When Nora uses the ShareCred option, the Δ for Bob and herself are -32 and 22 respectively. In addition, when the hacker uses the StealInfo operation, the Δ for Bob and himself are -37 and 47 respectively.

12. Determine risk.

This can be achieved by investigating each of the strategies with respect to sign and magnitude of the changes determined in the previous step. In our case study, when John uses the SecUse option, it results in a negative change in utility for both the players (falls in the third quadrant in the incentive graph as shown in Figure 9.2). Thus, we know it is an undesirable situation for both the players and they both want to move out of this quadrant. Thus, this might result in co-operation. However, Nora's degree of desirability to play the ShareCred is slightly more as it leads her to a better position with a gain of 22. In this case, 22 is the strength of the force that motivates Nora to send Bob to an undesirable state and -32 is the magnitude of this undesirability and the combination of these is the risk (-32, 22). Similarly, it is clear that the Hacker X's degree of desirability to play the StealInfo is high as it leads

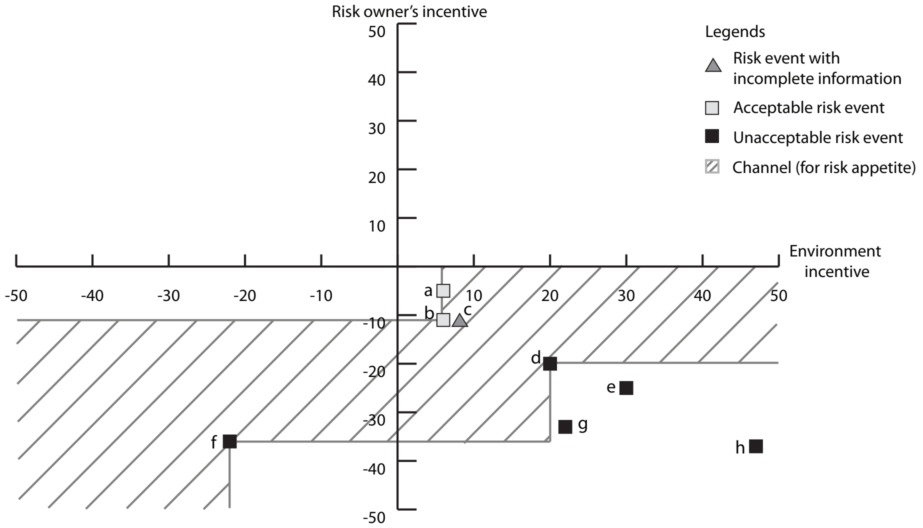


Figure 9.2: The Incentive graph

him to a better position with a gain of 47 and -37 is the magnitude of the undesirability faced by Bob, which results in the risk (-37, 47).

13. Evaluate risk.

We identity the risk acceptance and rejection criteria for the risk owner to determine whether a specified level of risk is acceptable or not. In our model, we make the simplifying assumption that all strategy owners will need the same time to act if they have the same magnitude of incentive. Strategies will be executed in decreasing order of utility as perceived by each of the strategy owners.

We presented Bob with following risk pairs: a. (-5,6), b. (-11, 6), c. (-11, 8), d. (-20,20), e. (-28, 30) along with the ones determined in Step 12, which are f. (-36,-22), g. (-32,22) and h. (-37, 47) obtained when the strategy owners execute the strategies SecUse, ShareCred and StealInfo respectively. The risk pairs are represented by (C, I_i) where C is the consequence for the risk owner and I refers to how strong is strategy owner i's incentive to make the first move or the magnitude of incentive. For instance, for the risk pair 'b', Bob gets the value of C as -11 when the final values for privacy, satisfaction and usability in the execution of any of the strategies would be 95%, 70% and 50% respectively (keeping the weights of the utility factors and their initial values as obtained before). Note that this is one of the several possible combinations that gives Bob the consequence of -11. Nora has an incentive of 6, when the final values for availability, freedom and trust are 90%, 10% and 53% respectively. Similarly for other stakeholders the possible combinations can be determined.

To determine the risk acceptance criteria, we asked Bob (User): 'How strong a temptation is it acceptable to give a strategy owner to execute the strategy, so as to cause him (i.e. Bob) a given loss?'. From the above risk pairs, he accepted the risk pairs a and b (represented by the light gray square) as shown in Figure 9.2. However, for the risk pair, c he was willing to accept the risk only if Nora was in the position of executing the strategy (represented by the triangle) and unsure in case other strategy owners executed their strategy. The remaining risk pairs were not acceptable to him (represented by the black square).

9.5 Results and Discussion

Our findings can be grouped in the following categories: (1) application of CIRA to NorgID IdMS, (2) feasibility of CIRA in terms of its complexity and risk analyst effort required, (3) improvements made and (4) some limitations of CIRA that require further work. Application of CIRA to NorgID IdMS, resulted in the determination of risks faced by the risk owner. We were further able to represent acceptable/ unacceptable risk events by means of an incentive graph which was easy to communicate to the risk owner.

Assuming we have n stakeholders, each stakeholder owns s strategies and has u utility factors that go into the computation of his utility, then the effort of the various tasks can be estimated as follows: The total number of strategies to be considered will be $n * s$. The total number of utility factors to be considered will be $n * u$. However, in practice, it is expected that utility factors will be taken from a limited set. To determine the risk acceptance criteria, it will suffice to ask the risk owner $n * s$ yes/no (i.e. accept/reject) questions. Thus the complexity of CIRA in terms of human effort will be in the order of

$$n * (u + s) \tag{9.5}$$

By instantiating (9.5) with the value of $n = 4$, $s = 1$ and $u = 3$ as in the NorgID case study, we obtain the estimate of complexity as 16. Furthermore, the effort in terms of total amount of time spent in doing the case study was determined to be approximately 27 hrs (which includes the time given in Table 9.1 along with the time for initial preparation (1 hr), scenario construction to provide the background information of the role to the participants (2 hrs), role play selection and guidance (2 hrs) and documentation (1.5 hrs)). The given hours are approximate values; the values were jotted down only after the actual process was completed. It is clear that steps for determining the changes to the utility factors with respect to the operations (Step 9) and the operationalization of utility factors (Step 7) required the highest amount of time i.e. approximately 280 and 240 mins respectively. When the problem space grows, for instance the values of $n = 8$, $s = 10$ and $u = 5$, we would expect that the risk analyst would have to spend in the order of 200 hours to complete the analysis. Note that the elapsed time may be longer. CIRA is still in development phase and the steps will be optimized. For e.g. a comprehensive library of utility factors will be developed. It is expected that this library will speed up the data collection phase. Moreover, tools will be developed to support the risk analyst.

Learning from the case study, we discovered the following issues that resulted in improvements: the procedure was updated to ease the data collection process and the data collection manual was developed for the risk analyst. Interviews/ survey responses indicated that it was essential that the risk analyst and the participants have the same understanding of the concepts (e.g. utility factors) used during the data collection phase. Thus, even though a lot of resources were required for instance, in the operationalization of the metrics for the utility factors and also determining their value, we focused on these key issues in order to improve data quality.

The following limitations of CIRA were identified: (1) We have assumed that all the participants are honest when interacting with the risk analyst. However, the fact that they might be reluctant to provide information or give wrong information during the interview/ survey needs further investigation. (2) As metrics have always been a challenge in information security, for some of the utility factors it was difficult to formulate the metrics. Hence, we need to collect definitions of utility factors and perform their validation. (3) To determine whether an obtained set of utility factors represents the complete set for a particular stakeholder in a given context requires further work. (4) More work is also needed in capturing the uncertainties in relation to estimates using interval arithmetic or bounded probabilities instead of point values. (5) When assigning weights, the same scale is used for all the stakeholders. The mapping of scale of one stakeholder with another also needs further investigation. (6) Finally, CIRA tool support.

9.6 Conclusion

In this paper, we have explored the feasibility of CIRA to analyze risk in a non-trivial setting. The CIRA method is still at an early stage of development. However, the results from our case study suggests that it is possible to use CIRA in such settings, and that the method helps the analyst to get a better understanding of the risks. Our work has contributed to the development of CIRA and helped to identify practical problems that can be addressed in future research.

Acknowledgement.

The work reported in this paper is part of the PETweb II project sponsored by The Research Council of Norway under grant 193030/S10. We would like to thank the anonymous reviewers for their valuable comments.

9.7 Bibliography

[1] *The Honeynet Project. Know Your Enemy*, 2 ed. Addison-Wesley, 2004.

[2] ASME INNOVATIVE TECHNOLOGIES INSTITUTE, LLC. *Risk Analysis and Management for Critical Asset Protection (RAMCAP): The Framework*, May 2006. Version 2.0.

[3] AS/NZS 4360. *Risk management*. AS/NZS, 2004.

[4] ATZENI, A., CAMERONI, C., FAILY, S., LYLE, J., AND FLECHAIS, I. Here's Johnny: a Methodology for Developing Attacker Personas. *ARES* (2011), 722–727.

[5] CHULEF, A., READ, S., AND WALSH, D. A Hierarchical Taxonomy of Human Goals. *Motivation and Emotion 25*, 3 (2001), 191–232(42).

[6] CLEMEN, R. T. *Making Hard Decision: An Introduction to Decision Analysis*, 2nd ed. Duxbury, 1996.

[7] COOPER, A. *The Inmates Are Running the Asylum*. Macmillan Publishing Co., Inc., Indianapolis, IN, USA, 1999.

[8] COX, JR., L. A. Some limitations of "Risk = Threat x Vulnerability x Consequence" for risk analysis of terrorist attacks. *Risk Analysis 28*, 6 (2008), 1749–61.

[9] DIFI (DIREKTORATET FOR FORVALTNING OG IKT). MinID. http://minid.difi.no/minid/minid.php?lang=en. [Online accessed: 06-2012].

[10] INFORMATION COMMISSIONER'S OFFICE (ICO). Privacy Impact Assessment Handbook, 2009. Version 2.0.

[11] ISACA. *The Risk IT Framework*, 2009.

[12] ISACA. *COBIT 5: A Business Framework for the Governance and Management of Enterprise IT*. IT Governance Institute, 2012.

[13] ISO 31000. *Risk Management – Principles and Guidelines*. ISO, 2009.

[14] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*, 1st ed. ISO/IEC, 2008.

[15] KARABACAK, B., AND SOGUKPINAR, I. ISRAM: information security risk analysis method. *Computers & Security 24*, 2 (2005), 147–159. doi:10.1016/j.cose.2004.07.004.

[16] Lund, M. S., Solhaug, B., and Stølen, K. A Guided Tour of the CORAS Method. In *Model-Driven Risk Analysis*. Springer Berlin Heidelberg, 2011, pp. 23–43.

[17] NIST. *NIST SP 800-39, Managing Information Security Risk - Organization, Mission, and Information System View*, 2011.

[18] NIST and U.S. Department of Commerce. *NIST SP 800-30 Revision 1, Guide for Conducting Risk Assessments*, September 2012.

[19] Rajbhandari, L., and Snekkenes, E. Intended Actions: Risk Is Conflicting Incentives. In *Information Security*, vol. 7483 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 370–386.

[20] Shanteau, J., and Stewart, T. R. Why study expert decision making? Some historical perspectives and comments. *Organizational Behavior and Human Decision Processes 53*, 2 (1992), 95–106. doi:10.1016/0749-5978(92)90057-E.

[21] Solove, D. J. A Taxonomy of Privacy. *University of Pennsylvania Law Review 154*, 3 (January 2006), 477. GWU Law School Public Law Research Paper No. 129.

[22] Stoneburner, G., Goguen, A., and Feringa, A. *NIST SP 800-30, Risk Management Guide for Information Technology*. NIST, 2002.

[23] Treasury Board of Canada Secretariat. Privacy Impact Assessment Guidelines: A Framework to Manage Privacy Risks Guidelines. http://www.tbs-sct.gc.ca, April 2012. [Online accessed: 1-2013].

[24] Wright, D. Should privacy impact assessments be mandatory? *Commun. ACM 54*, 8 (2011), 121–131. doi:10.1145/1978542.1978568.

[25] Yardley-Matwiejczuk, K. M. *Role play: theory and practice*. Sage Publications Limited, 1997.

[26] Yin, R. K. *Case Study Research: Design and Methods*, 4th ed., vol. 5 of *Applied Social Research Method Series*. Sage, 2009.

Chapter 10

Case Study Role Play for Research and Training[1]

Abstract

Typically, a risk analysis may identify and document sensitive and confidential information regarding threats, vulnerabilities, assets and their valuation, etc. The intrusive nature of the risk analysis process makes it difficult for researchers (or students) to gain access to scenarios from operational organizations for evaluating (or training on) risk analysis methods. In order to resolve these issues, we propose Case Study Role Play (CSRP). We elaborate the use of CSRP in combination with the Conflicting Incentives Risk Analysis (CIRA) method to analyze privacy risks to an end-user from using the eGovernment service. This paper contributes by demonstrating how CSRP helps to establish a platform for doing risk management related research and training in a 'reasonably' realistic environment, where confidentiality, sensitivity issues, red tape and the need for permissions do not create roadblocks. Furthermore, CSRP ensures that the time and resources needed to set up the required environment is low and predictable.

10.1 Introduction

Risk analysis helps to identify and estimate risks, and to provide insight suitable for deciding if risk exposure needs to be changed. That is, if a treatment action is needed, or a high risk exposure is more cost effective. Here, we focus on risk analysis in the context of information security and privacy management.

Typically, an information security risk analysis may identify and document sensitive and confidential information regarding threats, vulnerabilities, assets and their valuation, etc. Most of the risk analysis method evaluations are usually presented with a toy example. In [13], Kotulic et al. state that due to the sensitivity issues, they faced difficultly in validating their model. Their study was declined by majority of the organizations they had contacted. They write "we learned, the hard way, that developing a research stream in an emerging, organization-sensitive area requires major personal, financial and professional commitments far beyond what most researchers can afford to expend". Information Security Management (ISM) is about people rather than technology. Training students in core ISM skills requires access to a representative teaching environment, for e.g., an operational business setting. However, in the context of information security risk management, researchers/ students (referred to as *practitioners* from here on, in situations where it fits both) cannot usually be cleared for access to sensitive information and will not be permitted to perform representative vulnerability discovery activities. These issues create a blockage for public research or training on risk analysis methods.

Risk analysis case study research may have multiple objectives such as discovering new unknown vulnerabilities, validating a method (checking how well it performs, if it is usable, scalable) and for providing a real life like training platform. In this paper, we focus

[1] RAJBHANDARI, L., AND SNEKKENES, E. Case Study Role Play for Risk Analysis Research and Training. In Proceedings of the 10th International Workshop on Security in Information Systems. SciTePress, 2013, pp. 12-23.

on the latter two stated objectives. We can use case study research in two settings: non-interventive and interventive. In a non-interventive case study, the practitioner is basing his work on information that can be obtained without interacting with the organization in question. In most cases, the bulk of the information will be from written sources. In an interventive case study, individuals from the organization in question will participate in activities initiated by the practitioner. This typically includes answering questions or following procedures prescribed by the practitioner.

These two approaches have very different performance characteristics. The interventive case study may give rise to increased costs for the organization in terms of lost time. Furthermore, sensitive information regarding members of staff, technology, procedures, plans etc. may be disclosed to a third party (the practitioner). For example, the construction of psychologic profiles of a large number of members of staff to be used in a risk analysis will in general be time consuming, intrusive and sensitive. Kotulic et al. [13] state that information security research is highly intrusive in nature, thus it is hard for the organizations and individuals to trust an outsider trying to gain data about their security strategies or practices. Unless the researcher is able to convince the organization that he is providing significant value, there is no reason for the organization to consider offering access - not even conditioned on the signing on an NDA (Non- Disclosure Agreement). When security is at stake, access will be even more restricted. There is always the possibility that the parties involved will not respect the NDA. Clearly, a prudent organization will take this risk into account when deciding if a third party is to be offered access.

The non-interventive case study is only suitable for research where the required information is readily available and when one is not doing research into the interaction process itself. In a non-interventive case study, findings will typically be illustrated through examples rather than through aggregated data. This represents an additional challenge, since in many cases, the researcher would like to publish the results. This challenge manifests itself as a lack of published work reporting on the use of risk analysis for non-trivial scenarios.

A resolution towards the above issues is proposed by specifying and demonstrating the Case Study Role Play (CSRP) method. CSRP is developed from the integration of case study [23], persona [5] and role play [22]. In CSRP, data is collected from the individuals playing the role of the fictitious characters rather than from an operational setting. In our study, role play as a mechanism helps in mimicking behavior (that would yield sensitive output) for producing non-sensitive output. Apart from creating the persona for the users (which is normally done in user centered design), we create personas of a wide range of stakeholders e.g. CEO, IT Manager and System Administrator. Moreover, in our approach, each role as described in the persona and scenario is played by a real person. By doing this, we can extract information from the participant as required by the risk analysis method.

In this paper, we investigate- 'Can CSRP be used as a platform for risk analysis research and training, resolving the inherent problem of risk analysis findings' sensitivity and confidentiality?'. The Conflicting Incentives Risk Analysis (CIRA) method [19] addresses human aspects of information security. Thus, CIRA seems an ideal candidate for demonstrating CSRP. Typically, when applying CIRA in a real life risk analysis activity, it will produce sensitive information as depicted in Figure 10.1(a). Note that $CIRA_1$, $CIRA_2$ and $CIRA_3$ corresponds to the structural data collection phase, numerical data collection phase and analysis phase of the CIRA method (explained in Section 10.4). However, when using CSRP with CIRA, the information extracted will be non-sensitive as illustrated in Figure 10.1(b).

This paper contributes by demonstrating how CSRP helps to establish a platform for doing risk management related research and training in a 'reasonably' realistic environment, where confidentiality, sensitivity issues, red tape and the need for permissions do not create roadblocks. CSRP can be used as a platform for improving/ gaining new knowledge about risk analysis methods. Furthermore, CSRP ensures that the time and resources needed to set up the required environment is low and predictable. However, CSRP may

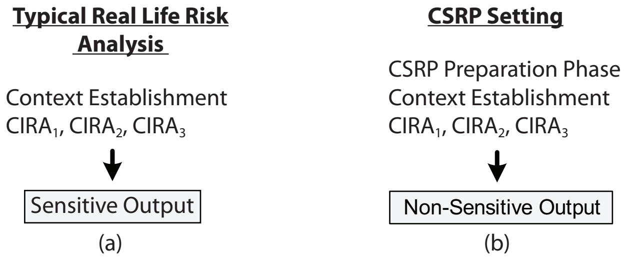


Figure 10.1: Steps and output of CIRA in (a) Typical real life risk analysis and (b) CSRP setting

not be particularly well suited for discovering actual risks in an operational system unless the role play closely matches a real organization.

The remainder of this paper is organized as follows. Related work is presented in Section 10.2 followed by the explanation of CSRP approach in Section 10.3 and summary of CIRA in Section 10.4. In Section 10.5, we provide an overview of how CSRP was utilized in one of our case studies to analyze the privacy risks. In Section 10.6, we discuss the results, limitations and lessons learned from using CSRP, and provide suggestions for future work. Finally, we conclude the paper in Section 10.7.

10.2 Related Work

Our study is inspired by the work on the use of case studies, personas and role play. These, along with risk analysis and management methods, are briefly described and discussed below.

10.2.1 Risk Analysis and Management Methods

There are many classical risk analysis and management approaches and guidelines in the context of information security such as the ISO/IEC 27005:2008 standard [12] (its new version ISO/IEC 27005:2011), the ISO 31000 standard [11] (that supersedes AS/NZS 4360:2004 [2]), NIST 800-39 [16] (that supersedes NIST SP 800-30 [21]; its revised version NIST 800-30 Rev. 1 [17] is a supporting document to NIST 800-39), CORAS [14], TVRA [6], Risk IT [10] and OCTAVE [1]. For conducting an effective risk analysis and management, the inputs from the stakeholders need to be considered and the results need to be communicated.

Depending on the depth of the analysis to be conducted, these methods involve gathering of sensitive and confidential data. For e.g., in the ISO/IEC 27005:2008 standard [12], when defining the scope of risk management, information about the organization is collected so as to determine its operational setting/ environment. The information includes the organization's strategic business objectives, strategies and policies, business processes, the organization's functions and structure, information assets, constraints affecting the organization, etc. Besides these, other relevant information necessary for conducting risk analysis is collected throughout the process. According to the ISO 31000 standard [11] risk management relies on the foundation of the best available information. The information sources as per the ISO 31000 standard include historical data, experience, stakeholder feedback, observation, forecasts and expert judgment. In Risk IT [10], the 'collect data' process under risk evaluation domain is dedicated to gathering data on the organization's operating environment and risk events in order "to enable effective IT related risk identification, analysis and reporting".

10.2.2 Case Study

According to Yin [23], a case study is “an empirical inquiry that investigates a contemporary phenomenon in depth and within its real life context, especially when the boundaries between phenomenon and context are not clearly evident”. Case study is a method that is widely used in research but many view it as being subjective, lacking rigor, requiring less effort than other research methods e.g. experiments. Yin [23] and Flyvbjerg [7] have pointed out and clarified these misunderstandings about case studies by providing relevant examples and explanation. Even though they emphasize these in relation to social science, we think their explanations are relevant for our area of focus. Flyvbjerg further states that the benefit of a case study is that “it can close in on real-life situations and test views directly in relation to phenomena as they unfold in practice” [7].

10.2.3 Persona

Persona [5] is typically employed in user centered design for system development projects. Persona are fictitious characters that are built to represent the users’ needs and goals. It was used by Cooper [5] in order to remove biases of the programmers that resulted in their own assumptions and opinions about the ‘user’ for which the system was being designed. He phrased this act as “designing for the elastic user”. For him, persona should be believable with specific details and each persona should be distinct by their descriptions and scenarios in relation to their respective goals. Personas are designed based on empirical data collected from representative users. Typically, data is collected using methods such as interviews, ethnography, workshops and observations to create good description of the user. However, when the empirical data might not be easily elicitable, e.g. in the case of an attacker, an assumption persona can be used [3]. According to Atzeni et al. [3], the assumptions may be derived from different sources of data for the type of individuals that are known to attack the systems.

Nielsen [15] explains how to write a good description of the user i.e. how to describe the ‘user’ as a character in a way that engages the readers. She states that it is important to consider the users’ environment, character traits, goals and tasks. A persona is brought to life by giving it “a name, a life, a personality as well as a portrait” [9]. Pruitt et al. [18] states that persona is not a science but a powerful tool that helps to engage people in an effective way. After all, one of the incentives behind using persona is to use it as a means for communication or discussion.

10.2.4 Role Play

Role play has been used in entertainment (theater/ movies), research, education, clinical training and therapies. Role play is the way of “deliberately constructing an approximation of aspects of a real life episode or experience” which is controlled (initiated and/ or defined) by the investigator [22]. According to Greenberg et al., role play is used in organizational research to learn about attitudes and behaviors of individuals in organizations and to understand about the basic psychological processes [8]. Even though role play can be conducted for various studies, they point out that in all of the cases, it differs in three dimensions: level of involvement, role being played (self/ other) and degree of response specificity provided.

10.3 Case Study Role Play

CSRP is obtained from the integration of case study [23], persona [5] and role play [22]. In CSRP, data is collected from the individuals playing the role of the fictitious characters rather than from an operational setting. In our study, role play as a mechanism helps in mimicking behavior (that would yield sensitive output) for producing non-sensitive output. The CSRP preparation phase consists of the following steps:

Determine the objective of the activity.

We first decide the objective of the activity e.g. whether it is to mimic an operational setting for the purpose of gaining knowledge about the performance of a risk analysis method or to provide a real life like training platform.

Select the organization.

We then select the organization that would be appropriate for the above identified activity e.g. bank, software company.

Familiarization with the method.

The practitioner needs to get familiar with the risk analysis method to be used for the analysis. The information and procedural requirements of the risk analysis method need to be identified.

Design and build the organization.

We design an abstract form of the selected organization considering all information and procedural requirements required to carry out the risk analysis activities. The requirements may include identifying the objective of the organization, stakeholders, service architecture, process flow, etc. To capture the essential features of that context, experimental setting with equipments may also be set up. After the stakeholders are identified, it is followed by persona and scenario construction, role play selection and guidance as given below.

Persona and Scenario Construction. For each of the identified stakeholders, we design the personas and develop the scenarios. The intention behind constructing persona is to communicate who the stakeholders are, what they are like, what their roles/ tasks are, what their needs are, etc. Attributes such as name, age and gender are assigned to the personas. The possible behavioral variable types proposed by Cooper [5] are used where applicable which includes activities, attitudes, aptitudes, motivation and skills. These are derived from the empirical data or assumptions of stakeholders in a certain situation or from existing data sources.

In our case, scenarios are written to provide the background information of the role to the participants. The background information includes the goals, needs of an individual and how one can accomplish it. Besides that, it also includes the details of the company or the system a person is working on for which the risk analysis is being conducted. While developing the scenario to be provided to the participant, the narrative is written such that it helps the participant to be able to imagine her/ himself in that position as required by the persona. The narrative makes the technological knowledge about the system and process/ flow easier to comprehend for a non-expert.

Role Play Selection and Guidance. The participants are selected to play the role of each persona. At first, the initial approval for participation needs to be gathered. It should be made clear that as the participants are playing a role/ character, the personal data of the participant will not be collected. Moreover, the participation should be voluntary and the results completely anonymous.

Each of the players are provided with a set of instructions explaining how to play the role that he has been assigned. For instance, when doing the risk analysis of an information system, the role of the user can be played by someone who is using a similar system or has some knowledge about it and hence should be representative of general users. However, in cases where a close match is not available, for instance, in the case of the hacker, someone who has knowledge or has done research about the hackers can be selected. Alternatively, the data can simply be collected based on research work about hackers.

Finally, the selected risk analysis method is applied and data as required by the method are gathered from the participants by conducting interviews, through questionnaires, etc.

10.4 Summary of the Conflicting Incentives Risk Analysis Method

CIRA [19] identifies the stakeholders, their actions and perceived expected consequences that characterize the risk situation. There are two classes of stakeholders: the strategy owner and the risk owner. The strategy owner is the stakeholder who is capable of triggering an action that will influence the risk owner. Typically, each stakeholder is associated with a collection of actions that he owns. The risk taker is the stakeholder whose perspective is taken when performing the risk analysis, that is, he is the stakeholder at risk. Utility is the benefit as perceived by the corresponding stakeholder and it comprises of utility factors. Chulef et. al. [4] identify utility factors relevant for our work. Each utility factor captures a specific aspect of utility, for e.g., prospect of wealth, reputation, social relationship. The procedure in CIRA consists of context establishment and three phases as shown in Figure 10.1(b), which are explained below:

$CIRA_1$ (Structural Data Collection).

This phase consists of the identification of stakeholders, their utility factors and actions. Based on the case description, the risk taker and his key utility factors are identified. Note that these utility factors are informally identified by the risk analyst from the case description and later finalized by interviewing the respective stakeholders. Secondly, the actions that can influence the risk taker's utility factors are identified. Then, the roles that may have the opportunities/ capabilities to perform these actions are identified. Finally, the strategy owners that can take on these roles and their utility factors are determined.

$CIRA_2$ (Numerical Data Collection).

This phase consists of determining how the utility factors can be operationalized, how the stakeholders weigh the utility factors and how the various actions result in changes to the utility factors for each of the stakeholders.

$CIRA_3$ (Analysis).

The risks to the risk owner are determined and evaluated.

10.5 Using CSRP for CIRA Research and Training

In this section, we explain how CSRP is utilized for evaluating the performance of CIRA and to provide a real life like CIRA training platform. We used CSRP in one of our studies, to analyze the privacy risks facing the end-user of an eGovernment service. The details on the case and the overall application are provided in [20]. Below, we provide an overview of the procedure as depicted in Figure 10.1(b).

CSRP Preparation Phase.

The objective of using CSRP was to mimic an operational setting for the purpose of gaining knowledge about the performance of the CIRA method. We selected A-SOLUTIONS (that

Table 10.1: Personas of the stakeholders

Role	Name	Description
End-user	Bob	30 years old, local school teacher, regular user of NorgID with general IT knowledge; aware of some privacy issues mainly due to the media coverage of data breaches.
CEO	John	50 years old, ensures the overall development and relationship with its stakeholders; has motivation to increase the company's service delivery capacity.
System Admin	Nora	29 years old, known for her friendly behavior and highly trusts her co-workers; ensures the system is functioning properly and secure; manages the access permission for internal staffs to the server; in her absence to assure that co-workers get proper system function, she usually lets them access servers and even shares important credentials to the server.

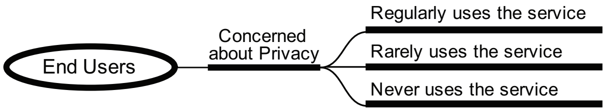


Figure 10.2: Categories of end-users for which the personas can be constructed

manages an Identity Management System) as an organization to carry out the study. The process of eliciting the data from the stakeholders involved in the actual operational setting using the CIRA method gave rise to sensitivity and confidentiality issues. This is because, when analyzing real life scenarios, personal information of the stakeholders such as their preferences, goals, actions and information about their wealth, reputation, status, etc. may need to be collected. In our case, this data is sensitive and data collection is considered intrusive.

We then created an abstraction of the fictitious organization anticipating information required by CIRA. We assumed A-SOLUTIONS manages an authentication system called NorgID and a portal (ID-Portal). Their goal is to provide secure access to digital public services. NorgID is one of the identity providers which provides authentication for logging on to a federation called 'ID-portal'. It provides the end-user cross-domain Single Sign-On (SSO), i.e., the end-user needs to authenticate only once and can gain access to many services by using the portal for eGovernment services such as tax, health care and pension. The stakeholders were identified followed by the persona and scenario construction, participants' selection as explained below.

Persona and Scenario Construction. During the study, we created the persona "Bob" to represent an end-user, based on the assumption that he is concerned about privacy. The synopsis of the persona for Bob is given in Table 10.1.

Instead of having someone that represents the general population of users, Bob portrays a specific end-user. Apart from Bob, we can construct personas, for e.g., for individuals that are concerned about privacy but use the service rarely or never as shown in Figure 10.2. We considered Bob as the primary persona i.e. the individual with the main focus in our analysis.

Similarly, for other stakeholders, the personas were created. The synopsis of the personas are given in Table. 10.1. Due to space constraint, we leave out the description of the hacker that was included in the study [20]. Then, for each of the stakeholders, the scenario was written to provide background information. For example, the scenario description for Bob included the brief description of the eGovernment service, for what purpose he uses it, how it works and what personal information is collected by the service. The scenario description for the other two stakeholders included the description of the company and the functionality of the system.

Role Play Selection and Guidance. The participants were selected for each of the persona. All the participants were IT literate. At first the initial approval for participation was gathered. Then, each of the participants were introduced to the persona and the scenario description was explained. The participants were also given an explanation about the process of data collection for the study.

Context Establishment.

CIRA is implemented starting with context establishment which includes defining the scope and boundaries of risk analysis, objectives of the organization, etc.

$CIRA_1$ **(Structural Data Collection).**

In our scenario, the risk taker is the end-user (Bob). We identified the key utility factors for Bob which were privacy, satisfaction from using the service and usability. Then, the actions that can influence Bob's utility factors were identified. Here, we focused on the actions that cause privacy risks to Bob such as secondary use (*SecUse*) and breach of confidentiality of his information caused by sharing credential (*ShareCred*) by the other stakeholder. We identified the roles that may have the opportunities to perform these actions and the stakeholders (i.e. the strategy owners) in those particular positions. We considered the strategy owners as the CEO (John) and the System Administrator (Nora) of the company operating the eGovernment service in the position to execute the SecUse and ShareCred strategies respectively. Then, we identified the utility factors of interest to these strategy owners.

$CIRA_2$ **(Numerical Data Collection).**

In our case, the necessary data (e.g. weights for the utility factors) as required by CIRA were collected through interviews and surveys from the participants representing Bob, John and Nora. In this way, data can be collected through role play from participants rather than from those in an operational setting.

$CIRA_3$ **(Analysis).**

Finally, the risks to Bob were determined and evaluated. It was determined that Bob faced higher risk in terms of breach of confidentiality of his information than that of his information being used for secondary purpose. This is because when John (CEO) executed the *SecUse* strategy, it resulted in negative utility for both himself and Bob. However, when Nora (System Administrator) executed the *ShareCred* strategy, it resulted in positive gain for Nora and loss for Bob.

10.6 Discussion

In this section, we include (a) results that are beneficial to the practitioners, (b) explain limitations and lessons learned from using CSRP and (c) provide suggestions for further work.

Results.

The application of CSRP proved beneficial as it helped to analyze the risks faced by an end-user when using the eGovernment service without having to worry about the intrusive nature of the study and the hassle of getting access to an operational case study and the stakeholders in an operational setting [20]. It was easy to communicate the details to the participants using CSRP because of the narrative nature of persona and scenario.

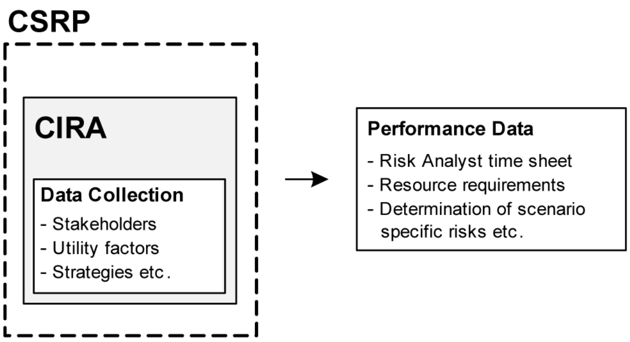


Figure 10.3: Samples of data collected to assess CIRA performance using CSRP

Figure 10.3 depicts data collected using CSRP to assess the performance of CIRA. The data collection in CIRA includes the structural data (e.g. stakeholders, their utility factors) and numerical data (e.g. values of the utility factors) as explained above. By using CSRP, we were able to collect CIRA performance data such as the approximate time required for the analysis, determination of scenario specific risks and resource requirements.

Limitations and Lessons Learned.

CSRP is an approach with a primary focus on method specific research with the objectives of validating a method or for providing a real life like training platform. It can also be used for discovering new unknown vulnerabilities of a system. However, it may not be particularly well suited for discovering actual risks in an operational system unless the role play closely matches a real organization.

One of the issue identified was- 'What happens if participants deviate from their roles (e.g. provide wrong information) ?'. This issue may or may not have an impact depending on the objective of using CSRP. If one is using CSRP for training the students to use a risk analysis method or determining the performance of a method (e.g. time required running the method), the issue of not having the 'correct' answer from the participant might not impact the study. However, if one is interested in determining a specific vulnerability of a system, then the issue of deviation might have a considerable impact on the result of the study.

The participant should have enough information to play the role correctly. For this, the personas and the scenario developed should be such that it is easy for the participants to engage or mimic it. Extensive research is needed to make sure that the assigned persona are good/ valid representations of the stakeholders rather than depicting the point of view of the person writing the persona (as pointed out by [9]). Thus, constructing the right persona(s) is a challenge for the practitioner.

Even though it is made clear to the participants that they are playing a role, it is likely that they feel their choices represent their personal opinion. Thus, it is important to communicate clearly about the process to the participants at the beginning and also during the implementation phase. Getting inspiration from processes used to train actors may be beneficial. However, there is always a limit to how well this will be grasped by the participants. We are also considering professional or student actors as players in CSRP.

Future Work.

More work is needed to reduce the above identified limitations. We think CSRP can be used with other approaches such as ISO/IEC 27005:2008 [12], Risk IT [10]. However, this needs to be explored further. Furthermore, one possible utilization of CSRP can be to assess the

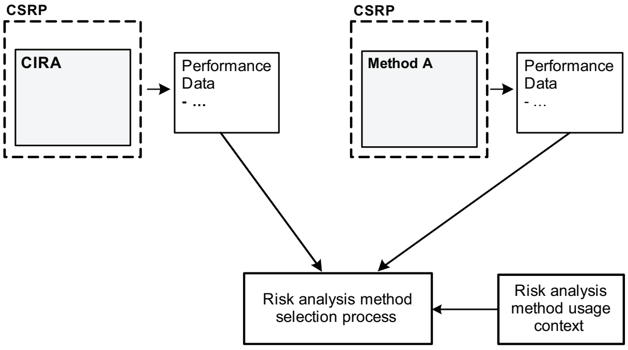

CSRP
CIRA
Performance Data
- ...
CSRP
Method A
Performance Data
- ...
Risk analysis method selection process
Risk analysis method usage context

Figure 10.4: CSRP as part of a risk analysis method selection process

performance of other risk analysis methods e.g. Method A as shown in Figure 10.4. Then, the collected performance data of CIRA and Method A can be compared allowing the most appropriate risk analysis method for a given project to be selected.

10.7 Conclusion

This paper provides an important contribution to information security management. It explains how CSRP can be used as a research and training platform for gaining new knowledge about risk analysis methods while resolving the inherent problem of risk analysis findings' sensitivity and confidentiality. We expect that CSRP will facilitate an increase in the body of published knowledge on the performance of risk analysis methods. It seems reasonable to expect that this platform will result in improved risk analysis methods.

Acknowledgement.

The work reported in this paper is part of the PETweb II project sponsored by The Research Council of Norway under grant 193030/S10.

10.8 Bibliography

[1] Alberts, C., and Dorofee, A. *Managing information security risks, The OCTAVE approach*. Addison Wesley, 2002. ISBN 0-321-11886-3.

[2] AS/NZS 4360. *Risk management*. AS/NZS, 2004.

[3] Atzeni, A., Cameroni, C., Faily, S., Lyle, J., and Flechais, I. Here's Johnny: a Methodology for Developing Attacker Personas. *ARES* (2011), 722–727.

[4] Chulef, A., Read, S., and Walsh, D. A Hierarchical Taxonomy of Human Goals. *Motivation and Emotion 25*, 3 (2001), 191–232(42).

[5] Cooper, A. *The Inmates Are Running the Asylum*. Macmillan Publishing Co., Inc., Indianapolis, IN, USA, 1999.

[6] ETSI TS 102 165-1 V4.2.3 (2011-03). *Method and proforma for Threat, Risk, Vulnerability Analysis*. ESTI, 2011.

[7] Flyvbjerg, B. Five Misunderstandings About Case-Study Research. *Qualitative Inquiry 12*, 2 (2006), 219–245. doi:10.1177/1077800405284363.

[8] Greenberg, J., and Eskew, D. E. The role of role playing in organizational research. *Journal of Management 19*, 2 (1993), 221–241.

[9] Gudjonsdottir, R. *Personas and Scenarios in Use*. Ph.D. thesis, KTH, Human - Computer Interaction, MDI, 2010. QC20100629.

[10] ISACA. *The Risk IT Framework*, 2009.

[11] ISO 31000. *Risk Management – Principles and Guidelines*. ISO, 2009.

[12] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*, 1st ed. ISO/IEC, 2008.

[13] Kotulic, A., and Clark, J. Why there aren't more information security research studies. *Information & Management 41*, 5 (2004), 597–607.

[14] Lund, M. S., Solhaug, B., and Stølen, K. A Guided Tour of the CORAS Method. In *Model-Driven Risk Analysis*. Springer Berlin Heidelberg, 2011, pp. 23–43.

[15] Nielsen, L. From user to character: an investigation into user-descriptions in scenarios. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques* (New York, NY, USA, 2002), DIS '02, ACM, pp. 99–104. doi:10.1145/778712.778729.

[16] NIST. *NIST SP 800-39, Managing Information Security Risk - Organization, Mission, and Information System View*, 2011.

[17] NIST and U.S. Department of Commerce. *NIST SP 800-30 Revision 1, Guide for Conducting Risk Assessments*, September 2012.

[18] Pruitt, J., and Grudin, J. Personas: practice and theory. In *Proceedings of the 2003 conference on Designing for user experiences* (2003), ACM, pp. 1–15.

[19] Rajbhandari, L., and Snekkenes, E. Intended Actions: Risk Is Conflicting Incentives. In *Information Security*, vol. 7483 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 370–386.

[20] Rajbhandari, L., and Snekkenes, E. Using the Conflicting Incentives Risk Analysis Method. In *Security and Privacy Protection in Information Processing Systems*, vol. 405 of *IFIP Advances in Information and Communication Technology*. Springer Berlin Heidelberg, 2013, pp. 315–329. doi:10.1007/978-3-642-39218-4_24.

[21] Stoneburner, G., Goguen, A., and Feringa, A. *NIST SP 800-30, Risk Management Guide for Information Technology*. NIST, 2002.

[22] Yardley-Matwiejczuk, K. M. *Role play: theory and practice*. Sage Publications Limited, 1997.

[23] Yin, R. K. *Case Study Research: Design and Methods*, 4th ed., vol. 5 of *Applied Social Research Method Series*. Sage, 2009.

Chapter 11

Risk Acceptance and Rejection for Threat and Opportunity Risks in Conflicting Incentives Risk Analysis[1]

Abstract

Classical methods for risk analysis usually rely on probability estimates that are sometimes difficult to verify. In particular, this is the case when the system in question is non-stationary or does not have a history for which reliable statistics is available. These methods focus on risks in relation to threats failing to consider risks in relation to opportunity. The Conflicting Incentives Risk Analysis (CIRA) addresses both these issues. Previously, CIRA has been investigated in analyzing threat risks. The paper contributes by illustrating the concept of opportunity risk in the context of CIRA. We give some theoretical underpinnings of risk acceptance and rejection of CIRA, addressing both risks. Furthermore, the paper explains the extension of CIRA to risk management by outlining the risk treatment (response) measures for threat (opportunity) risks.

Keywords: threat risk, opportunity risk, risk acceptance, risk rejection, risk analysis

11.1 Introduction

The Conflicting Incentives Risk Analysis (CIRA) method provides an alternative notion of risk. That is, risk is specified in terms of conflicting incentives between the stakeholders (the risk owner and the strategy owner(s)) in regards to the execution of actions. The risk owner is the stakeholder whose perspective we consider when performing the risk analysis, i.e., he is the stakeholder at risk. The strategy owner is the stakeholder who is capable of triggering an action to increase his perceived benefit. For e.g., when analyzing risks to an end-user of a social networking service, the risk owner is the end-user while the strategy owners can be system administrator, hacker, etc.

Risk is the subjective concern that an individual can feel towards the outcome of events. Taking this perspective, we have two kinds of risks: concern that something undesirable might happen and concern that something desirable might not happen. In the following, we use the term threat (opportunity) to refer to the former (latter). To date, CIRA has been used in analyzing threat risks [9], [10]. In [10], CIRA is used for analyzing privacy risks faced by an end-user in a fictitious case study of an identity management system. Threat risks are caused by intentional execution of strategies by the strategy owner which results in gain for himself and loss for the risk owner. However, the use of CIRA for opportunity risk has not been investigated yet. Opportunity risks are caused by the strategy owners' potential failure to trigger a strategy that the risk owner could reasonably expect that he should trigger.

[1] RAJBHANDARI, L., AND SNEKKENES, E. Risk Acceptance and Rejection for Threat and Opportunity Risks in Conflicting Incentives Risk Analysis. In Trust, Privacy, and Security in Digital Business, vol. 8058 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2013, pp. 124-136.

Most of the works on information security risks analysis and management such as NIST 800-30 [11], ISO/IEC 27005:2008 [7] and CORAS [3] consider only threat risks. This view is backed by researchers in [4], [12], [8]. Hillson [4] agrees that most of the classical risk management methods consider threats while the opportunities are ignored or addressed only reactively. In [8], Olsson also provides the evidence that the current risk management approaches focus on risk rather than opportunity. In economics, one tends to include opportunity risk. We agree with the economic perspective. For instance, we think that the risk that members of staff may fail to take advantage of new security technologies is a risk that must be included in the Chief Information Security Officer's bag of concerns.

To our knowledge, no works have been published addressing how risk acceptance and rejection criteria can be captured and analyzed in the context of CIRA. This paper gives some theoretical underpinnings of risk acceptance and rejection of CIRA method, addressing both threat and opportunity risks. In particular, it highlights and goes some way towards resolving a serious limitation present in other works on risk management- identification and management of opportunity risk in the context of information security management. Furthermore, the paper explains the extension of CIRA to risk management by outlining the risk treatment (response) measures for threat (opportunity) risks.

The remainder of the paper is organized as follows. In Section 11.2, we discuss the related work followed by the overview of CIRA. We explain the threat and opportunity risks in the context of CIRA in Section 11.4 and the details on computing the risk acceptance and rejection bounds are provided in Section 11.5. Section 11.6 outlines the risk treatment (response) measures for threat (opportunity) risks. In Section 11.7, we discuss some issues for further research. Finally, we conclude the paper in Section 11.8.

11.2 Related Work

In classical risk management, risk is often calculated as a combination of the likelihood of an incident and its consequence. The events are usually associated with having adverse/ unfavorable effect. This is further endorsed by the definition: *"risk is a function of the likelihood of a given threat-sources exercising a particular potential vulnerability, and the resulting impact of that adverse event on the organization"* [11]. Thus, typically most of the risk analysis and risk management approaches such as NIST 800-30 [11], ISO/IEC 27005:2008 [7], CORAS [3], OCTAVE [1] and RAMCAP [2] focus on threat risks. At the top level, the risks sources are categorized into human and non-human. Human threat sources further include both the intended and unintended actions of human beings. The non-human threat sources consist of natural threats (e.g. flood, earthquake) and environmental threats (e.g. power failure, system failure). In NIST 800-30, risk from the given threats sources are considered: human (unintentional or deliberate actions), natural and environmental [11]. The ISO/IEC 27005:2008 standard also categorizes the origin of threat into accidental, deliberate and environmental (natural) [7].

The usage of the terms risk and opportunity varies among the risk analyst/ researchers. Some view risk as a term capturing both opportunity and threat [4]. On the other hand, some view opportunity as the opposite of risk [13], [8]. The latter view is usually captured by the term uncertainty; risk is defined as the uncertainty with negative consequences while opportunity is defined as the uncertainty with positive consequences. Even though the view on the definition of opportunity differs, most of the researchers agree that opportunity should be considered, whether being integrated into risk management [4], transforming risk management to uncertainty management [12] or as a separate field which is referred to as opportunity management [13], [8]. This issue has been emphasized mainly in the field of project management.

Hillson [4] explains how an existing risk management method can be extended to incorporate both threats and opportunities by: (1) adding new identification ways to effectively identify opportunities, (2) using double probability-impact matrix for representing both risks and (3) incorporating new strategies to respond to opportunities which are exploit, share, enhance and ignore. Ward et al. [12] argue that both threats and opportunities should be managed and proposes to transform the current project risk management processes into project uncertainty management. Further, White [13] suggests that at the enterprise level, more attention should be given to opportunity management than risk management. The Risk IT [5] framework looks at both IT risk and opportunity in an enterprise. The opportunity is concerned with the benefits that can be achieved (for e.g. identifying new business opportunities from using IT). In ISO 31000 [6], risk is defined as the effect of uncertainty on objectives whether positive or negative. Thus, the guideline can be used to determine risks having both positive or negative consequences.

11.3 Overview of CIRA

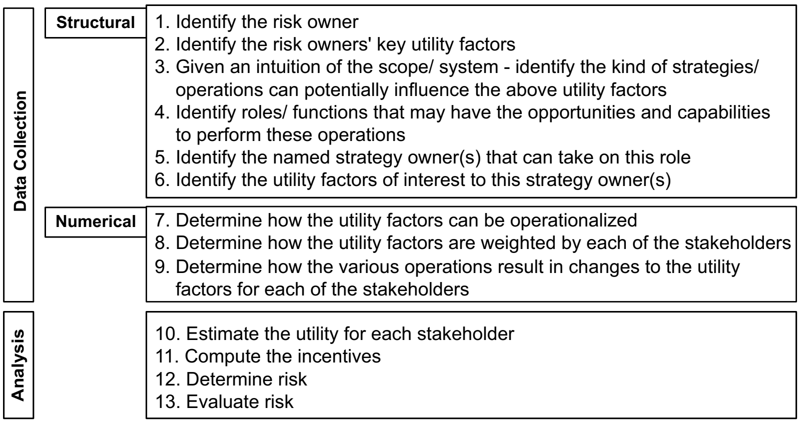
In this section, we provide an overview of CIRA explaining the terms, concepts and procedure. CIRA identifies stakeholders, their actions and perceived expected consequences that characterize the risk situation. As mentioned before, there are two classes of stakeholders: the strategy owner and the risk owner. Typically, each stakeholder has associated a collection of actions that he owns. Risk is modeled in terms of conflicting incentives between the risk owner and the strategy owners in regards to the execution of actions. The actions of the strategy owners may cause threat/ opportunity risks to the risk owner.

Human related risks is the focus of CIRA. This corresponds to understanding the human behavior and incentives that influence their actions. An incentive is something that motivates a stakeholder to take an action to increase his expected/ predicted utility. Utility is the benefit as perceived by the corresponding stakeholder and it comprises of utility factors. Each utility factor captures a specific aspect of utility e.g. prospect of wealth, reputation, legal compliance, ego. Thus, utility can be approximated as the sum of weighted values for utility factors using Multi Criteria Decision Analysis.

When we (the risk analysts) interact with the stakeholders, we assume that their preferences have been determined according to their thinking/ perception about the available options and what is good for them with respect to how others choose their actions. Thus, game theory comes into play in our model as it helps us understand these strategic settings that influence the stakeholders behavior. However, we do not employ game theoretic modeling (i.e. constructing models of strategic settings) and computations.

People consider their self-interest when they interact in a strategic setting and this often gives rise to conflicts. However, we cannot ignore the fact that individuals acting in their self interest sometimes do cooperate to maximize the benefit of all involved. In our setting, it is assumed the stakeholders choose their utility factors according to their self interest and in turn their strategies/ actions to maximize their utility. In other words, the stakeholder will make the move that he thinks will give him the highest gratification.

When identifying and modeling strategy owner actions, the actions modeled correspond to the first move of a game and the effect of the move is modeled as the value of the game as perceived by the strategy owner and risk owner. Note that the strategy owner may be playing a game with multiple stakeholders. For e.g., an attacker will play a game with law enforcement and the legal system. However, we will model this as a strategy that modifies the utility factors of the strategy owner and the risk owner taking into account how the attacker perceives the outcome of the attack including the uncertainty relating to his capture and prospect of penalty. Thus, we do not engage in game theoretic computations of stakeholder behavior but rely on data collection to capture expectations relating to strategy outcomes. The procedure in CIRA is divided into: structural data collection phase (1-6), numerical data collection phase (7-9) and analysis phase (10-13) as depicted in Figure 11.1.



Data Collection

Structural

1. Identify the risk owner
2. Identify the risk owners' key utility factors
3. Given an intuition of the scope/ system - identify the kind of strategies/ operations can potentially influence the above utility factors
4. Identify roles/ functions that may have the opportunities and capabilities to perform these operations
5. Identify the named strategy owner(s) that can take on this role
6. Identify the utility factors of interest to this strategy owner(s)

Numerical

7. Determine how the utility factors can be operationalized
8. Determine how the utility factors are weighted by each of the stakeholders
9. Determine how the various operations result in changes to the utility factors for each of the stakeholders

Analysis

10. Estimate the utility for each stakeholder
11. Compute the incentives
12. Determine risk
13. Evaluate risk

Figure 11.1: Procedure in CIRA

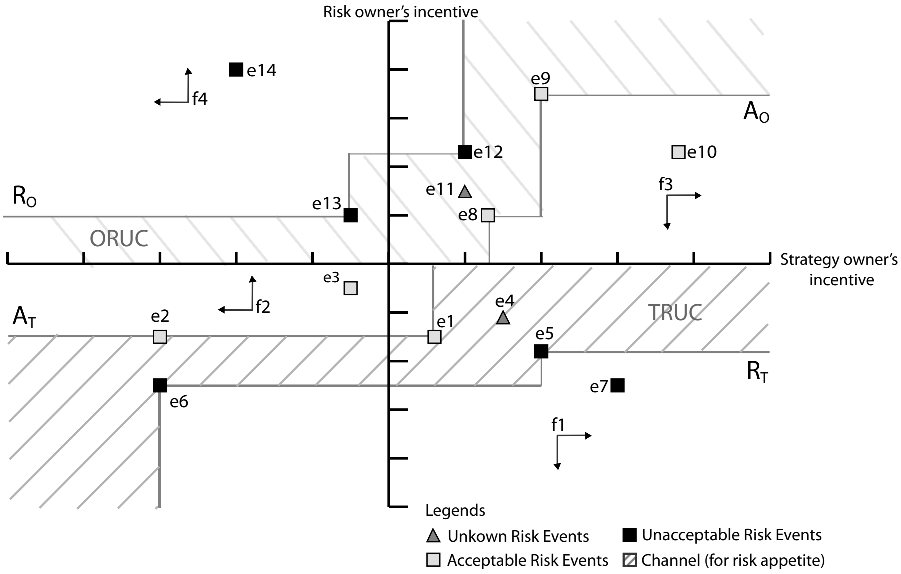


Figure 11.2: CIRA risk visualization using the incentive graph.

11.4 Explaining Risk in the Context of CIRA

In this section, we explain the risks caused by intentional execution of strategies by the strategy owner: threat risk and opportunity risk in the context of CIRA. In classical risk analysis, risk is usually determined as the combination of likelihood and consequence resulting in the unit of Ut^{-1}, where U represents utility and t represents time. On the other hand, in CIRA, risk is the result of conflicting incentives and its unit is U^2.

11.4.1 Risk Visualization

In CIRA, we visualize risk using an incentive graph which is a simple 2 axis coordinate system corresponding to the set of incentives (s, r), where $s(r)$ corresponds to the incentive of the strategy owner (risk taker) as shown in Figure 11.2. Note that all events above (below) the X-axis belong to the collection of opportunity (threat) events. These concepts are described in Table 11.1. The graphs defined by R_O, A_O, A_T, R_T, X-axis in Figure 11.2 partitions the risk plane into 6 non-overlapping areas as described in Table 11.2.

Table 11.1: Legend for CIRA Risk Visualization.

For Threat Risk	
R_T	Rejection boundary of threat risks
A_T	Acceptance boundary of threat risks
e_1 - e_3	Outcome of a typical threat event having an acceptable risk
e_4	Outcome of a typical threat event for which it is not known if the risk is acceptable or not
e_5 - e_7	Outcome of a typical threat event having an unacceptable risk
f_1	Rejection rationality for threat risks
f_2	Acceptance rationality for threat risks
For Opportunity Risk	
R_O	Rejection boundary of opportunity risks
A_O	Acceptance boundary of opportunity risks
e_8 - e_{10}	Outcome of a typical opportunity event having an acceptable risk
e_{11}	Outcome of a typical opportunity event for which it is not known if the risk is acceptable or not
e_{12} - e_{14}	Outcome of a typical opportunity event having an unacceptable risk
f_3	Acceptance rationality for opportunity risks
f_4	Rejection rationality for opportunity risks

Table 11.2: CIRA Plane Partition Legend.

Area bounding	**Explanation**
A_O and X-axis	Acceptable risk from opportunity events (e_{10})
Left and above R_O	Unacceptable risks from opportunity events (e_{14})
A_O and R_O	Information is lacking with respect to the acceptability of the risk of these opportunity events (e_{11}). It is called as the Opportunity Risk Uncertainty Channel (ORUC).
A_T and X-axis	Acceptable risk from threat events (e_3)
Below and to the right of R_T	Unacceptable risk from threat events (e_7)
A_T and R_T	Information is lacking with respect to the acceptability of the risk of these threat events (e_4). It is called as the Threat Risk Uncertainty Channel (TRUC).

11.4.2 The Threat Risk

Previously, CIRA [9], [10] has been restricted to analyzing threat risks i.e. risk facing the risk owner caused by the intentional execution of strategies by the strategy owner which results in gain for himself and loss for the risk owner. The idea being that, risk is the combination of the strength of the force that motivates the strategy owner to send the risk owner to an undesirable state and the magnitude of this undesirability. These risks are usually the consequence of some personal motivations of the strategy owner such as gaining wealth, status, free time, etc. It is reasonable to make the assumption that the strategy owner will be rational in a behavioral economic sense. For e.g., for an end-user of a social networking service, there is uncertainty to the protection of his privacy as his information could be exploited for secondary purposes. In this case, he is facing a threat risk.

11.4.3 The Opportunity Risk

Opportunity risk is the concern that something desirable might not happen. The risk owner is facing opportunity risk when he is concerned that the strategy owner may fail to trigger a strategy that he could reasonably expect that the strategy owner should trigger. The reason being that the strategy owner would have to take a loss in utility and the risk owner would have the prospect of a gain. Likewise, in the threat risk case, it is reasonable to make the assumption that the strategy owner will be rational in a behavioral economic sense. For e.g., if there is uncertainty as to the willingness of a member of staff to spend some effort to identify and deploy more cost effective security products (while maintaining its security posture), the organization is facing opportunity risk.

11.5 Computing Risk Acceptance and Rejection Bounds

Assume we have a collection point sets $P = 2^{U \times U}$, where U denotes the set of utilities. We select a set $D \in P$ of (Incentive, Consequence) pairs and for each of these pairs, we ask the risk owner if he finds the risk associated with this pair acceptable or unacceptable (reject). A risk pair with a negative (positive) consequence is referred to as a threat (opportunity) risk. Then we define $D_{AT}, D_{RT}, D_{AO}, D_{RO} \subseteq D$ such that D_{AT} denotes the set of threat risks to be accepted, D_{RT} the set of threat risks to be rejected, D_{AO} the set of opportunity risks to be accepted and D_{RO} the set of opportunity risks to be rejected. Thus, the risk owner partitions D into the disjoint subsets $D_{AT}, D_{RT}, D_{AO}, D_{RO}$. We require that D_{AT}, D_{RT}, D_{AO}, D_{RO} are non-empty.

We stipulate that the risk owner is rational in the following sense: Let i, j, d be any positive utilities and c be negative for threat consequences and positive for opportunity consequences, then

R1 If a threat consequence c is accepted by the risk owner for a given strategy owner incentive i, then the consequence $c + d$ is acceptable for the incentive $i - j$.

R2 If a threat consequence c is unacceptable (i.e. rejected) by the risk owner for a given strategy owner incentive i, then the consequence $c - d$ is unacceptable for the incentive $i + j$.

R3 If an opportunity consequence c is accepted by the risk owner for a given strategy owner incentive i, then the consequence $c - d$ is acceptable for the incentive $i + j$.

R4 If an opportunity consequence c is unacceptable (i.e. rejected) by the risk owner for a given strategy owner incentive i, then the consequence $c + d$ is unacceptable for the incentive $i - j$.

For example, in the case of opportunity risks, if the gain is relatively modest, you may be prepared to forfeit the gain generated by the strategy. However, if you stand to gain a lot, you will be less inclined to accept the possibility that you may not receive the benefit. Thus, you will require that the strategy owner has strong incentives to implement strategies that would give you large benefits.

The functions defined below compute bounds for acceptance and rejection for both opportunity and threat risks under the assumption that the risk owner is rational in the above sense.

Definition 11.1 (Risk acceptance and rejection bounds)
The lower bound of consequences for threat risks to be accepted, specified as a function of strategy owner incentive is:

$$A_T(x) = \mathit{Min}(\{y' | \exists x' \cdot (x', y') \in D_{AT} \wedge x \leq x'\})$$

The upper bound of consequences for threat risks to be rejected, specified as a function of strategy owner incentive is:

$$R_T(x) = \mathit{Max}(\{y' | \exists x' \cdot (x', y') \in D_{RT} \wedge x' \leq x\})$$

The upper bound of consequences for opportunity risks to be accepted, specified as a function of strategy owner incentive is:

$$A_O(x) = \mathit{Max}(\{y' | \exists x' \cdot (x', y') \in D_{AO} \wedge x' \leq x\})$$

The lower bound of consequences for opportunity risks to be rejected, specified as a function of strategy owner incentive is:

$$R_O(x) = \mathit{Min}(\{y' | \exists x' \cdot (x', y') \in D_{RO} \wedge x \leq x'\})$$

We can then define the rationality closures for the risk acceptance and rejection sets as follows:

Definition 11.2 (Risk acceptance and rejection rationality closures)

$$\begin{array}{rcll} A_T^c & = & \{(i, c) | c \geq A_T(i)\} & \textit{All acceptable threat risks (R1)} \\ R_T^c & = & \{(i, c) | c \leq R_T(i)\} & \textit{All unacceptable threat risks (R2)} \\ A_O^c & = & \{(i, c) | c \leq A_O(i)\} & \textit{All acceptable opportunity risks (R3)} \\ R_O^c & = & \{(i, c) | c \geq R_O(i)\} & \textit{All unacceptable opportunity risks (R4)} \end{array}$$

Theorem 11.1 (Rationality closures)
All elements in $D_{AT}, D_{RT}, D_{AO}, D_{RO}$ belong to the corresponding rationality closures, i.e.

$$\begin{array}{rcl} D_{AT} & \subseteq & A_T^c \\ D_{RT} & \subseteq & R_T^c \\ D_{AO} & \subseteq & A_O^c \\ D_{RO} & \subseteq & R_O^c \end{array}$$

PROOF By expansion and noting that for any closed boolean expressions $P(.)$, a and b:

$$\begin{array}{rcl} P(a) & \Rightarrow & a \leq \mathrm{Max}(\{y | P(y)\}) \\ P(b) & \Rightarrow & b \geq \mathrm{Min}(\{y | P(y)\}) \end{array}$$

Theorem 11.2 (The rationality closures extends the acceptance and rejection bounds) *The rationality closures extends the acceptance and rejection bounds. I.e. for all $i, j, d \geq 0$ then*

$$\begin{aligned} c \leq 0 \wedge (i, c) \in A_T^c \Rightarrow (i - j, c + d) \in A_T^c \quad &\textit{(R1)} \\ c \leq 0 \wedge (i, c) \in R_T^c \Rightarrow (i + j, c - d) \in R_T^c \quad &\textit{(R2)} \\ c \geq 0 \wedge (i, c) \in A_O^c \Rightarrow (i + j, c - d) \in A_O^c \quad &\textit{(R3)} \\ c \geq 0 \wedge (i, c) \in R_O^c \Rightarrow (i - j, c + d) \in R_O^c \quad &\textit{(R4)} \end{aligned}$$

PROOF By expansion and noting that $a + b \geq a$ for $b \geq 0$, $e - f \leq e$ for $f \geq 0$ and for any closed boolean expression $P(.)$:

$$\begin{aligned} \text{Min}(\{y | \exists x \cdot P(x, y) \wedge (i - j) \leq x\}) &\leq \text{Min}(\{y | \exists x \cdot P(x, y) \wedge i \leq x\}) \\ \text{Max}(\{y | \exists x \cdot P(x, y) \wedge x \leq i\}) &\leq \text{Max}(\{y | \exists x \cdot P(x, y) \wedge x \leq i + j\}) \end{aligned}$$

when $j \geq 0$. ■

The acceptance and rejection sets are mutually consistent for threats (opportunities) iff their rational extensions are non-overlapping. Given a risk acceptance (rejection) set A (R), we define opportunity (OC(.)) and threat (TC(.)) risk acceptance/rejection consistency as

Definition 11.3 (Risk acceptance and rejection consistency)

$$OC(R, A) = TC(A, R) = \forall i, j, c, d \cdot (i, c) \in A \wedge (j, d) \in R \Rightarrow i < j \vee c > d$$

Theorem 11.3 (The rationality closure is consistency preserving) *The rationality closure is consistency preserving. I.e.*

$$\begin{aligned} TC(A_T, R_T) &\Rightarrow TC(A_T^c, R_T^c) \\ OC(A_O, R_O) &\Rightarrow OC(A_O^c, R_O^c) \end{aligned}$$

PROOF Since the acceptance and rejection sets are mutually consistent, noting that there is a transitivity property between position of the point in A_T, R_T and the point in the closure, it suffices to show that each element in the accept (reject) closure is 'on the correct side' of some point in the corresponding partition of D. But this holds by Lemma 11.4. ■

Lemma 11.4 (All elements in a closure are bounded by some element) *All elements in a closure are bounded by some element in the corresponding partition of D:*

$$\begin{aligned} \forall i, c \cdot (i, c) \in A_T^C &\Rightarrow \exists j, d \cdot (j, d) \in D_{AT} \wedge j \geq i \wedge d \leq c \\ \forall i, c \cdot (i, c) \in R_T^C &\Rightarrow \exists j, d \cdot (j, d) \in D_{RT} \wedge j \leq i \wedge d \geq c \\ \forall i, c \cdot (i, c) \in A_O^C &\Rightarrow \exists j, d \cdot (j, d) \in D_{AO} \wedge j \leq i \wedge d \geq c \\ \forall i, c \cdot (i, c) \in R_O^C &\Rightarrow \exists j, d \cdot (j, d) \in D_{RO} \wedge j \geq i \wedge d \leq c \end{aligned}$$

PROOF By expansion, using the existential witness obtained from the antecedent, we can easily construct the existential witness required in the consequent. Noting that $\forall x \cdot x \leq max(y | P(y)) \Rightarrow \exists z \cdot x \leq z \wedge P(z)$ and $\forall x \cdot x \geq min(y | P(y)) \Rightarrow \exists z \cdot x \geq z \wedge P(z)$. ■

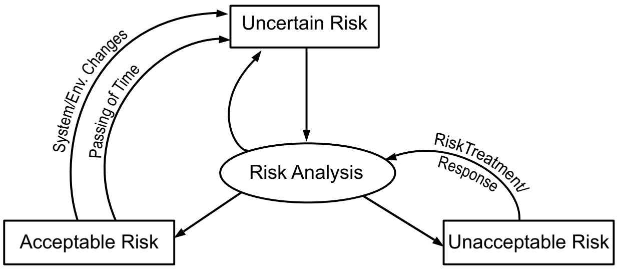


Figure 11.3: CIRA Risk Exposure States and Transitions.

We can easily restrict acceptance and rejection closure sets to the corresponding consistent subset as follows:

$$\begin{aligned} A^c_{TC} &= A^c_T \setminus R^c_T \\ R^c_{TC} &= R^c_T \setminus A^c_T \\ A^c_{OC} &= A^c_O \setminus R^c_O \\ R^c_{OC} &= R^c_O \setminus A^c_O \end{aligned}$$

However, in practice, rather than restricting the acceptance and rejection sets, one would engage in a dialogue with the risk owner such as to ensure that the partitions of D (i.e. $D_{AT}, D_{RT}, D_{AO}, D_{RO}$) are mutually consistent.

11.6 Risk Treatment (Response) Measures for Threat (Opportunity) Risks

In this section, we explain the risk treatment (response) measures for threat (opportunity) risks. The overall process for risk management in CIRA is depicted in Figure 11.3. Risk analysis helps to identify and estimate risks, and provide insight suitable for deciding if risk exposure needs to be changed. That is, if a treatment/ response action is needed, or risk exposure may be increased. Further, risk management is taking actions to treat/ respond to those risks that are not within the risk acceptance criteria. The treatment measures for the threat risks include: mitigate, avoid, transfer and accept (i.e. accept the risk without taking any action). On the other hand, the response measures for the opportunity risks include enhance, exploit, share and ignore [4].

The risk exposure is either acceptable, unacceptable or uncertain/ unknown. By doing risk analysis, we can determine if the exposure is acceptable or not. If the system or the environment changes, we may no longer have sufficient evidence to conclude that the exposure is acceptable. Similarly, over time, the system and its environment may be exposed to changes that are not easily discoverable. Thus, we are drifting towards a state of unknown exposure. When implementing a risk treatment/ response measure, this measure may only have a partial effect and it may also have side effects giving rise to new vulnerabilities. The risk analysis process further helps to decide the risk exposure.

In classical risk management, risk is managed through reduction of incident likelihood and consequence. Since CIRA does not adopt the likelihood and consequence paradigm, our risk management goals will be somewhat different. In CIRA, risk treatment/ response amounts to the modification of perceived utility caused by the strategies in questions. That

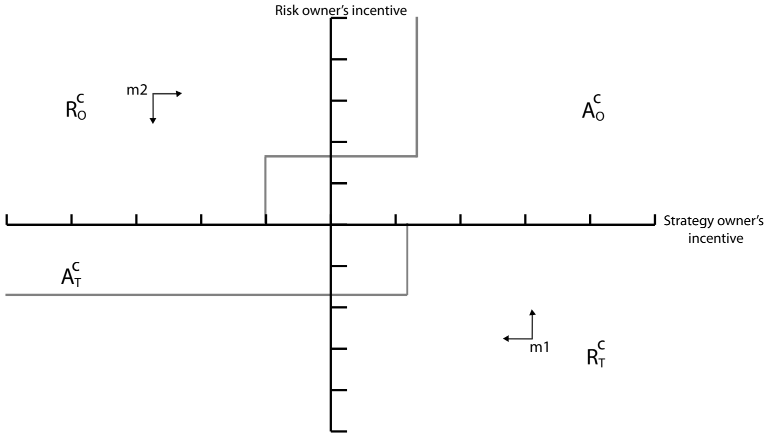


Figure 11.4: CIRA Risk Management Strategies.

is, a risk treatment/ response measure aims to modify the weights that the stakeholders assign to the relevant utility factors or modify the incentives of the stakeholders. This is illustrated in Figure 11.4 and described in more detail below. Threat risks can be mitigated through a combination of strategy owner incentive reduction and risk owner incentive increase. In Figure 11.4, the arrows in m_1 represent the desired direction that we want to move the outcome through mitigation. In the above threat risk example, risk faced by the end-user can be reduced if privacy rules and regulations (that govern when and how the services use/ collect personal information of customers) are established and enforced. In the case of opportunity risks, the primary risk response strategy is to increase the strategy owner incentives to implement the strategy (represented by the rightward arrow in m_2). We may also respond to the risk by reducing the utility of the risk owner (represented by the downward arrow in m_2). However, sometimes we may also want to increase the utility of the risk owner. In the above opportunity risk example, the staff can be incentivized by enforcing and communicating rules within the organization or providing free training so that they are willing to deploy more cost effective security products which in turn benefits the organization.

11.7 Future Work

Some of the potential issues for further work include the exploration of the newly introduced theoretical concepts on opportunity risk, risk treatment and response measures. More case studies needs to be conducted to explore the different risks, and to validate and improve the method.

In CIRA, risk acceptance criteria is determined after the decision input parameters are determined unlike in most of the classical methods. For instance, in the ISO/IEC 27005:2008 standard, it is required that the risk acceptance criteria should be determined before the threat and vulnerability discovery is carried out. However, this may not be economically rational. For e.g., we may have a situation where a risk that is low is unacceptable because the mitigation effort required to control the risk is very low. Similarly, we may accept a very high risk if all response options have little or no effect. Thus, an economically rational actor will determine his risk acceptance threshold on at least the following information: incident risk, risk mitigation cost and the effectiveness of mitigation measures. Thus, it is questionable if it is a good strategy to fix the risk acceptance criteria as an expected value before the relevant decision input parameters have been determined.

This issue can be further investigated.

11.8 Conclusion

This paper has explained the key concepts of risk acceptance and rejection in the CIRA method. Definitions have been formalized and we have included some theorems establishing some consequences of our definitions. Our definitions and model introduce the concept of opportunity risk in the context of information security management. The opportunity risk remains overlooked and needs more emphasis by current research on risk management. A comprehensive Conflicting Incentives Risk Analysis and Management (CIRAM) method which considers and addresses both threat and opportunity risks has the potential to enhance the overall risk management process.

Acknowledgement.

The work reported in this paper is part of the PETweb II project sponsored by The Research Council of Norway under grant 193030/S10.

11.9 Bibliography

[1] Alberts, C., and Dorofee, A. *Managing information security risks, The OCTAVE approach*. Addison Wesley, 2002. ISBN 0-321-11886-3.

[2] ASME Innovative Technologies Institute, LLC. *Risk Analysis and Management for Critical Asset Protection (RAMCAP): The Framework*, May 2006. Version 2.0.

[3] Braber, F., Hogganvik, I., Lund, M. S., Stølen, K., and Vraalsen, F. Model-based security analysis in seven steps — a guided tour to the CORAS method. *BT Technology Journal 25*, 1 (2007), 101–117. doi:10.1007/s10550-007-0013-9.

[4] Hillson, D. Extending the risk process to manage opportunities. *International Journal of Project Management 20*, 3 (2002), 235–240. doi:10.1016/S0263-7863(01)00074-6.

[5] ISACA. *The Risk IT Framework*, 2009.

[6] ISO 31000. *Risk Management – Principles and Guidelines*. ISO, 2009.

[7] ISO/IEC 27005. *Information technology -Security techniques -Information security risk management*, 1st ed. ISO/IEC, 2008.

[8] Olsson, R. In search of opportunity management: Is the risk management process enough? *International Journal of Project Management 25*, 8 (2007), 745–752. doi:10.1016/j.ijproman.2007.03.005.

[9] Rajbhandari, L., and Snekkenes, E. Intended Actions: Risk Is Conflicting Incentives. In *Information Security*, vol. 7483 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 370–386.

[10] Rajbhandari, L., and Snekkenes, E. Using the Conflicting Incentives Risk Analysis Method. In *Security and Privacy Protection in Information Processing Systems*, vol. 405 of *IFIP Advances in Information and Communication Technology*. Springer Berlin Heidelberg, 2013, pp. 315–329. doi:10.1007/978-3-642-39218-4_24.

[11] Stoneburner, G., Goguen, A., and Feringa, A. *NIST SP 800-30, Risk Management Guide for Information Technology*. NIST, 2002.

[12] Ward, S., and Chapman, C. Transforming project risk management into project uncertainty management. *International Journal of Project Management 21*, 2 (2003), 97–105. doi:10.1016/S0263-7863(01)00080-1.

[13] White, B. E. Enterprise Opportunity and Risk. *INCOSE Symposium, Orlando, FL* (2006).

Nomenclature

CIRA	Conflicting Incentives Risk Analysis
CIRAM	Conflicting Incentives Risk Analysis and Management
CRAMM	CCTA Risk Analysis and Management Method
CSRP	Case Study Role Play
ENISA	European Network and Information Security Agency
ERM	Enterprise Risk Management
GUC	Gjøvik University College
IdMS	Identity Management System
ISO/IEC	International Organization for Standardization/International Electrotechnical Commission
ISRAM	Information Security Risk Analysis Method
IST	Information Society Technologies
MAUT	Multi Attribute Utility Theory
NISlab	Norwegian Information Security laboratory
NIST	National Institute of Standards and Technology
OCTAVE	Operationally Critical Threat, Asset, and Vulnerability Evaluation
PIA	Privacy Impact Assessment
PRA	Probabilistic Risk Analysis
QuERIES	Quantitative Evaluation of Risk for Investment Efficient Strategies
RAMCAP	Risk Analysis and Management for Critical Asset Protection
RCN	Research Council of Norway
SAM	System-Action-Management
TISPAN	Technical Committee Telecommunications and Internet converged Services and Protocols for Advanced Networking
TVRA	Threat Vulnerability and Risk Analysis
UML	Unified Modeling Language

Index