2016

Evidential Reasoning for Forensic Readiness

Yi-Ching Liao
Norwegian University of Science and Technology

Hanno Langweg
Norwegian University of Science and Technology

EVIDENTIAL REASONING FOR FORENSIC READINESS

Yi-Ching Liao and Hanno Langweg

Norwegian University of Science and Technology

Teknologivn. 22, 2815 Gjøvik, Norway

{yi-ching.liao2, hanno.langweg}@ntnu.no

ABSTRACT

To learn from the past, we analyse 1,088 "computer as a target" judgments for evidential reasoning by extracting four case elements: decision, intent, fact, and evidence. Analysing the decision element is essential for studying the scale of sentence severity for cross-jurisdictional comparisons. Examining the intent element can facilitate future risk assessment. Analysing the fact element can enhance an organization's capability of analysing criminal activities for future offender profiling. Examining the evidence used against a defendant from previous judgments can facilitate the preparation of evidence for upcoming legal disclosure. Following the concepts of argumentation diagrams, we develop an automatic judgment summarizing system to enhance the accessibility of judgments and avoid repeating past mistakes. Inspired by the feasibility of extracting legal knowledge for argument construction and employing grounds of inadmissibility for probability assessment, we conduct evidential reasoning of kernel traces for forensic readiness. We integrate the narrative methods from attack graphs/languages for preventing confirmation bias, the argumentative methods from argumentation diagrams for constructing legal arguments, and the probabilistic methods from Bayesian networks for comparing hypotheses.

Keywords: legal knowledge extraction, forensic readiness, evidential reasoning, kernel tracing, attack description languages, argumentation diagrams, Bayesian networks

1. INTRODUCTION

"A case is only as strong as its evidence" (Torpey, 2009). Many computer crime cases are closed due to the lack of evidence, and the most common reason is not preparing for upcoming legal disclosure at all, which is equivalent to covering the tracks for perpetrators (Rowlingson, 2004). Moreover, covering the tracks has become the standard operating procedure for perpetrators (Graves, 2007). To meet the burden of proof, we should establish an information retention strategy to prepare for upcoming legal disclosure (International Organization for Standardization, 2015), instead of merely relying on the evidence remained.

Since digital evidence is more susceptible to tampering and subsequent modification than traditional documents, it causes more uncertainties in legal cases, such as using the Trojan horse defense to avoid a conviction (Brenner, Carrier, & Henninger, 2004). What is worse, forged digital evidence can result in miscarriages of justice, such as

framing victims by planting digital evidence. To prevent miscarriages of justice caused by false evidence, we have to consider evidence reliability early in the collection phase by choosing more reliable evidence sources, instead of merely securing evidence with tamper resistance and detection techniques later in the preservation phase.

To facilitate the preparation of reliable evidence before the occurrence of legal actions, we emphasise the domain of computer crime and analyse the "computer as a target" judgments for evidential reasoning through case element extraction in Section 2. Inspired by the effectiveness of case elements and the feasibility of transforming them into argumentation diagrams, we integrate the narrative methods, the argumentative methods, and the probabilistic methods to produce scenarios through attack graphs/languages, generate hypotheses through argumentation diagrams, and assess probabilities through Bayesian networks in Section 3. Section 4 concludes the paper with the discussion of the effectiveness of method integration for evidential reasoning.

2. ANALYSING JUDGMENTS FOR EVIDENTIAL REASONING

To learn from the past, we analyse 1,088 "computer as a target" judgments through language engineering techniques for evidential reasoning. We explain our analysis methods and findings as follows:

2.1 Related Work

Distinguished from previous studies on legal knowledge extraction, we emphasise the analysis of "computer as a target" judgements written in Chinese characters from four case elements: decision, intent, fact,

and evidence. Table 1 presents a comparative analysis of previous research on legal text summarization and analysis from three perspectives: language, corpus, and elements. Previous studies mostly collect and analyse judgments written in English, and mainly focus on general criminal cases. To extend the research scope, we emphasise the domain of computer crime and analyse the "computer as a target" judgments for evidential reasoning.

2.2 Collecting "Computer as a Target" Judgments

The criminal activities targeting at computers demand more technical knowledge than the activities utilizing computers as tools. Therefore, we emphasise the analysis of "computer as a target" judgments for evidential reasoning. Since different jurisdictions have their own computer crime laws, we need to employ different search terms based on the corresponding computer crime laws for judgment collection.

Chinese judgments are available at the website wenshu.court.gov.cn (China Judgements Online) for public access. The Criminal Law of the People's Republic of China clearly differentiates between the "computer as a target" crime and the "computer as a tool" crime in separate articles. Therefore, we employ "computer information system" as the search term for "case name" to collect judgments regarding "illegal invasion of computer information system" and "deleting, altering, adding or jamming the functions of the computer information system" (Council of Europe, 2008). We employ `Scrapy` (Scrapinghub, Ltd., 2015), an open source Web crawler, and utilize the XML path language to locate and retrieve the judgment content. We categorize the collected Chinese judgments into two judgment types: first instance and appeal. We

Table 1. A Summary of Research on Legal Text Summarization and Analysis

Author(s)	Language	Corpus	Elements
Gelbart and Smith	English	more than 1000 economic loss cases and general British Columbia cases	legal concepts, statute citations, case citations, and facts
Schweighofer et al.	English	75 full text documents of court decisions from the European Community law database	index with the thesaurus entries
Uyttendaele et al.	Dutch	more than 3000 decisions from the correctional Court of Leuven	superscription, victim, accused, alleged offenses, transition formulation, opinion of the court, legal foundations, verdict, and conclusion
Farzindar and Lapalme	English	3500 judgments from the Federal Court of Canada	four thematic segmentations: introduction, context, juridical analysis, and conclusion
Hachey and Grover	English	188 judgments from the UK House of Lords	seven rhetorical annotations: fact, proceedings, background, framing, disposal, textual, and other
Saravanan et al.	English	200 judgments related to rent control, income tax and sales tax	seven rhetorical roles: identifying the case, establishing facts of the case, arguing the case, history of the case, arguments, ratio of the decision, and final decision
Chieze et al.	English French	14,380 historical decisions from Canadian federal courts and provincial tribunals	four thematic segments (Mailhot & Carnwath, 1998): introduction, context, reasoning, and conclusion
Wyner	English	47 criminal cases drawn from the California Supreme Court and State Court of Appeal	case citations, names of parties, roles of parties, and final decision
Yousfi-Monod et al.	English French	3715 decisions from the Canadian courts	four decision sections: introduction, context, reasoning and conclusion
Galgani et al.	English	5705 case reports from the Federal Court of Australia	collected catchphrases, such as courts and judges, corporations, costs, etc.

found 26 duplicated judgments, two judgments without content, and four irrelevant judgments while collecting Chinese judgments.

As for Taiwan, the Judicial Yuan makes legislative and judicial texts accessible at the website jirs.judicial.gov.tw, and we utilize "offenses against the computer security" as the search term for "cause of action" and "criminal case" as "judgment type" for judgment collection. Since certain character in the case number represents different judgment types, we can categorize the collected Taiwanese judgments into six judgment types: prosecution, summary judgment, appeal, civil action, private prosecution, and mediation. The most common types of Taiwanese judgments are prosecution and summary judgment, which form 88.9% of the collected Taiwanese judgments. The average character count of summary judgments is around half of the prosecution judgments, which are 2820.58, and 5801.13 respectively. We also found two duplicated judgments while collecting Taiwanese judgments. Until May 29th 2016, we totally col-

lect 1,088 "computer as a target" judgments: 358 Chinese judgments and 730 Taiwanese judgments.

Figure 1 shows the number of "computer as a target" judgments by judgment year. The Criminal Law of the People's Republic of China incriminates "illegal invasion of computer information system" and "deleting, altering, adding or jamming the functions of the computer information system" in 1997, whereas Criminal Code of the Republic of China incriminates "offenses against the computer security" in 2003. However, there is only one Chinese judgment concerning "computer as a target" crime before 2011, and the number of judgments collected in China is less than in Taiwan. The number of "computer as a target" judgments is much less than expected considering the number of incidents reported.

2.3   Extracting Case Elements

To analyse the "computer as a target" judgments for evidential reasoning, we determine four case elements for legal knowledge extraction: decision, intent, fact, and evi-
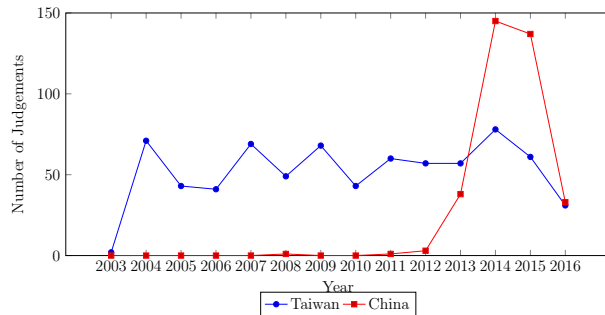
Figure 1. Number of "computer as a target" Judgments by Year

dence. Examining the decision element is essential for studying the scale of sentence severity, which can serve as a criterion for cross-jurisdictional comparisons. Analysing the intent element is indispensable to threat analysis, which can facilitate future risk assessment. Examining the fact element can enhance an organization's capability of understanding criminal behaviours for future offender profiling. Analysing the evidence used against a defendant from previous judgments can facilitate the preparation of evidence before the occurrence of legal actions.

To extract case elements, we employ five processing resources (Cunningham et al., 2014) provided by GATE (Cunningham, 2002), an open source language engineering software, to locate annotations from the "computer as a target" judgments, which we list as follows:

1. RegEx Sentence Splitter annotates sentences according to the files containing regular expressions of sentence splits, which we utilize as the indicators of different judgment stages: decision, fact, and reasoning. Regarding a judgment as a transaction, Cheng (2010) determines generic structures of Chinese and Taiwanese judgments by identifying various stages and legal functions. The orders of stages are different between Chinese and Taiwanese judg-

ments, which are shown in Table 2. Chinese judgments establish the fact first, then justify the decision, and finally declare the decision. On the other hand, Taiwanese judgments declare the decision first, then establish the fact, and eventually justify the decision.

Despite different stage orders, the case elements exist within the same stages. For Chinese judgments, we utilize RegEx Sentence Splitter to locate the linguistic markers of the fact and evidence elements, which we later extract the first paragraph between them as the fact element by JAPE Transducer. As for Taiwanese judgments, we employ RegEx Sentence Splitter to locate the linguistic markers as the separator between the decision and intent elements.

2. GATE Unicode Tokeniser divides judgments into smaller tokens, such as punctuations and numbers. We customize GATE Unicode Tokeniser to annotate one or more occurrences of Unicode character between punctuations as a token, which we regard as a sentence for further higher-level annotations by JAPE Transducer.

3. ANNIE Gazetteer aims to search the pre-defined words for further annotation by JAPE Transducer. We set the "wholeWordsOnly" parameter to false to allow ANNIE Gazetteer to match not only whole words. Due to the different judgment structures, we employ different linguistic markers to different jurisdictions, which are summarized in Table 2. Currently we generate these linguistic markers manually through analysing the indicators of different judgment stages and case elements. We only update the linguistic markers to processing resources after

Table 2. The Analysis of Judgments for Case Element Extraction

Stage	Element	Linguistic Markers	Processing Resources
		China	
fact	intent	*intent*=[purpose, for, used]	ANNIE Gazetteer
		the sentence which contains *intent*	JAPE Transducer
	fact	*fact_start*=[the prosecutor accused, procuratorate accused, after being examined] *evidence_start*=[the above fact, given with evidence, to support the above allegation]	RegEx Sentence Splitter
		the first paragraph between *fact_start* and *evidence_start*	JAPE Transducer
reasoning	evidence	*evidence_start*=[the above fact, given with evidence, to support the above allegation] *evidence_end*=[this court thinks, sufficient to justify, is consistent with fact]	ANNIE Gazetteer
		the sentences between *evidence_start* and *evidence_end*	JAPE Transducer
decision	decision	*decision_start*=[the decision is as follows, based on the defendant] *decision*=[is sentenced]	ANNIE Gazetteer
		after *decision_start*, the sentence which contains *decision*	JAPE Transducer
		Taiwan	
decision	decision	*fact_start*=[reason, crime fact]	RegEx Sentence Splitter
		guilty=[detention, imprisonment, probation, fine] *length*=[day, month, year, dollar] *other*=[not guilty, dismissal, overrule, dismissal, jurisdictional error, summary judgement, grounds for appeal]	ANNIE Gazetteer
		before *fact_start*, the sentences which contain *guilty* and *length* or contain *other*	JAPE Transducer
fact	intent	*fact_start*=[reason, crime fact]	RegEx Sentence Splitter
		intent=[intent, intent, mens rea, based, for not, discontent]	ANNIE Gazetteer
		after *fact_start*, the sentence which contains *intent*	JAPE Transducer
	fact	the paragraph after the sentence which contains *intent*	JAPE Transducer
reasoning	evidence	*evidence_list_start*=[evidence, evidence of fact, basis of fact, evidence name, evidence part, based on the evidence, the following evidence, sufficient evidence to prove, the following evidence to prove] *evidence_start*=[based on, the above fact, the above crime fact, mainly based on, basis of crime fact] *evidence_ref*=[indictment, as reference]	ANNIE Gazetteer
		after *evidence_list_start*, the sentences which contains parentheses the paragraph which contains *evidence_start* the sentence which contains *evidence_ref*	JAPE Transducer

verifying the enhanced annotation coverage. However, the diverse usages of legal terms by different judges obstruct the automatic generation of linguistic markers.

4. `JAPE Transducer` recognises higher-level annotations through regular expressions. For example, after we utilize `ANNIE Gazetteer` to locate the linguistic markers for the start of the decision element "the decision is as follows" and the keyword of the decision element "is sentenced" from Chinese judgments, we extract the sentences which contain "is sentenced" after "the decision is as follows" as the decision elements by `JAPE Transducer`.

5. `Flexible Exporter` saves the annotations to files. After locating four case elements from judgments, we employ `Flexible Exporter` to store the annotations for further analysis and visualization.

2.4 Verifying Annotation Coverage

Table 3 presents the number of judgments collected, the average character count, and the number of case element not found according to different judgment types. Since there is no frequently used term to indicate the intent in Chinese judgments, we fail to extract the intent element from 134 Chinese judgments, from which we can observe that the pre-defined words in `ANNIE Gazetteer` lists greatly influence the accuracy of element extraction. Analysing the annotation coverage from the judgment types, we fail to extract almost all the case elements from 26 Taiwanese judgments of civil action and mediation types due to the lack of element in these judgments. Moreover, since we utilize `RegEx Sentence Splitter` to locate the linguistic markers as the separator between the decision and intent elements for Taiwanese judgments, extracting the decision element will fail if there is no separator for `RegEx Sentence Splitter` to

locate. Not every Taiwanese judgment has the same judgment stages, even though Taiwanese judgments have more identifiable linguistic markers to separate judgment stages than Chinese judgments.

Table 3. A Summary of Collected Chinese and Taiwanese Judgments

China						
Judgment Type	Avg. Char Count	Judgment Number	Number of Case Element not Found			
			Decision	Intent	Fact	Evidence
first instance	4160.06	302	4	110	3	22
appeal	5922.93	56	42	24	12	17
Taiwan						
prosecution	5801.13	331	32	119	121	199
summary judgment	2820.58	318	14	97	97	149
appeal	5138.14	50	2	13	13	22
civil action	677.52	23	18	23	23	23
private prosecution	5145.40	5	1	1	1	2
mediation	513.00	3	3	3	3	3

2.5 Judgment Analysis Findings

After analysing 1,088 "computer as a target" judgments through case element extraction, we describe the major findings as follows:

2.5.1 Effectiveness of Case Elements

The decision element serves as a good criterion to analyse the scale of sentence severity for cross-jurisdictional comparisons. Regarding not guilty as zero days, the average length of imprisonment is 962.07 days based on the decision elements extracted from Chinese judgments. On the other hand, the average length of imprisonment is 170.46 days according to Taiwanese decision elements, which is not that severe compared to the maximum three or five years imprisonment regulated in criminal code on offenses against the computer security. We can observe the defendants who commit computer crime tend to be sentenced more severely in China than those in Taiwan, even though the number of judgments collected in China is less than in Taiwan. Note that the written numbers are different between Taiwanese and Chinese judgments; most of Taiwanese judgments use financial characters, whereas Chinese judgments use normal characters.

Studying the intent element is beneficial for future risk assessment, even though it is difficult to have solid evidence to prove the intent of a defendant. Since threat is one of the indispensable components of risk, and threat possesses intent and capability, examining the intent element is advantageous for risk assessment and determining the level of security control to protect information assets. According to 224 intent elements extracted from Chinese judgments, the most common intent is to gain the illegal benefits. On the other hand, based on 474 intent elements extracted from Taiwanese judgments, the intent "for purpose to exercise unlawful control over others property" is more popular than the intent of gaining the illegal benefits.

The fact elements extracted are well constructed and describe the events that previously occurred in narrative forms, which can facilitate the analysis of criminal activities for further crime scenario construction and offender profiling. Based on 343 fact elements extracted from Chinese judgments, the common activities include installing Trojan horse for stealing money from on-line accounts, and even selling bots for distributed denial-of-service attacks. According to 472 fact elements extracted Taiwanese judgments, the most common behaviour is the illegal possession of others' accounts.

Analysing the evidence element facilitates the preparation of evidence before the occurrence of legal actions. An organization should establish an information retention strategy to prepare for upcoming legal disclosure (International Organization for Standardization, 2015), which can be easily clarified by analysing the evidence used against a defendant from previous judgments. However, based on the total 651 evidence elements extracted Chinese and Taiwanese judgments, we observe the digital evidence

supporting the "computer as a target" cases rarely stands alone. Even the "computer as a target" cases still depend heavily on non-digital evidence, such as testimonials, documents, or confessions.

2.5.2 Transformation into Argumentation Diagrams

Based on the case elements extracted from judgments, we can follow the concepts of datum and claim from the Toulmin model of argument (Toulmin, 2003), which are shown in Figure 2, to develop an automatic judgment summarizing system. Taking the file generated by `Flexible Exporter` as input, we visualize the case elements with DOT, a plain text graph description language, and employ `Graphviz` (Ellson, Gansner, Koutsofios, North, & Woodhull, 2001), an open source graph drawing tool, to convert the DOT files to PDF files.
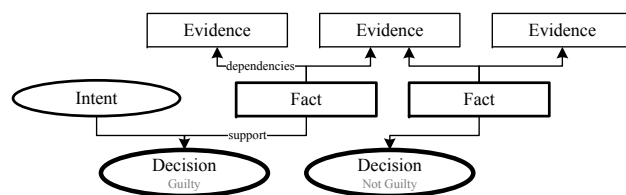
Figure 2. Transformation of Case Elements into Argumentation Diagrams

Figure 3 presents an overview of the measures and procedures for automatic judgment summarization. Figure 4 presents an automatically generated argumentation diagram from a Chinese judgment. Using rectangles to represent a fact or an observation and ovals to present a consequent assertion (Shum, 2003), we summarize the "computer as a target" judgments automatically in a consistent format, which can enhance the accessibility and readability of judgments. Since the fact element describes the previously occurred events in chronological order, we divide the fact element into shorter phrases, and list the phrases from top to bottom for better comprehensibility of criminal activities. We also categorize the evidence element into digital evidence and non-digital evidence to demonstrate that the "computer as a target" judgments do not depend heavily on digital evidence as expected.
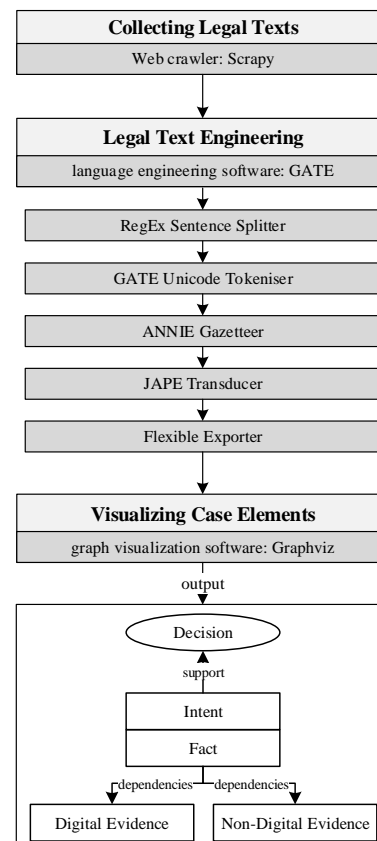
Figure 3. An Overview of an Automatic Judgment Summarizing System

2.5.3 Grounds of Inadmissibility

The gathered evidence only become admissible when it meets the criteria of relevance and reliability. Based on collected Chinese and Taiwanese judgments, the grounds of inadmissibility include processed by unqualified personnel, diverse malware analysis result, inaccurate calculation of monetary value, no evidential connection, reasonable doubt, etc. Analysis outcomes of the admissibility and inadmissibility of evidence from
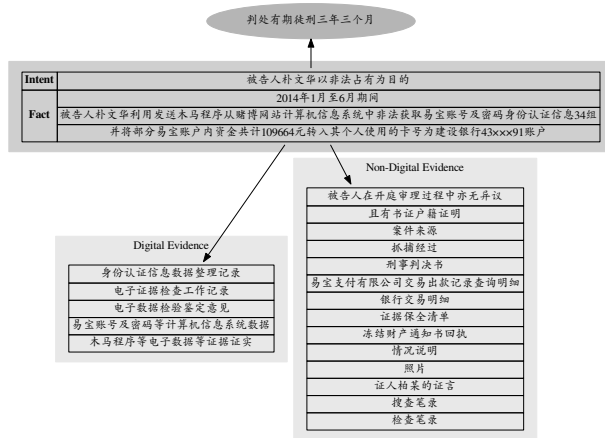
Figure 4. An Automatically Generated Argumentation Diagram from a Chinese Judgment

the "computer as a target" judgments are valuable inputs for probability assessment. Extracting 5W1H information and other factors affecting the scale of sentence severity can also assist the calculation of probabilities.

2.5.4 Need for Forensic Readiness

Since the criminal activities targeting at computers demand more technical knowledge than general criminal cases, the "computer as a target" judgments are supposed to depend heavily on digital evidence. However, based on the evidence elements extracted, we discover the digital evidence rarely stands alone, even for supporting the "computer as a target" cases. Moreover, since the number of incidents reported are much more than the number of judgments, there must be many "computer as a target" cases closed before being brought to courts. It is essential for organisations to prepare themselves for digital forensics with reliable digital evidence, in other words- forensic readiness.

3. EVIDENTIAL REASONING OF KERNEL TRACES

To learn from the past, we analyse previous "computer as a target" judgments for evidential reasoning in Section 2. To set out the future, organisations can prepare themselves for upcoming legal disclosure with reliable digital evidence, such as kernel traces, the low-level execution logs of operating systems (Giraldeau, Desfossez, Goulet, Dagenais, & Desnoyers, 2011). To conduct evidential reasoning of kernel traces for forensic readiness, we propose a framework by integrating the narrative methods from attack graphs/languages, the argumentative methods from argumentation diagrams, and the probabilistic methods from Bayesian networks, which is shown in Figure 5.

3.1 Related Work

Distinguished from previous studies on method integration for evidential reasoning and interpretation, which are summarized in Table 4, we employ the narrative methods from attack graphs and attack description languages to emphasise the investigation of information security incidents in this paper.

Table 4. A Summary of Method Integration for Evidential Reasoning

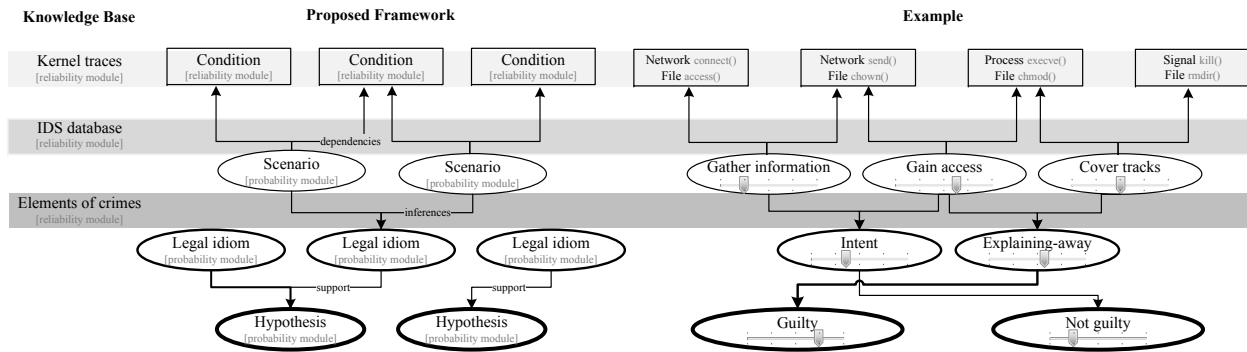| Author(s)	Methods		
	Narrative	Argumentative	Probabilistic
Hepler et al.		x	x
F. J. Bex et al.	x	x	
Condliffe et al.		x	x
Vlek et al.	x		x
Verheij	x	x	x
Timmer et al.		x	x

Figure 5. The Framework for Evidential Reasoning of Kernel Traces

3.2   Producing Scenarios through Attack Graphs/Languages

Producing scenarios automatically can reduce the possibility of bias and errors. Once a suspect has been targeted, the evidence gathering process tends to become restricted to supporting the guilt of the suspect. To keep investigators from the confirmation bias, we suggest to employ the essential elements of attack description languages (Cuppens & Ortalo, 2000; Templeton & Levitt, 2000; Cheung, Lindqvist, & Fong, 2003): pre-conditions (require capabilities) and post-conditions (provide capabilities) as datums, and scenarios as claims, to produce scenarios automatically from kernel traces.

Attack graphs aim to represent the potential intrusion paths for security vulnerability analysis, whereas attack description languages usually address the generic description issues for knowledge interchange. Despite the differences, both of them are effective narrative methods to demonstrate information security incidents. Table 5 summarizes the elements to present information security incidents in a narrative form. The pre-condition is the set of resources required for an attack to occur (e.g. the host vulnerability information and the communication connectivity), the scenario is the set of

events of an attack, and the post-condition is the set of resources obtained by a successful attack.

Table 5. A Summary of Attack Graphs and Languages

Author(s)	Graph/Language Elements		
	Nodes	Edges	Edge Weight
Phillips and Swiler	stages of attack	actions, conditions	success probability
Ortalo et al.	privileges	vulnerabilities	efforts required
Cuppens and Ortalo	pre-condition, post-condition, scenario, detection, verification		
Ritchey and Ammann	hosts, connectivity matrix, exploits		
Templeton and Levitt	require capabilities, provide capabilities, actions		
Eckmann et al.	states, transitions		
Michel and M	exploit, detection, response		
Cheung et al.	pre-condition, post-condition		
Noel et al.	exploits, conditions	dependencies	hardening costs
Ou et al.	attack goal, derivation, fact	dependencies	n/a

The scenario producing process requires the knowledge similar to the database of intrusion detection systems, which holds the security-related events correlated based on different system resources (e.g. file system, process management, and network). Using system resources for correlation not only increases the possibility of integrating kernel traces from other operating systems, but also enhances the comprehensibility of attack stages. Since the post-condition of a scenario can be the pre-condition of the following scenario, it is effortless to understand the intent of acquiring specific system resources.

For instance, if the database of security-related events defines the scenario of gathering information as pre-conditions of initiating a connection on a socket and check-

ing user's permissions for a file, and post-conditions of sending a message on a socket and changing ownership of a file, we can produce the scenario of gathering information automatically from gathered kernel traces. Similarly, we can produce the scenario of covering tracks with pre-conditions of executing program and changing permissions of a file, and post-conditions of sending signal to a process and deleting a directory.

3.3 Generating Hypotheses through Argumentation Diagrams

Convicting a defendant demands constructing legal arguments. To assist investigators to construct legal arguments, we suggest to employ the concepts of datum and claim from the Toulmin model of argument (Toulmin, 2003) and legal idioms, such as evidence-accuracy idioms and intent idioms (Fenton, Neil, & Lagnado, 2013), for hypothesis generation.

Argumentation diagrams aim to represent the structure of an argument visually, which can facilitate the crime scenario construction, crime investigation, and decision support from a legal perspective. However, argument diagramming requires facts, arguments, and evidence as input data, which mainly depends on manual user input. Table 6 summarizes the elements to represent information security incidents in an argumentative form.

Table 6. A Summary of Argumentation Diagrams

| Author(s) | Graph Elements | | |
	Nodes	Edges	Edge Weight
Goodwin	facts	probative processes	belief strength
Toulmin	datum, warrant, claim, backing, rebuttal	certainty	n/a
Verheij	statements	support/attack	n/a
Reed and Rowe	reconstructed enthymemes, refutations	support	n/a
Chesevar et al.	claims, applications of schemes	support	n/a
Keppens and Schafer	facts, evidence, assumptions, hypothesis	causal relations	n/a
F. Bex et al.	data, inference, scheme	inferential relations	n/a
Gordon	statements, testimonial evidence	inferential relations	argument strength

An automatic judgment summarizing system, which we have developed in Section 2,

can minimize the manual effort required for hypothesis generation and provide knowledge for developing legal idioms from scenarios. Examining the evidence used against a defendant from previous judgments constantly can also solidify legal cases and avoid repeating past mistakes. The decision element corresponds to hypotheses, the intent and fact elements fit into legal idioms, and the scenarios can be considered as evidence in narrative forms. The legal idioms developed can establish attack or support relations with the hypotheses generated, which are normally mutually exclusive. For example, Figure 5 consists of two mutually exclusive hypotheses: the intruder is guilty or not guilty.

3.4 Assessing Probabilities through Bayesian Networks

From a legal perspective, Bayesian networks can facilitate the interpretation of complicated relations between evidence, the modelling of legal arguments and crime scenarios, and the process of evidential reasoning. To assist investigators to discover of the most supported hypotheses, we suggest to employ the qualitative probability (Keppens, 2007) from Bayesian networks for probability assessment.

Bayesian networks aim to present variables and their probabilistic relations in a directed acyclic graph. Table 7 summarizes the elements to present information security incidents in a probabilistic form. Even though Bayesian networks can present variables and their probabilistic relations in accurate numbers, these numerical probabilities demand huge amount of knowledge as input for calculation.

The gathered evidence only become admissible when it meets the criteria of relevance and reliability. Since the developed

Table 7. A Summary of Bayesian Networks Regarding Legal Reasoning

Author(s)	Graph Elements		
	Nodes	Edges	Edge Weight
Zukerman et al.	propositions	inferences	n/a
Hepler et al.	modules, variables	inferred hierarchy	n/a
Keppens	variables	influences	qualitative derivatives
F. J. Bex et al.	observations, arguments, evidence	inferences	n/a
Condliffe et al.	evidence variables	inferred hierarchy	n/a
Keppens	variables	inference	probative force
Fenton et al.	basic causal structures	causal inferences	probability
Vlek et al.	scenarios	dependencies	n/a
Verheij	hypotheses, evidence	inferences	strengths
Timmer et al.	support factors	support	likelihood ratio

legal idioms attack or support the generated hypotheses, the datums and the claims connected to the legal idioms should be considered relevant. As for reliability, since a datum is a fact or an observation, and a claim is a consequent assertion that depends on datums (Shum, 2003), based on the object oriented concept (Hepler et al., 2007), we suggest to assign a reliability module and a qualitative probability module to each datum and claim respectively.

A reliability module should contain a conditional probability table from Bayesian networks, which can be derived from potential errors identified and vulnerability assessment outcomes. For instance, we can use the detection rate and false alarm rate as input for the reliability module of the intrusion detection system database required for producing scenarios. Thus, we can make errors present and transparent for a fair trial.

Since the comparison between different hypotheses does not require precise numerical probabilities, the qualitative approach is enough in the context of forensic analysis (Keppens, 2007). We suggest employing the analysis of inadmissibility grounds as input for calculation, assigning a qualitative probability module to a claim, and using different levels of edge thickness to indicate different levels of probability. For instance, we can utilize sliders to present scales of probability, which are shown in Figure 5.

4. CONCLUSIONS AND FUTURE WORK

Inspired by the feasibility of extracting legal knowledge for argument construction and employing grounds of inadmissibility for probability assessment, we integrate the narrative methods from attack graphs/languages, the argumentative methods from argumentation diagrams, and the probabilistic methods from Bayesian networks, to conduct evidential reasoning of kernel traces for forensic readiness.

Employing the elements of pre-condition and post-condition from attack description languages, we can keep investigators from confirmation bias, clarify the stages of attacks, and prevent miscarriages of justice. Following the concepts of datum and claim from argumentation diagrams, we can assist investigators to construct legal arguments. Moreover, the automatic judgement summarizing system developed can help investigators to solidify legal cases and avoid repeating past mistakes. Utilizing the ideas of legal idioms and qualitative probability from Bayesian networks, we can assist investigators to discover of the most supported hypotheses and assess the admissibility of gathered evidence.

In addition to broadening the judgment analysis scope for conducting comprehensive cross-jurisdictional comparisons, we plan to conduct a case study for evaluating the effectiveness of the proposed framework, and broaden the element scope for deeper legal knowledge extraction, such as 5W1H information. We also need to ensure the privacy implications for users are minimized to acceptable levels to alleviate the conflicts between accountability and privacy. We will make the raw data and source code available to interested researchers under an appropriate open source licenses.

ACKNOWLEDGEMENT

REFERENCES

Bex, F., Van den Braak, S., Van Oostendorp, H., Prakken, H., Verheij, B., & Vreeswijk, G. (2007). Sense-making software for crime investigation: how to combine stories and arguments? *Law, Probability and Risk*, *6*(1-4), 145–168. Retrieved 2016-02-28, from `http://lpr.oxfordjournals.org/content/6/1-4/145.short`

Bex, F. J., Koppen, P. J. v., Prakken, H., & Verheij, B. (2010, July). A hybrid formal theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law*, *18*(2), 123–152. doi: 10.1007/s10506-010-9092-x

Brenner, S. W., Carrier, B., & Henninger, J. (2004). The Trojan Horse Defense in Cybercrime Cases. *Santa Clara Computer and High Technology Law Journal*, *21*, 1–1.

Cheng, L. (2010). A semiotic interpretation of genre: Judgments as an example. *Semiotica*, *2010*(182), 89–113.

Chesevar, C., Modgil, S., Rahwan, I., Reed, C., Simari, G., South, M., ... Willmott, S. (2006). Towards an argument interchange format. *The Knowledge Engineering Review*, *21*(04), 293–316.

Cheung, S., Lindqvist, U., & Fong, M. W. (2003, April). Modeling multistep cyber attacks for scenario recognition. In *DARPA Information Survivability Conference and Exposition, 2003. Proceedings* (Vol. 1, pp. 284–292 vol.1). doi: 10.1109/DISCEX.2003.1194892

Chieze, E., Farzindar, A., & Lapalme, G. (2010). An Automatic System for Summarization and Information Extraction of Legal Information. In E. Francesconi, S. Montemagni, W. Peters, & D. Tiscornia (Eds.), *Semantic Processing of Legal Texts* (pp. 216–234). Springer Berlin Heidelberg. (DOI: 10.1007/978-3-642-12837-0_12)

Condliffe, P., Abrahams, B., & Zeleznikow, J. (2010). An OWL Ontology and Bayesian Network to Support Legal Reasoning in the Owners Corporation Domain. In *ODR* (pp. 51–62). Retrieved 2016-02-28, from `http://ceur-ws.org/Vol-684/paper5.pdf?1323a5d8`

Council of Europe. (2008, March). *Cybercrime legislation-country profile: People's Republic of China* (Tech. Rep.). Retrieved 2015-09-18, from `https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016803042ef`

Cunningham, H. (2002, May). GATE, a General Architecture for Text Engineering. *Computers and the Humanities*, *36*(2), 223–254. doi: 10.1023/A:1014348124664

Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V., Ursu, C., Dimitrov, M., ... others (2014). *Developing Language Processing Components with GATE Version 8*. University of Sheffield Department of Computer Science. Retrieved 2015-10-04, from `https://gate.ac.uk/sale/tao/tao.pdf`

Cuppens, F., & Ortalo, R. (2000, October). LAMBDA: A Language to Model a Database for Detection of Attacks. In H. Debar, L. M, & S. F. Wu (Eds.), *Recent Advances in Intrusion*

Detection (pp. 197–216). Springer Berlin Heidelberg. (DOI: 10.1007/3-540-39945-3_13)

Eckmann, S. T., Vigna, G., & Kemmerer, R. A. (2002). STATL: An attack language for state-based intrusion detection. *Journal of computer security*, *10*(1, 2), 71–103. Retrieved 2016-02-28, from `http://content.iospress.com/ articles/journal-of-computer -security/jcs158`

Ellson, J., Gansner, E., Koutsofios, L., North, S. C., & Woodhull, G. (2001, September). Graphviz Open Source Graph Drawing Tools. In P. Mutzel, M. Jnger, & S. Leipert (Eds.), *Graph Drawing* (pp. 483–484). Springer Berlin Heidelberg. (DOI: 10.1007/3-540-45848-4_57)

Farzindar, A., & Lapalme, G. (2004). Letsum, an automatic legal text summarizing system. *Legal knowledge and information systems, JURIX*, 11–18.

Fenton, N., Neil, M., & Lagnado, D. A. (2013). A general structure for legal arguments about evidence using Bayesian networks. *Cognitive science*, *37*(1), 61–102. Retrieved 2016-02-28, from `http://onlinelibrary.wiley.com/ doi/10.1111/cogs.12004/full`

Galgani, F., Compton, P., & Hoffmann, A. (2012, March). Towards Automatic Generation of Catchphrases for Legal Case Reports. In A. Gelbukh (Ed.), *Computational Linguistics and Intelligent Text Processing* (pp. 414–425). Springer Berlin Heidelberg. (DOI: 10.1007/978-3-642-28601-8_35)

Gelbart, D., & Smith, J. C. (1993). FLEXICON: An Evaluation of a Statistical Ranking Model Adapted to Intelligent Legal Text Management.

In *Proceedings of the 4th International Conference on Artificial Intelligence and Law* (pp. 142–151). New York, NY, USA: ACM. doi: 10.1145/158976.158994

Giraldeau, F., Desfossez, J., Goulet, D., Dagenais, M., & Desnoyers, M. (2011). Recovering system metrics from kernel trace. In *Linux Symposium* (Vol. 109). Retrieved 2016-02-28, from `http://landley.net/kdocs/ mirror/ols2011.pdf#page=109`

Goodwin, J. (2000). Wigmore's Chart Method. *Informal Logic*, *20*(3).

Gordon, T. F. (2007). Visualizing Carneades argument graphs. *Law, Probability and Risk*, *6*(1-4), 109–117. Retrieved 2016-02-28, from `http://lpr.oxfordjournals.org/ content/6/1-4/109.short`

Graves, K. (2007). *CEH: Official Certified Ethical Hacker Review Guide* (1st ed.). Sybex.

Hachey, B., & Grover, C. (2007, March). Extractive summarisation of legal texts. *Artificial Intelligence and Law*, *14*(4), 305–345. doi: 10.1007/s10506-007-9039-z

Hepler, A. B., Dawid, A. P., & Leucari, V. (2007). Object-oriented graphical representations of complex patterns of evidence. *Law, Probability and Risk*, *6*(1-4), 275–293. Retrieved 2016-02-28, from `http://lpr.oxfordjournals.org/ content/6/1-4/275.short`

International Organization for Standardization. (2015, March). *ISO/IEC 30121:2015 - Information technology – Governance of digital forensic risk framework* (Tech. Rep.).

Keppens, J. (2007). Towards Qualitative Approaches to Bayesian Evidential Reasoning. In *Proceedings of the 11th*

International Conference on Artificial Intelligence and Law (pp. 17–25). New York, NY, USA: ACM. doi: 10.1145/1276318.1276322

Keppens, J. (2012, March). Argument diagram extraction from evidential Bayesian networks. *Artificial Intelligence and Law*, *20*(2), 109–143. doi: 10.1007/s10506-012-9121-z

Keppens, J., & Schafer, B. (2006, February). Knowledge based crime scenario modelling. *Expert Systems with Applications*, *30*(2), 203–222. doi: 10.1016/j.eswa.2005.07.011

Mailhot, L., & Carnwath, J. D. (1998). *Decisions, Decisions: A Handbook for Judicial Writing.* Cowansville, Qubec: ditions Y. Blais.

Michel, C., & M, L. (2002). ADeLe: An Attack Description Language for Knowledge-Based Intrusion Detection. In M. Dupuy & P. Paradinas (Eds.), *Trusted Information* (pp. 353–368). Springer US. (DOI: 10.1007/0-306-46998-7_25)

Noel, S., Jajodia, S., O'Berry, B., & Jacobs, M. (2003, December). Efficient minimum-cost network hardening via exploit dependency graphs. In *Computer Security Applications Conference, 2003. Proceedings. 19th Annual* (pp. 86–95). doi: 10.1109/CSAC.2003.1254313

Ortalo, R., Deswarte, Y., & Kaaniche, M. (1999, September). Experimenting with quantitative evaluation tools for monitoring operational security. *IEEE Transactions on Software Engineering*, *25*(5), 633–650. doi: 10.1109/32.815323

Ou, X., Boyer, W. F., & McQueen, M. A. (2006). A Scalable Approach to Attack Graph Generation. In *Proceedings of the 13th ACM Conference on Computer and Communications Security* (pp. 336–345). New York, NY, USA: ACM. doi: 10.1145/1180405.1180446

Phillips, C., & Swiler, L. P. (1998). A Graph-based System for Network-vulnerability Analysis. In *Proceedings of the 1998 Workshop on New Security Paradigms* (pp. 71–79). New York, NY, USA: ACM. doi: 10.1145/310889.310919

Reed, C., & Rowe, G. (2004). Araucaria: Software for argument analysis, diagramming and representation. *International Journal on Artificial Intelligence Tools*, *13*(04), 961–979. Retrieved 2016-02-28, from `http://www.worldscientific.com/doi/abs/10.1142/S0218213004001922`

Ritchey, R. W., & Ammann, P. (2000). Using model checking to analyze network vulnerabilities. In *2000 IEEE Symposium on Security and Privacy, 2000. S P 2000. Proceedings* (pp. 156–165). doi: 10.1109/SECPRI.2000.848453

Rowlingson, R. (2004). A ten step process for forensic readiness. *International Journal of Digital Evidence*, *2*(3), 1–28.

Saravanan, M., Ravindran, B., & Raman, S. (2008). Automatic Identification of Rhetorical Roles using Conditional Random Fields for Legal Document Summarization. In *Third International Joint Conference on Natural Language Processing* (p. 481).

Schweighofer, E., Winiwarter, W., & Merkl, D. (1995). Information Filtering: The Computation of Similarities in Large Corpora of Legal Texts. In *Proceedings of the 5th International Conference on Artificial Intelligence and Law* (pp. 119–126). New York, NY, USA: ACM. doi:

10.1145/222092.222205

Scrapinghub, Ltd. (2015, June). *Scrapy.* Retrieved from `http://scrapy.org`

Shum, S. B. (2003). The Roots of Computer Supported Argument Visualization. In *Visualizing Argumentation* (pp. 3–24). Springer London. (DOI: 10.1007/978-1-4471-0037-9_1)

Templeton, S. J., & Levitt, K. (2000). A Requires/Provides Model for Computer Attacks. In *Proceedings of the 2000 Workshop on New Security Paradigms* (pp. 31–38). New York, NY, USA: ACM. doi: 10.1145/366173.366187

Timmer, S. T., Meyer, J.-J. C., Prakken, H., Renooij, S., & Verheij, B. (2015). A Structure-guided Approach to Capturing Bayesian Reasoning About Legal Evidence in Argumentation. In *Proceedings of the 15th International Conference on Artificial Intelligence and Law* (pp. 109–118). New York, NY, USA: ACM. doi: 10.1145/2746090.2746093

Torpey, E. M. (2009). Careers in Forensics: Analysis, Evidence, and Law. *Occupational Outlook Quarterly*, *53*(1), 14–19. Retrieved 2016-02-28, from `http://eric.ed.gov/?id=EJ875430`

Toulmin, S. E. (2003). *The uses of argument.* Cambridge University Press.

Uyttendaele, C., Moens, M.-F., & Dumortier, J. (1998, March). Salomon: Automatic Abstracting of Legal Cases for Effective Access to Court Decisions. *Artificial Intelligence and Law*, *6*(1), 59–79. doi: 10.1023/A:1008256030548

Verheij, B. (2003, November). Artificial argument assistants for defeasible argumentation. *Artificial Intelligence*, *150*(12), 291–324. doi: 10.1016/S0004-3702(03)00107-3

Verheij, B. (2014). To catch a thief with and without numbers: arguments, scenarios and probabilities in evidential reasoning. *Law, Probability and Risk*, *13*(3-4), 307–325. Retrieved 2016-02-28, from `http://lpr.oxfordjournals.org/content/13/3-4/307.short`

Vlek, C., Prakken, H., Renooij, S., & Verheij, B. (2013). Modeling Crime Scenarios in a Bayesian Network. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law* (pp. 150–159). New York, NY, USA: ACM. doi: 10.1145/2514601.2514618

Wyner, A. Z. (2010). Towards annotating and extracting textual legal case elements. *Informatica e Diritto: special issue on legal ontologies and artificial intelligent techniques*, *19*(1-2), 9–18.

Yousfi-Monod, M., Farzindar, A., & Lapalme, G. (2010, May). Supervised Machine Learning for Summarizing Legal Documents. In A. Farzindar & V. Keelj (Eds.), *Advances in Artificial Intelligence* (pp. 51–62). Springer Berlin Heidelberg. (DOI: 10.1007/978-3-642-13059-5_8)

Zukerman, I., McConachy, R., & Korb, K. B. (1998). Bayesian reasoning in an abductive mechanism for argument generation and analysis. In *AAAI/IAAI* (pp. 833–838). Retrieved 2016-02-28, from `http://www.aaai.org/Papers/AAAI/1998/AAAI98-118.pdf`