Identifying Vulnerabilities and Critical Requirements Using Criminal Court Proceedings

Travis D. Breaux[1], Jonathan D. Lewis[1], Paul N. Otto[1,2], Annie I. Antón[1]

Department of Computer Science, North Carolina State University[1]
School of Law, Duke University[2]
{tdbreaux,jdlewis,pnotto,anton}@ncsu.edu

Abstract

Information systems governed by laws and regulations are subject to both civil and criminal violations. In the United States, these violations are documented in court records, such as complaints, indictments, plea agreements, and verdicts, which constitute a source of real-world software vulnerabilities. This paper reports on an exploratory case study to identify legal vulnerabilities and provides guidance to practitioners in the analysis of court documents. As legal violations occur after system deployment, court records reveal vulnerabilities that were likely overlooked during software development. We evaluate the effectiveness of established requirements engineering techniques, including sequence and misuse case diagrams and goal models, as applied to criminal court records to identify mitigating requirements. In a sustainable world, these techniques, when properly applied, can help organizations focus their risk-management efforts on emerging vulnerabilities. We illustrate our analysis using criminal indictments involving two separate systems governed by the U.S., Health Insurance Portability and Accountability Act (HIPAA) and U.S. Sarbanes-Oxley Act (SOX).

1. Introduction

The requirements of information systems are increasingly being affected by United States government laws and regulations, such as the Health Insurance Portability and Accountability Act of 1996[1] (HIPAA), which governs the privacy of electronic patient medical records, and the Sarbanes-Oxley Act of 2002[2] (SOX), which governs corporate accounting. In addition to new laws that govern existing practices, businesses that innovate by introducing new products and services subject themselves to new interpretations of existing laws that have civil and criminal consequences. For example, a recent U.S. administrative investigation involving ChoicePoint [9], an information broker, resulted in a $10 million civil penalty and $5 million in consumer redress [8] and illustrates how information products can be designed without sufficient consideration for the impact of existing laws, in this case, the Fair Credit Reporting Act[3] (FCRA) [17, 27]. Similar cases have drawn the attention of chief privacy and security officers (CPOs, CSOs) in organizations worldwide. The 10th annual 2007 Ernst & Young survey of nearly 1,300 organizations revealed that the top two drivers of information security practice are compliance with regulations (63%) and privacy and data protection (58%) [7].

Compliance frameworks and standards such as COBIT and ISO 17799 provide organizations broad guidance in how to plan, implement, and monitor security controls in information systems. However, these frameworks and standards do not address specific issues in laws and regulations that can leave organizations exposed to civil and criminal penalties. One approach to help organizations focus their compliance efforts is to identify specific vulnerabilities and the necessary requirements to address them. We distinguish between *preventative requirements*, which prevent or reduce the possibility of vulnerabilities, and *mitigating requirements*, which seek to reduce the effect of vulnerabilities after they are exploited. In this paper, we propose the use of civil and criminal court documents as an information source for identifying vulnerabilities and corresponding mitigating requirements. These requirements are *critical requirements* because they are based on established vulnerabilities documented in civil and criminal cases and resulting in civil and criminal penalties, including regular audits, fines, and incarceration.

[1] Pub. L. No. 104-191, 110 Stat. 1936
[2] Pub. L. No. 107-204, 116 Stat. 745

[3] 15 U.S.C. §§ 1681–81x (2000)

Whereas the ChoicePoint case illustrates violations of law due to outsider threats, an equally important but less-studied problem involves insider threats in which an employee or contractor engages in activities that increase civil or criminal liability. In two successive years, CSO Magazine reported on the annual E-crime Watch Survey, showing that over one-third of surveyed security executives and law enforcement officials identified insider threats as the greatest cause of damage to information systems [10, 38]. Herein, we illustrate how to identify critical requirements by highlighting two criminal cases that describe insider threats: United States v. Ferrer, involving a HIPAA violation [12]; and United States v. Fumo, involving a SOX violation [11].

The contributions in this paper are: (1) background on the U.S. federal legal process and terminology that requirements engineers must know to replicate this analysis; (2) our experiences acquiring relevant court documents needed to perform this analysis and guidance for how future researchers can save time during this effort; and (3) a comparative evaluation of three notations that we use to analyze legal vulnerabilities with examples that link vulnerabilities to mitigating requirements.

The remainder of the paper is organized as follows: we review related work in Section 2; present background including relevant legal terminology, procedures and documents from the U.S. federal court system in Section 3; in Section 4, we present the methods we used to conduct this study and a description of the documents that we analyzed; in Section 5, we present our analysis results, comparing the different notations; in Section 6, we present an example that links legal vulnerabilities to mitigating requirements; and finally, we discuss our observations and conclude with future work in Section 7.

2. Related Work

Misuse and abuse cases are used in security to elicit and reason about system vulnerabilities. McDermott and Fox first defined abuse cases as user-system interactions that result in harm to the system [25]. Engineers use the abuse case notation to elicit these interactions from customers and document them. A related concept to the abuse case is the misuse case introduced by Sindre and Opdahl [34]. An important difference from abuse cases is that misuse cases provide links to security use cases that are intended to mitigate the case for misuse. Misuse cases have been popular in engineering practice as important artifacts to be used early in the software development process [18, 30], for performing trade-off analysis [1] and risk analysis [36], and to analyze business processes [31].

In requirements engineering, goals describe intended states to be maintained or achieved by the system [6]. Conflicts or obstacles are relationships between goals that lead to inconsistency in software specifications [22]. Obstacles include hazards to safety goals or threats to security goals [21]. Similar to misuse cases, obstacles introduced by a threatening agent are called *anti-goals* [23]. Anti-goals are resolved by creating new goals to mitigate or prevent the obstructing goals [21, 23]. In this paper, we contrast the use of misuse case diagrams with KAOS diagrams for representing legal vulnerabilities using anti-goals.

Regnell et al. propose using hierarchical use case models to iteratively decompose goals into sequential user actions using use case, flow and sequence diagrams [32]. In Section 5, we show that certain court documents can lack sufficient detail to create accurate sequence diagrams but can provide enough information to create motivating misuse case and KAOS diagrams that aid in identifying mitigating requirements.

In recent years, researchers have been drawn to the challenges that legal requirements pose to information systems [28]. These challenges include accurately acquiring legal requirements [4, 5], maintaining traceability [4, 15], checking consistency [26], and realizing legal requirements in business practices [19]. To the authors' knowledge, this paper represents the first time that vulnerabilities or software requirements have been identified from criminal and civil case law.

3. Legal Background

To help the reader understand the context surrounding criminal proceedings, we provide a cursory overview of the process governing United States criminal law. A few simplifying assumptions have been made for clarity and brevity. The discussion focuses on the federal criminal court system, which is governed by the Federal Rules of Criminal Procedure (FCRP)[4] [13]; various states may have different rules at each stage of the process. We describe how the federal system handles non-capital felonies, glossing over differences in misdemeanor proceedings. Many white-collar crimes are initially investigated by administrative agencies, rather than police; the discussion briefly notes how the process differs in such cases. Our discussion emphasizes phases involving the specific court documents analyzed for this paper.

Once a violation is suspected, the investigative process begins: enforcement officials (e.g., police, prosecutors) determine whether a crime was committed, identify the perpetrators, gather evidence linking the perpetrators to the crime, and locate the

[4] Abbreviated as "Fed. R. Crim. P." in legal works.

perpetrators [20]. Once probable cause is established, police typically arrest and book the suspected perpetrators. After the arrest is made, further investigation typically takes place, including interviewing witnesses and subpoenaing documents as required by the situation.

If substantial evidence exists linking the suspects with the crime, official charges will be filed against the suspects within days of the initial arrest. A *criminal complaint* – a formal document accusing a suspect of committing some criminal act – is typically filed at this stage of the process [14]. The complaint, governed by FCRP 4 [13], tends to be a brief assessment of the specific acts performed by the accused that constituted a criminal statute's violation [20]. At this stage in the process, the accused becomes a defendant and the judicial process begins. The court first creates an official *docket*, a record of all proceedings and filings involved in the case [14]. With the passage of the E-Government Act of 2002[5], all federal courts must provide online access to court information, including full case dockets.

In many cases, the next step in the process is to convene a grand jury, a function specified by FCRP 6 [13]. A grand jury determines whether there is sufficient evidence against the defendant, based solely on the evidence proffered by the prosecution, to justify advancing to trial [20]. If the grand jury finds the evidence sufficient, it can issue an *indictment*, or "formal written accusation" of criminal conduct [14]. The indictment, specified by FCRP 7(c), "shall be a plain, concise and definite statement of the essential facts constituting the offense charged" [13]. The indictment must reference the specific law or regulation violated by the defendant, as well as substantiating the existence of facts for each element of the crime. Note that an indictment may contain several counts (distinct criminal charges), or also may charge that one act violated multiple criminal statutes. The indictment thus provides the first full account of the suspected violation(s) leading to criminal prosecution. In some cases, due to the sensitive nature of the criminal act or other considerations, an indictment may be released either in a redacted form or sealed permanently from public viewing. The indictment becomes the official accusatory document against the defendant, superseding the complaint [20]. In some cases, superseding indictments will be filed; such amended indictments replace previous indictments on record.

For crimes occurring in highly-regulated industries, the steps leading up to the grand jury may be entirely performed by an administrative agency. In the context of investigating an organization for civil violations, an administrative agency may discover that criminal acts may have occurred. Such agencies can refer the case to a prosecutor to determine whether criminal charges are warranted. An investigative grand jury may then be convened, bypassing the police investigation and criminal complaint.

After an indictment is issued by a grand jury, the defendant faces the first opportunity to enter a plea at arraignment. *Arraignment* entails bringing the defendant, who is asked to enter a *plea* of either guilty, not guilty, or nolo contendere (no contest to the charges, with the same sentencing as a guilty verdict), before the trial court [14]. After arraignment, *plea bargaining* may begin in which a substantial majority of defendants will exchange a guilty plea for reduced charges or lesser sentencing [20]; if accepted, the offer will be detailed in a *plea agreement*. The trial judge must approve any plea agreement before the case is closed; the judge ensures that the defendant has voluntarily entered into the plea agreement and understands all its terms [20]. In 2006, over 95% of criminal charges resulted in a guilty plea before reaching trial [35].

Defendants have a right to a *jury trial* in all felony prosecutions; a trial by judge is called a *bench trial* [20]. Around 70% of defendants who do not plead guilty in the U.S. elect to exercise their jury right [20]. At either trial, defendants enjoy several key rights, including the presumption of innocence, right to avoid self-incrimination, and requirement for the prosecution to establish guilt beyond a reasonable doubt. If a guilty verdict is entered, whether by judge or jury, a judge generally determines the defendant's sentence. Probation officers will provide a *presentence investigation report* for sentencing purposes, which details the "convicted defendant's educational, criminal, family, and social background" [14]. In some cases, the statute or regulation establishing the crime will also dictate the appropriate sentence. There are three broad categories of sanctions: restitution, probation, or incarceration [20].

4. Legal Case Selection and Description

This research employed an exploratory, multiple case study design [39] to answer a two-part research question: can we identify software vulnerabilities from civil and criminal cases and, if so, which notation best represents the information contained in relevant case materials? In this section, we describe the materials that we used to conduct the case study, how we identified and purposefully selected these materials, the units of analysis and our analysis procedure that was used to obtain our findings.

[5] Pub. L. No. 107-347, § 205, 116 Stat. 2899, 2913-15

4.1. Identifying Relevant Cases

Relevant civil and criminal cases and corresponding court documents can be identified and acquired in different ways. In the United States, the federal government centrally manages federal court records through the Public Access to Court Electronic Records (PACER) database system. In addition, privately managed databases, such as LexisNexis and WestLaw are available. These databases charge a subscription-based or per-page fee to retrieve court documents. The per-page fee includes the number of pages for each requested document, in addition to each page that appears in search results leading up to the document request. The search costs to identify relevant cases by topic or keyword, as opposed to looking up unique case numbers, may be prohibitive for businesses or engineers with a small discretionary budget.

An alternative lower-cost, indirect method to identify cases is through news reports and press releases. During this study, we employed the relevant sections of the United States Code (U.S.C.) for HIPAA and SOX to identify cases in news reports that correspond to violations of specific laws, then using the case number to identify information on specific cases. In addition, the regional offices of the U.S. Department of Justice will often post online press releases announcing indictments and convictions. In rare situations, these offices will also post complaints and indictments online from specific cases or make these documents available to individuals upon request through electronic mail. Based upon our experience, however, this is not standard practice.

Although PACER and private databases provide the most comprehensive source of court records, court transcripts may not be included in these databases. Unlike other court documents, transcripts are recorded and prepared by stenographers, known as court reporters, who record the verbal communication of judges, attorneys, witnesses, and other parties during the trial by using a shorthand form of writing. The shorthand notation utilizes symbols corresponding to spoken phonemes (discrete speech sounds) rather than letters. The shorthand documents are transcribed into standard English upon request, at which time the requesting agent may pay the per-page cost of transcription.

For this study, we chose to examine recent civil and criminal cases that include at least one violation of HIPAA regulatory law, because of our experience in analyzing HIPAA regulations [4, 5]. In addition, we examined cases that include one violation of SOX regulatory law for contrast and to identify any domain effects separately due to information privacy and corporate accounting law. We acquired the full dockets for the following eight cases using PACER:

1. United States v. Gibson – a hospital insider acquires patient medical records to commit wire fraud.
2. United States v. Ferrer – an insider acquires patient medical records to commit Medicare fraud.
3. United States v. Hungerford – a health insurance insider acquires patient medical records to commit wire fraud.
4. United States v. Occident– a hospital insider acquires patient medical records to commit wire fraud.
5. United States v. Ramirez – a primary care provider insider attempts to sell a patient medical record to a drug trafficker.
6. United States v. Williams – a healthcare clearinghouse insider acquires and sells patient medical records.
7. United States v. Williams and Adjei – a healthcare clearinghouse insider acquires patient medical records to file fraudulent tax returns.
8. United States v. Fumo – insiders destroy documents to obstruct a federal investigation.

Cases 1-7 were reported as the only seven HIPAA-related cases to date by an Assistant U.S. District Attorney for the Western District of Washington who summarized recent HIPAA-related cases [37]. Case 8 was selected because it is a high-profile SOX-related case. Each case has a full docket, which is a list of all relevant court documents for the case.

In all but one case (Case 7), the defendants were indicted together, although in all eight cases each defendant was tried separately. There is a distinct docket for each defendant, resulting in 22 dockets in all. Reviewing each docket, there are a total of 1141 entries; Table 1 presents a subset of the 238 different types of entries we identified from these dockets. Among these eight cases, only four originated with official complaints; the remaining four cases originated through other means, such as administrative hearings.

United States v. Fumo has been continued until September 8th, 2008. Of the remaining 18 defendants represented by these dockets, only two did not agree to a plea agreement with the prosecuting district attorneys. A bench trial was held for defendant Occident (Case 4) in which she was found guilty on multiple charges. A jury trial was held for defendant Ferrer (Case 2) in which he was also found guilty. Thus only two trial transcripts could be acquired as of this paper. There is no indictment for Case 1, due to a condition in Gibson's plea agreement.

Type of Docket Entry	No. of Entries
Complaint	8
Indictment	38
Plea Agreement	16
Transcript	25
Minute Entry	154
Judgment	26

Table 1: Types of available documents for the eight cases examined in this study

4.2. Selecting Relevant Court Documents

As our analysis focuses on identifying legal vulnerabilities that affect software, we selected documents that describe actors and events, as well as how software systems might directly or indirectly be used to commit the crimes as charged.

Complaints and indictments include a general account of events as charged. Because indictments supersede complaints as the official account of criminal violations [20], we included indictments in our detailed analysis. On the other hand and based upon our observations, we found that plea agreements contain no more detail than the corresponding indictments and, in fact, contain additional information irrelevant to this analysis (e.g., waivers of specific rights, penalties imposed).

While many transcripts were sealed or not accessible for various reasons, we did acquire one sentencing transcript, which provided insight into the seriousness of the crime. In United States v. Gibson, Gibson admits to using the identity of a terminal cancer patient to commit wire fraud. The sentencing transcript in this case describes the judge's reaction to Gibson's behavior as "the most deplorable" by stating that "it's true that [Gibson] didn't murder anyone, [he] didn't physically assault anyone; but in a very real sense [he] committed a vicious attack on someone who was fighting for his life at the most vulnerable point in his life." The judge decided this detail was substantial and cause for imposing the harshest sentence. As a result of this additional emphasis, security analysts might consider taking additional steps to protect records of patients who are more vulnerable due to the severity of their illnesses. Because of the cost and time required to obtain transcripts, we did not obtain a complete record of transcripts for all eight cases in time for this paper.

Other documents that did not contain substantial information relevant to this study include minute entries and judgments. Minute entries provided many notes regarding meetings and hearings. The information would be important to those involved in the legal process, but minute entries do not provide any information regarding the facts surrounding the alleged crimes. The judgments that we analyzed were template-based forms with checkboxes and small write-in blanks. The information contained in these documents did not contribute substantially to identifying vulnerabilities.

We focused our analysis primarily on the indictments from each case, or secondarily on the plea agreements if an indictment was not available, as was the case in United States v. Gibson. Several different versions of an indictment may exist; this explains the 38 indictment entries in the dockets for only 22 defendants. The various types of indictments are described in Section 3. Lastly, we found that the sentencing transcripts can be used to generally rate the severity of the case.

4.3. Units of Analysis and Procedure

The units of analysis for this study consist of descriptions of actors, actions and events involved in civil and criminal charges. These units were prescribed by our choice of notations: sequence, misuse case and KAOS diagrams. This limited focus supported the goal of our analysis, which was to identify legal vulnerabilities in software systems, but may have caused us to overlook other important features that are also relevant to engineering software systems that comply with the law.

The analysis procedure was performed in two passes over selected court documents by two researchers working in tandem. The first pass consisted of a complete document reading and identification of all relevant actors and events. The actors and events are identified using heuristics from the Goal-Based Requirements Analysis Method [2]. The second pass is limited to parts of the document in which events are identified. During this pass, the analyst reads each event description and integrates the corresponding actor and their actions into the target notation. This integration requires analyzing phrases in each description and utilizes heuristics for handing cross-references [4, 5] and to identify purposes and instruments [3]. The second pass is repeated for each notation. This repetition, as opposed to deriving one diagram from another, avoids bias introduced by the limitations of any one notation. Domain-specific linguistic devices in the text and limitations in the notation are identified and recorded for discussion. Finally, to check consistency and completeness, the actors from the first pass are cross-checked with the actors from the second pass to identify missing events.

5. Scenario and Goal Analysis

We illustrate the results of our analysis using criminal indictments to identify real-world software

vulnerabilities.[6] Our objective was to develop software requirements that will thwart future insider efforts to exploit these vulnerabilities. This analysis entailed deriving sequence, misuse case and KAOS diagrams from the indictments and charges.

We introduce the notations and illustrate the analysis using United States v. Ferrer (Case 2) that describes an insider threat [12]. Consider the following excerpt from the corresponding indictment, paragraph (6) in which actors are *italicized* and events are underlined:

> "6. From on or about May 23, 2005, and continuing through on or about June 26, 2006, at Broward County, in the Southern District of Florida, and elsewhere, the defendants,
>
> *FERNANDO FERRER, JR.*,
> and
> *ISIS MACHADO*,
>
> did knowingly and willfully combine, conspire, confederate, and agree with *each other* and with *others* known and unknown to the Grand Jury, to defraud the *United States* and to commit certain other offenses against the *United States*, namely:
>
> a. to knowingly and with intent to defraud, exceed authorized access to a protected computer, and by such conduct further the intended fraud to obtain things of value exceeding $5,000, in violation of Title 18, United States Code, Sections 1030(a)(4) and (c)(3)(A);
> b. during and in relation to a felony violation of Title 18, United States Code, Chapter 47, to wit, Title 18, United States Code, Sections 1030(a)(4) and (c)(3)(A), to knowingly possess and use, without lawful authority, a means of identification of *another person*, in violation of Title 18, United States Code, Section 1028A(a)(1); and
> c. to knowingly and for a reason other than permitted by Title 42, United States Code, Chapter 7, Subchapter XI, Part C, obtain individually identifiable health information relating to an *individual*, with the intent to sell, transfer, and use, and cause to be used, individually identifiable health information for personal gain, in violation of Title 42, United States Code, Sections 1320d-6(a)(2) and (b)(3)."

This excerpt highlights several key findings that we observed throughout our analysis of all eight cases. First, at the start of paragraph (6), the dates of the violations are reported as a period of time and the summary violations in paragraphs (6)(a)-(c) do not include specific dates. At this point in the legal process, the exact dates may not be known or they may not be relevant for the purpose of the indictment. Second, the number of parties involved in the violation may not be known, as illustrated in the above excerpt "others… unknown to the Grand Jury." As we show in the following sub-sections, this missing information affects the quality of scenario and goal analysis in different ways. Finally, the indictments trace from each violation in paragraphs (6)(a)-(c) to specific paragraphs in corresponding laws that were violated. These references can be useful to identify potential "hotspots" in regulations and prioritize related requirements by surveying multiple indictments.

An important observation not shown in this excerpt is that subsequent, numbered paragraphs include backward references to this paragraph. These cross-references are used to refer back to details that are shared across these different contexts, including actors and events. Similar to the analysis method that we employ on regulations [4, 5], analysts must incorporate these details in each new context to accurately represent the individual charges.

We now discuss each notation using examples drawn from the above excerpt. For simplicity, we do not separate independent events identified in a single phrase from the original text into separate statements. For example, the events in the phrase "sell, transfer and use" are not separated unless necessary; however, this separation is desirable in practice to independently reason about different prevention and mitigation strategies. In addition, we simplify and generalize event descriptions where appropriate by removing phrases for presentation purposes.

5.1. Sequence Diagrams

Sequence diagrams are an Object Management Group (OMG) standard included in the popular Unified Modeling Language (UML). Using sequence diagrams, engineers can describe the functions of individual objects in a linear-time, total-order notation that does not support modeling concurrent events; see related work on state charts for a partial-order notation that supports concurrency [16]. Because engineers are familiar with sequence diagrams, others have used this notation to describe scenarios and the actions of actors in an analogous manner [32, 33].

Figure 1 shows a sequence diagram acquired from paragraph (6), above. The primary disadvantage of sequence diagrams, observed in modeling this excerpt and observed in the other indictments we considered, is the missing temporal information necessary to create a sequence of events. For example, in paragraphs (6)(a)-(c), the engineer must make additional inferences to temporally order the events "exceed authorized access," "possess and use" and "sell, transfer and use" presented in Figure 1. These inferences include deciding that the phrase "things of value" in paragraph (6)(a) includes both "a means of identification of another person" and "individually identifiable health information" in paragraphs (6)(b) and (6)(c), respectively, which may or may not be accurate. Therefore, it is impossible to use sequence diagrams to

[6] DISCLAIMER: Any statements made in this paper are intended to reflect the actual charges stated in the indictments and are not intended to suggest guilt or innocence of the defendants.

accurately represent these events because the text does not explicitly state the order of these events. The excerpt also contains other ambiguities. For example, the excerpt states that "access to a protected computer" occurred "to obtain things of value;" however, we cannot assume the "things of value" were obtained from the computer. It may have been that the computer was a means to access other computers or contact other individuals. Because sequence diagrams model transactions as functions between two objects, we address this ambiguity by introducing anonymous parties A and B in Figure 1 that may be the computer, the defendants or another actor. We believe this application of the notation, while accurate in our explanation, is exceptional and potentially misleading.

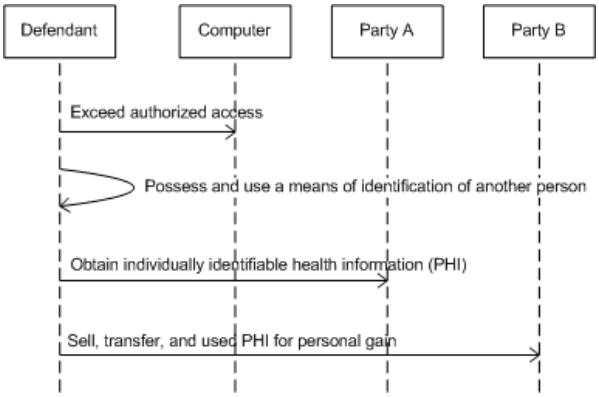

Figure 1: Inaccurate sequence diagram from paragraph (6), United States v. Ferrer

The detail in sequence diagrams, which allows us to model the interactions between the actors, can also be an advantage to this notation. The value of using sequence diagrams must be determined based on the detail of information provided by the document and the detail necessary to accurately describe a scenario. We have highlighted an example where sequence diagrams are potentially inaccurate and misleading however, paragraphs 8-12 (Manner and Means) explicitly state an order of events. These paragraphs connected events using keywords "then" and "next." Using sequence diagrams to model the scenario described by these paragraphs, we can accurately model the order of events and interactions between actors with the same level of detail provided in the indictment.

5.2. Misuse Case and KAOS Diagrams

Sindre and Opdahl introduced the misuse case diagram [34]. In misuse case diagrams, actors are linked to misuse cases that represent misuses of the system. As in goal-oriented requirements engineering [6], these cases can be refined into sub-cases using "includes" links. Misuse cases are typically used to motivate discussions and elicit potential misuses of the system from stakeholders. In this section, we sought to use misuse cases to identify relationships that map actors to actions that result in legal vulnerabilities.

Figure 2 presents the same subset of events from paragraph (6) that appear in the sequence diagram in Figure 1. In Figure 2, the phrase "agree to defraud the United States" from paragraph (6)(a) is mapped to a misuse case and refined by the sub-cases "exceed authorized access to a protected computer" from paragraph (6)(a) and "sell, transfer and use… information for personal gain" from paragraph (6)(c). We identified these sub-cases using the phrase heuristics for identifying purposes and instruments in an activity description [3].

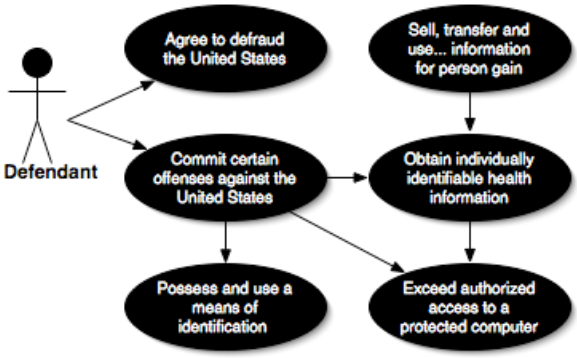

Figure 2: A misuse case diagram from paragraph (6), United States v. Ferrer

Similar to misuse case diagrams, KAOS diagrams can represent anti-goal and associated threat agents [23]. Unlike misuse case diagrams, anti-goals are refined using logical "AND" and "OR" nodes to represent possible alternatives. In KAOS goal models, agents are uniquely associated with the lowest-level goals in the refinement hierarchy to support operationalization [23]. However, in KAOS anti-goal models, the threat agents must be associated with anti-goals at various levels within the hierarchy. During refinement of an anti-goal model the threatening agents are refined to be responsible for a leaf level anti-goal. Figure 3 presents the same subset of events from paragraph (6) that appeared in Figure 2 using the KAOS method to represent anti-goals.

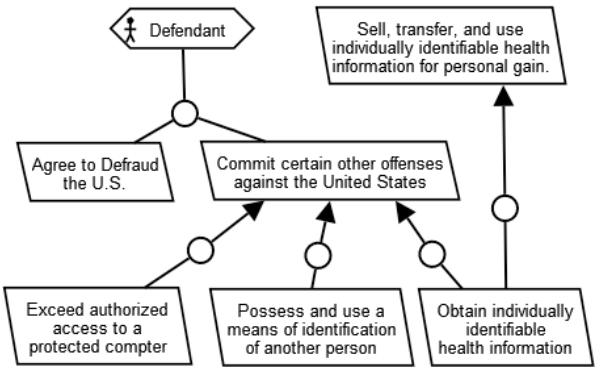

Figure 3: A KAOS diagram from paragraph (6), United States v. Ferrer

To illustrate the benefit of refinement using "AND" and "OR" links, we separate the events "sell, transfer and use" into three anti-goals. The English conjunction "and" is logically ambiguous and can be interpreted as a logical disjunction [4, 5]. Thus, the analyst can create mitigating requirements for each event as if they occur independently. For example, one can prevent "use" by encrypting the information, assuming the threat agent does not have the means to decrypt the information. However, sales and transfers are not prevented or mitigated by encryption; thus another goal is required to prevent or mitigate these threats. This benefit, available in KAOS, of independently considering alternatives is required to diagram legal vulnerabilities and is not present in misuse cases diagrams. Alternatively, misuse case diagrams provide the "excludes" link, which we did not employ in this study, but which may be relevant in the analysis of other indictments.

6. Identifying Vulnerabilities

Vulnerabilities describe something that is open to attack or damage. In this paper, we are interested in legal vulnerabilities that increase exposure to civil or criminal liability. We are primarily concerned with the role of software systems in the vulnerability, either as the cause, means, or target of the attack or damage. To illustrate, we use a KAOS diagram to identify vulnerabilities and propose mitigating requirements. Consider the following excerpt from paragraph (4) in United States v. Fumo [11] in which the actors are *italicized* and the events are underlined. This excerpt follows a description in the indictment of a U.S. Federal Bureau of Investigation (FBI) and Internal Revenue Service (IRS) investigation into allegations of misusing public assets, providing illegal political favors, fraud and extortion.

"4. It was a part of the conspiracy that, in both contemplation of and with actual knowledge of the investigation described above, and for the purpose of destroying e-mail and other electronic evidence in order to prevent the FBI, the IRS, and this federal grand jury from receiving or reviewing such evidence in the course of the investigation, *defendants VINCENT J. FUMO, RUTH ARNAO, LEONARD P. LUCHKO, MARK C. EISTER*, *Person No. 1*, *Person No. 11*, and *other persons*, both known and unknown to the grand jury (collectively, "*the conspirators*"),

(a) systematically destroyed e-mail communications sent to or received from FUMO and ARNAO;
(b) created and implemented a formal schedule to run specialized computer programs known as Secure Clean Deep Clean and PGP Free Space Wipe that erased any trace of deleted electronic files on computer hard drives, servers, PC cards, and other electronic storage devices;
(c) instructed FUMO's employees that under no circumstances, without FUMO's permission, were they permitted to save any e-mail sent to or received from FUMO;
(d) logged into the e-mail accounts of FUMO's employees to scan their e-mail to determine that they were, in fact, not saving any e-mail relating to FUMO; and
(e) deleted and wiped other electronic equipment, such as the Blackberry communication devices used by FUMO and ARNAO, among other persons."

The above excerpt from paragraph (4) describes a number of means in sub-paragraphs (4)(a)-(e) by which the charges state that Fumo et al. tried to prevent the FBI, IRS, etc. from receiving evidence. Figure 4 shows a partial KAOS diagram depicting the charges that were derived from paragraph (4) as anti-goals and their refinements. An analyst who seeks to identify preventative or mitigating requirements must decide which anti-goals they can affect through requirements. For example, the anti-goal "Implement a formal schedule to erase any trace of files" cannot be mitigated without preventing users from deleting files. However, one can mitigate the anti-goals to wipe electronic equipment and destroy e-mail by performing secure off-site backups, as shown in Figure 4. Similarly, one can mitigate (but not prevent) unauthorized individuals from logging into e-mail accounts by restricting access to e-mail accounts, for example by having an independent administrator centrally manage e-mail accounts and assign unique usernames and passwords.

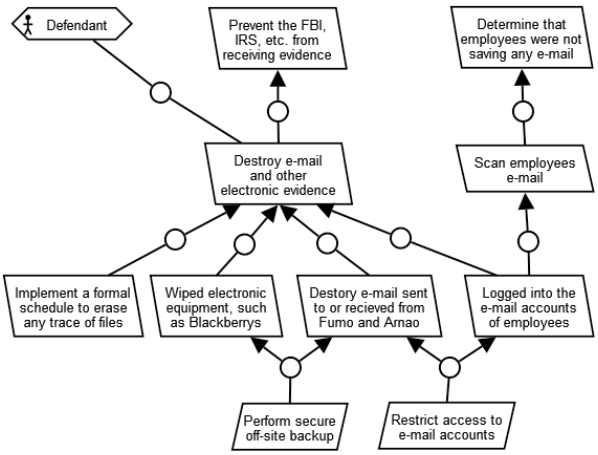

Figure 4: A KAOS diagram from paragraph (4), United States v. Fumo

The ability to integrate mitigating requirements into misuse case and KAOS diagrams provides important traceability between requirements and the rationale behind legal vulnerabilities.

7. Discussion and Future Work

This study yielded several important insights that motivate using all of sequence, misuse case and KAOS diagrams to identify vulnerabilities from criminal case documents. To a great extent, the applicability of each diagram notation depends heavily on the presence or absence of relevant information.

In Section 5.1, we observe an instance in which sequence diagrams cannot be accurately created from criminal indictments. While not always the case, it was generally found that misuse case and KAOS diagrams could represent actors and events using refinement hierarchies. As discussed in Section 5.2, the notable difference was that KAOS diagrams provide an additional distinction through AND/OR-refinement links that are necessary to capture the exclusivity of separate charges described in the indictment, despite the appearance of events occurring in conjunction to achieve some overall goal. This approach is still amenable to identifying mitigation strategies that address a single event, even if this event occurs repeatedly. For example, the act of exceeding authorized access to health information, as a single event, may be difficult to mitigate. This is especially true if observable behavior includes authorized access that is normally granted to the malicious user. However, as a sequence of multiple events of the same type, it may be possible to discern that access is in excess of normal behavior similar to *operational profiles* in software testing, in which normal frequency of use helps determined behaviors that are out of the norm [24].

However, goal-oriented models that do not express temporal relations, such as the misuse case and KAOS diagrams that we examined, will fail to capture a class of vulnerabilities that is only exploitable through transactions or a sequence of dissimilar events. For example, the acts of exceeding authorized access (for unauthorized purposes and in numbers beyond the average operational profile) to health information and subsequently using the information to file fraudulent insurance claims is a complex vulnerability: the health information is usable to file claims independent of the health care provider maintaining the information. Requiring that claims be filed using a secret known only to the provider and the agency, called a *shared secret*, mitigates this vulnerability as it renders the act of acquiring the information useless in the process of filing claims. The insurance agency would thus reject claims filed without the shared secret. Observing the applicability of this mitigation strategy, however, benefits from the explicit representation of temporal relations between events. While this information may be available in trial transcripts, the value of expending this additional effort to analyze these transcripts must be determined by future work.

The limits of this study reveal fertile ground for future work. For example, this case study did not examine cases that went through the U.S. federal appeals process. Cases that are appealed are used to decide legal precedent and constitute an extension or retraction to statutory law and/or case law. The decisions in these cases can be used to reinforce prior decisions regarding known vulnerabilities or to yield insight into new vulnerabilities through new interpretations of existing laws.

Acknowledgments

We thank the United States Attorney's Office for the Western District of Washington, Professor Sara Sun Beale in the Duke University School of Law, and ThePrivacyPlace.org for their feedback. This work was funded by the IBM PhD Fellowship (RTP CAS) and NSF #032-5269 and NSF #043-0166.

References

[1] I. Alexander, "Initial industrial experience of misuse cases in trade-off analysis," *IEEE Joint Int'l Conf. Req'ts Engr.*, pp. 61-68, 2002.

[2] A.I. Antón. *Goal-based Requirements Analysis Method*, PhD Thesis, Georgia Tech, 1996.

[3] T.D. Breaux, A.I. Antón. "Analyzing goal semantics for rights, permissions and obligations," IEEE Int'l Conf. Req'ts Engr., pp. 177-186, 2005.

[4] T.D. Breaux, M.W. Vail, A.I. Antón. "Towards compliance: extracting rights and obligations to align requirements with regulations," *IEEE Int'l Conf. Req'ts Engr.*, pp. 49-58, 2006.

[5] T.D. Breaux, A.I. Antón. "Analyzing regulatory rules for privacy and security requirements," *IEEE Trans. Soft. Engr., Special Issue on Soft. Engr. for Secure Sys.*, 34(1): 5-20, 2008.

[6] A. Dardenne, A. van Lamsweerde, S. Fickas. "Goal-directed requirements acquisition", Science of Computer Programming. 20:3-50, 1993.

[7] Ernst & Young, 10th Annual Global Information Security Survey: Achieving a balance of risk and performance, 2007.

[8] C.B. Farrell, "ChoicePoint settles data security breach charges; to pay $10 million in civil penalties and $5 million for customer redress," FTC File No. 052-3069, Office of Public Affairs, U.S. Federal Trade Commission, 2006.

[9] United States v. ChoicePoint, Inc., Case No. 1:06-CV-00198-JTC, N.D. Ga., Feb. 15, 2006.

[10] K. Fogerty, K. Kimberland, "2006 E-crime watch survey from CSO Magazine reveals insider threats are on the rise," *CSO Magazine*, Sep. 2006.

[11] United States v. Luchko, et al, Case No. 2:06-CR-00319-WY, E.D. Pa., Feb. 6, 2007.

[12] United States v. Ferrer, et al. Case No. 0:06-CR-60261-JIC, S.D. Fl., Dec. 7, 2006.

[13] Federal Rules of Criminal Procedure, last amended December 2007.

[14] B.A. Garner, Ed., *Black's Law Dictionary*, 8th ed., Thompson West, 2004.

[15] S. Ghanavati, D. Amyot, L. Peyton, "Towards a framework for tracking legal compliance in healthcare," *19th Int'l Conf. Adv. Info. Sys. Engr.*, pp. 218-232, 2007.

[16] M. Glinz. "Improving the quality of requirements with scenarios", *2nd World Congress for Software Quality*, 55-60, 2000.

[17] C.J. Hoofnagle, D.J. Solove, "Re: Request for investigation into data broker products for compliance with the FCRA," Electronic Privacy Information Center, Washington, D.C., 2004.

[18] P. Hope, G. McGraw, A. I. Anton, "Misuse and abuse cases: getting past the positive", *IEEE Security & Privacy*, 2(3): 90-92, 2004.

[19] D. Karagiannis, J. Mylopoulos, M. Schwab. "Business process-based regulatory compliance: the case of the Sarbanes-Oxley act," *IEEE Int'l Req'ts Engr. Conf.*, pp. 315-321, 2007.

[20] Y. Kamisar et al. *Modern Criminal Procedure: Cases, Comments, and Questions*, 11th ed., St. Paul, Minn.: Thomson/West, 2005, pp. 2-20.

[21] A. van Lamsweerde, E. Letier, "Handling obstacles in goal-oriented requirements engineering," *IEEE Trans. Soft. Engr.*, 26(10): 978-1005, 2000.

[22] A. van Lamsweerde, R. Darimont, E. Letier, "Managing conflicts in goal-driven requirements engineering," *IEEE Trans. Soft. Engr.,* 24(11): 908-926, 1998.

[23] A. van Lamsweerde, "Elaborating Security Requirements by Construction of Intentional Anti-Models," *IEEE 26th Int'l Conf. Soft. Engr.*, pp. 148-157, 2004.

[24] M.R. Lyu. Handbook of Software Reliability Engineering, McGraw-Hill, 1995.

[25] J. McDermott, C. Fox, "Using abuse case models for security requirements analysis", 15th Computer Security Applications Conf.,, pp. 55-64, 1999.

[26] F. Massacci, M. Prest and N. Zannone. "Using a security requirements engineering methodology in practice: the compliance with the Italian data protection legislation," *Computer Standards & Interfaces,* 27(5):445-455, 2005.

[27] P.N. Otto, A.I. Antón, and D. Baumer. The Choicepoint Dilemma: How Data Brokers Should Handle the Privacy of Personal Information. *IEEE Security & Privacy*, 5(5), pp. 15-23, September/October 2007.

[28] P.N. Otto, A.I. Antón, "Addressing legal requirements in requirements engineering," *15th IEEE Int'l Req'ts Engr. Conf.*, pp. 5-14, 2007.

[29] L. Payton, S. Ghanavati, D. Amyot. "Designing for privacy compliance and performance management in health care", Design Principles and Practices: An International Journal. Volume 1, number 3, 2007.

[30] G. Peterson, J. Steven, "Defining misuse within the development process," *IEEE Security & Privacy*, 4(6): 81-84, 2006.

[31] G. Regev, I. F. Alexander, A. Wegmann. "Modeling the regulative role of business processes with use and misuse cases". Business Process Management Journal. Vol. 11 No. 6, 2005.

[32] B. Regnell, M. Andersson, J. Bersrand. "A hierarchical use case model with graphical representation", *IEEE Int'l Symp. and Workshop on Engr. of Computer-based Sys.*, pp. 270-277, 1996.

[33] H. Saiedan, P. Kumarakulasingam, M. Anan. "Scenario-based requirements analysis techniques for real-time software systems: a comparative evaluation", *Req'ts Engr*. 10:22-23, 2005.

[34] G. Sindre, A.L. Opdahl. "Eliciting security requirements with misuse cases", Req'ts Engr. 10:34-44, 2005.

[35] United States Sentencing Commission. "2006 annual report," p34, 2006. http://www.ussc.gov/ ANNRPT/2006/ figc.pdf

[36] D. Verdon, G. McGraw, "Risk analysis in software design," IEEE Security & Privacy, 2(4): 79-84, 2004.

[37] P. Winn, "Confronting the threats of medical identity theft," *Health Information Privacy/ Security Alert*, July 24, 2007.

[38] S. Yanovitch, K. Kimberland, "2007 E-crime watch survey shows security incidents, electronic crimes and their impact steady versus last year," *CSO Magazine*, Sep. 2007.

[39] R.K. Yin. *Case Study Research*, 3rd ed. Applied Social Research Methods Series, v.5, Sage Pubs., 2003.