Central Global University, CGU-Georgia

Thesis

Title:

Risk in Cyber Systems and Security Administration

Submitted by:

Mailwasagan Udaya Jeewan

(CGU SID NUMBER)

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Cybersecurity

Supervisor:

Mrs. Joona Komban

Date:

10-October-2025

RISK IN CYBER SYSTEMS AND SECURITY ADMINISTRATION

DISSERTATION / THESIS SUBMITED TO THE DEPARTMENT OF MANAGEMENT

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UDAYA JEEWAN

P.hd

Abstract

Significant uncertainty surrounds cyber security investments. Chief information officers (CIOs) operate with limited resources and typically do not know the relative risk of different cyber attack vectors, such as malicious email, website attacks, or lost laptops. Therefore, CIOs currently have difficulty in assessing the risk reduction associated with different cyber security investments, possibly resulting in a poor allocation of resources. For example, an organization might dedicate significant resources to detecting malicious insiders, even though its risk from website hacking is much larger.

Presently, cyber risk is managed qualitatively in most organizations. Current best practices rarely incorporate quantitative risk tools and instead largely advocate the use of risk matrices, which are ambiguous and lack the ability to incorporate system dependencies. This dissertation discusses the application of probabilistic risk analysis (PRA) to cyber systems, which allows decision makers to rigorously assess the value of cyber security safeguards. First, different classes of attack scenarios are modeled. For example, laptops are lost or stolen, websites are defaced, phishing emails attempt to steal employee credentials, and malware infects machines via web browsing. Next, the rate and consequences of each scenario are assessed, drawing heavily from historical data at organizations, academic literature, publicly available data, and expert knowledge. In the case of large or rare cyber incidents where sufficient data do not exist, scenario analysis is used to obtain probabilistic assessments. These data initialize a Monte Carlo simulation to calculate probability distributions of monetary losses resulting from cyber incidents at the organization. Next, safeguards are considered that change the rate or impact of the scenarios. Changing the model structure or the model inputs shows how each safeguard affects the consequence distribution, essentially demonstrating the value of each safeguard. Sensitivity analysis can also be performed to identify the important uncertainties and the robustness of different safeguard implementation decisions.

The process described above is a framework for the quantitative assessment of cyber risk in dollar terms. The result is that cyber security safeguards can be valued and prioritized. To demonstrate this framework in action, this dissertation describes a general model combined with a detailed case study of cyber risk quantification at a large organization. Over 60,000 cyber security incidents from that organization are analyzed and used to initialize the model to determine the cost-effectiveness of security safeguards including full disk encryption, two-factor authentication, and network segmentation. These data provide useful statistics for low and medium level incidents, but some incidents may be absent from the data because large incidents have not yet occurred, or have occurred too rarely to obtain good estimates for the probabilities. In this case, classes of scenarios are modeled and initialized with conditional probabilities elicited from experts. The data driven

model is combined with the scenario based model by overlapping the two cost curves to ensure that incidents are not double counted, resulting in a complete and comprehensive assessment of cyber risk at the organization.

Risk quantification is a critical requirement for organizations. A lack of real-world data and massive uncertainty about cyber impacts has limited progress, but organizations can now be armed with the information and tools needed to measure cyber risk. Cyber security continues to be a rapidly evolving domain, but risk quantification illuminates the cyber landscape and enables defenders to improve resource allocation and optimize decision making.

Acknowledgements

The author would like to thank his advisor Elisabeth Paté-Cornell, and his committee members Nick Bambos, Herb Lin, and John Mitchell. Additionally, Thomas Kenny, who was chair of the oral defense. A large number of other people made this work possible, and the author is grateful for the support and mentorship of these individuals. In particular, this dissertation would not exist without Elise Kuypers. NB, AK, PK, CM, and JR also deserve special thanks.

Portions of this work were done while the author was supported by a grant from the Jet Propulsion Lab, a fellowship from the Center for International Security and Cooperation (CISAC), and the Burt and Deedee McMurtry fellowship. The author is grateful for this support.

Contents

List of Illustrations

1 Introduction

Quantifying cyber risk is a major challenge to some organizations. For example, it is unclear whether chief information security officers (CISOs) should be more concerned about laptop theft or phishing emails.[1] Security vendors offer many solutions, but rarely discuss how the costs and benefits of different technologies impact the overall risk exposure of an organization. As a result, bad security investments are probably common. For example, one security analyst was asked to name the largest cyber risk to her organization and where the majority of the security budget should be focused during the next year. The analyst replied that as far as she knew, malicious insiders were the greatest risk and that resources should be concentrated on detecting employee misconduct. However, cyber security incident data from that organization showed that over a five-year period, only one malicious insider attack had occurred.[2] The organization had fired an employee on a particular day, but did not revoke that employee's login credentials, leading to the employee gaining access to IT resources later in the evening and disrupting several services.

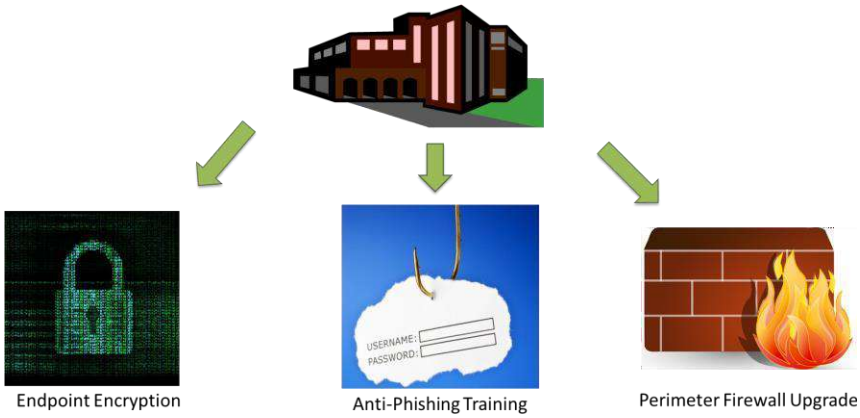


Figure 1: Potential organizational investment decisions. Organizations are currently uncertain about the effectiveness of security investments. For example, an organization could invest in encryption software to secure information on lost devices, conduct anti-phishing training to educate its workforce about clicking on malicious emails, or upgrade network equipment to enhance its firewall. Currently, given the lack of specific information, it is difficult to compare the risk reduction associated with these investments and to prioritize security investments.

[1] The title of the person responsible for cyber security varies across organizations. The basic guidelines used in this paper are consistent with the following definitions. Chief information security officers (CISO) manage the daily technical IT operations at an organization. CISOs typically have good technical skills and extensive knowledge of an organization's network and defenses. Chief information officers (CIO) act in an executive role, working to set strategy and policy for IT operations. Note that other organizations define these positions differently or may have other titles (chief technology officer, chief security officer, chief risk officer, etc.).

[2] Other sources define *insider attack* as both the malicious and accidental incidents. For the purposes of this paper, accidental incidents are referred to as *data spillage*, while *malicious insider* incidents refer to deliberate acts intended to cause harm.

During the same time period (five years), hundreds of successful website compromises occurred at the same organization, including website defacements[3], SQL injections[4], and even persistent adversaries compromising web applications to obtain sensitive information. After a careful analysis, it was clear that website attacks occur more often than insider attacks. Further, while the impacts from each type of incident differ, the analysis showed that insider attacks (including potential incidents that had not yet occurred) were not significantly more impactful than high-impact website incidents. Faced with this information, the analyst realized that her intuition did not match the ground truth about risk.

New methods are needed to provide decision support for cyber security investments. Organizations can combine ground-truth statistics on historical incidents to obtain data-driven risk assessments of cyber risk. In some cases, these historical data are insufficient because of changing trends, or because certain large impact incidents have yet to occur. In these cases, the data can be supplemented with a Bayesian analysis of different scenarios. Organizations armed with the data and the tools to analyze cyber security can proactively respond to emerging threats and make efficient tradeoffs between security and usability, optimize scarce resources including time and money, and adapt to an evolving cyber domain. Given these complicated tradeoffs, it is extremely difficult for organizations to compare the actual risk reduction associated with different cyber security investments via the current qualitative methods that are widely used. Qualitative risk analysis driven by sometimes biased perceptions can result in poor returns on security investment. The method presented in this dissertation allows organizations to make better security investment decisions by rigorously quantifying cyber risk and the value of different security safeguards.

1.1 Research Motivation

Cyber risk is a major concern to organizations. Cyber systems are now crowded with criminals, amateur hackers, government actors, hacktivists, and other adversaries. Media attention to cyber security issues has grown dramatically over the past several years as well. During the holiday shopping season in 2013, over 40 million credit cards were stolen from Target's point-of-sale terminals (Krebs, 2014a). The following year, details on 56 million credit cards were stolen from Home Depot in a similar attack. In February 2015, personal information from nearly 80 million people was stolen from the healthcare company Anthem (Krebs, 2015a).[5] At first, small businesses

[3] Website defacements consist of an attacker making unauthorized visual changes to a webpage.
[4] SQL (structured query language) injections are specially crafted queries designed to obtain information from online databases by using commands in unintended ways.
[5] Details on this data breach and others listed in this paper can be found at the website for Privacy Rights Clearinghouse, at privacyrights.org

could rely on a low profile to avoid cyber attacks, but criminals pivoted to banking fraud and ransomware[6] attacks in 2104 and 2015 to target small organizations (Krebs, 2015b). New classes of criminals also began to emerge, including self-organized activist groups and nation state–tolerated groups. Even businesses that specialize in hacking have become victims; for instance, in 2015 Hacking Team experienced a damaging data breach that exposed activities at the company (Greenberg, 2015).

Along with a large increase in media attention, cyber attacks have been shown to impact businesses and individuals in unique and surprising ways. Credit monitoring and breach notifications are major drivers of data breach costs; distributed denial of service (DDoS)[7] attacks cause business interruption and cyber attacks have led to the destruction of physical components (Zetter, 2015). Cyber security is now recognized as a substantial threat to organizations and to national security more generally. In the Worldwide Threat Assessment delivered annually to Congress by the US Director of National Intelligence, cyber has been the first threat listed each year since 2013. Organizations are likely to face a dynamic and rapidly evolving cyber threat landscape for several years to come, and organizations are struggling to hire well-trained cyber professionals due to a shortage of qualified individuals. Vendor solutions have increased as well, but organizations have a limited ability to evaluate the benefit of different security products. Cyber security is clearly a major concern for organizations in the foreseeable future.

1.1.1 Decision Makers Have Limited Tools to Assess Cyber Risks

Most decisions about security investments at organizations are made using heuristics. The CISO might rely on intuition, industry reports, or a security product salesperson to determine which technologies to implement. If an actual process exists to assess cyber risk, it is most often qualitative. For example, a risk matrix might be used to assess the likelihood and impact of a certain scenario or class of risks.

[6] Ransomware is a type of malicious software that will encrypt a victim's hard drive. Once the encryption is complete, the software demands a ransom (often between $500 and $2,000) to decrypt the files. If the victim has not retained periodic backups, then the ransom must be paid or the information is lost.

[7] DDoS attacks occur when internet traffic is funneled to a webpage. The huge increase in traffic can overwhelm the page, making it unavailable to legitimate users.

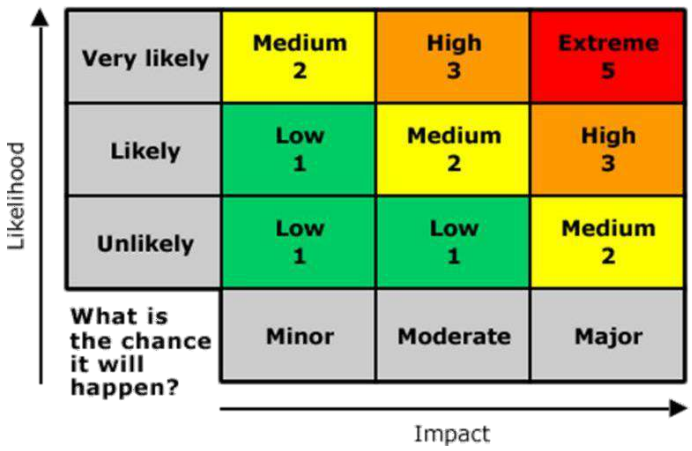


Figure 2: Risk matrices. A typical risk matrix used to evaluate risk. Risk matrices have important limitations and are not sufficient for conducting a risk analysis of cyber systems.

Risk matrices can be effective tools for stakeholders to brainstorm different risks, but have important limitations (Cox, 2008). Risk matrices do not address combinations of risks or dependencies among them. Also, the rich toolset associated with probabilistic risk analysis is not usable in a qualitative risk framework. Sensitivity analysis, optimization, and many other powerful methods cannot be applied to a risk matrix.

Qualitative tools can also lead to significant confusion. The terms used in a risk matrix, such as *unlikely*, are ambiguous and can be interpreted differently by different people. In 2008, researchers surveyed petroleum engineers to study the perception of different terms, such as *reasonably certain* and *proved* (McLane, Gouveia, Citron, MacKay, & Rose, 2008). Unsurprisingly, they found that the engineers had different interpretations of many of these words. For example, some engineers stated that oil would be found in 100% of fields with *proved* reserves, while others said that oil would be found on only 25% of the sites. This large discrepancy in the interpretation of subjective words is incompatible with a rigorous analysis of risk. In fact, significant literature exists that demonstrates biases and heuristics in humans that lead to suboptimal decision making (Tversky & Kahneman, 1974). Even seemingly simple details such as the phrasing of a question can lead to significant biases.[8]

Qualitative risk methods are particularly susceptible to biases and can lead to poor decision making. Cyber security is especially prone to these biases given the disproportionate amount of media attention placed on certain incidents, the sparseness of certain types of events (leading to a

[8] For example, if one group of doctors is asked if they recommend surgery when there is a 90% survival rate in the first month, 84% recommend surgery. Another group of doctors is asked if they recommend surgery when there is 10% mortality in the first month and only 50% recommend surgery. The question is rephrased, leading to the change (Kahneman, 2011). This bias has been studied in many other contexts and is often called the framing bias.

recency bias[9]), and large amounts of uncertainty. For example, the CISO of one organization was discussing security investments with an analyst. The organization had recently implemented full disk encryption and a data backup program, and was considering an asset recovery program.[10] The analyst performed a quick analysis to determine if the asset recovery software was a good investment. The asset recovery software was relatively expensive, costing about $150,000 in licensing per year. However, based on the number of laptops stolen over the past five years and assuming the technology could recover 100% of the stolen devices (an extremely optimistic assumption), the analyst determined that the organization would recover only about $50,000 in stolen equipment per year.[11] Therefore, asset recovery was a bad deal for the organization, because it would cost $150,000 to recover just $50,000 in lost assets.[12] It made much more sense to simply replace the laptops. Even some of the most basic data-driven methods can alleviate these biases by eliminating ambiguity in risk assessments, leading to better decision making.

1.1.2 Probabilistic Risk Analysis Potential

Probabilistic risk analysis (PRA) has been used to manage risk in a number of other fields including nuclear safety, space systems, and medical devices (Garrick et al., 1967; Paté-Cornell & Fischbeck, 1993; Paté-Cornell, Lakats, Murphy, & Gaba, 1997).[13] However, many of the powerful PRA tools and techniques have not yet been applied to the cyber domain. PRA is especially needed in the cyber domain to eliminate biases, calculate benefit-cost tradeoffs, and to address uncertainty in events that have not yet occurred.

As discussed earlier, biases permeate cyber security. Increased media attention has created the perception that data breaches are increasing in frequency, although evidence shows that this is not the case (Wheatley, Maillart, & Sornette, 2016; Maillart & Sornette, 2010; Edwards, Hofmeyr, & Forrest, 2015). Certain attack vectors that occur rarely (e.g., supply chain attacks or malicious insiders[14]) are areas of intense media focus, while more frequent and harmful vectors (phishing and website attacks) generate less attention. PRA provides a rigorous way to assess cyber risk and can

[9] Recency bias involves placing more emphasis on events that recently occurred.
[10] Asset recovery software is loaded onto laptops and beacons out to law enforcement if it is stolen. Law enforcement uses this information about its location to recover the device.
[11] The CISO and the analyst discussed other factors and determined that the information on the stolen devices was irrelevant because it was encrypted, and the organization made regular backups. Further, downtime would not be eliminated with asset recovery because the lost laptops would often take many days or weeks to be recovered.
[12] See Chapter 5 for a full case study of asset recovery programs at a different organization.
[13] Probabilistic risk analysis uses systems analysis, Bayesian probability, and other methods to quantitatively assess risk. This can involve the integration of data (statistics when they exist), expert opinion, and other sources of information.
[14] Note again that *malicious insiders* refers to intentional actions by an insider, not mistakes (which are labeled *data spillage* incidents. See Chapter 3 for more details.

eliminate many of the false perceptions analysts may have about cyber security. PRA also provides a rigorous, repeatable, and accurate way to assess the benefits and costs of cyber security investments. The number of security vendors offering services has risen sharply, but organizations cannot accurately quantify the value of safeguard investments via qualitative methods.

PRA is also needed to address the pervasive uncertainty in cyber security stemming from two areas: epistemic uncertainty arising primarily from the probability of large impact incidents that may not have occurred yet and the losses associated with these severe cyber incidents, and the aleatory uncertainty of cyber impact severity in cases where data exist. Many organizations are driven to qualitative assessments because of the fundamental difficulty of assessing cyber impacts, such as reputation damage. An organization might assess that losses could range between $10 million and $50 million, so this prospect is simply labeled "high impact." Assessing the probability of rare, large incidents can be even more challenging. Instead of masking the uncertainty in qualitative buckets, PRA incorporates these uncertain assessments directly into the analysis, which often leads to different results compared to the case where point estimates of costs are used. Probabilistic assessments lead to a much higher quality conversation about risk.

PRA is also needed to deal with the aleatory uncertainty of cyber impact severity. Surveys and industry reports are saturated with calculations of the "average" cost of a data breach, or a "typical" impact. However, some consequences of cyber impacts are extremely heavy-tailed, meaning that using the average is misleading at best and harmful at worst. Figure 3 shows the distribution of hours of investigation for lost devices at a large organization (historical data spanning six years), which is one of many impacts that result from an incident. The median incident size is 0.5 hours of investigation, but some incidents occur (although much more rarely) that are three orders of magnitude larger. In heavy-tailed distributions, the mean or median is a poor representation of the distribution. Some cyber impacts even have mathematically undefined averages.

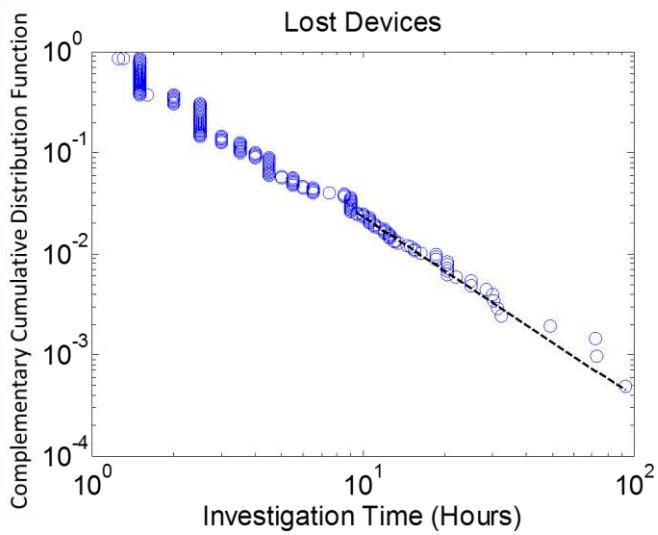


Figure 3: Heavy-tailed distributions. The complementary cumulative distribution function for the investigation time of lost devices at a large organization. Note that the distribution is linear on a log-log plot, meaning the distribution is heavy-tailed.

One of the main benefits of probabilistic risk analysis is the ability to look at distributions instead of a single loss metric. Explicitly incorporating uncertainty into the cyber risk assessment process allows decision makers to trade off low-frequency, high-impact incidents with high-frequency, low-impact incidents. Poor treatment of the uncertainty in cyber systems leaves decision makers prone to misunderstanding risks.

1.2 Research Scope

This dissertation presents a method for conducting a quantitative risk assessment of cyber systems at a fictitious but realistic organization, called Space Corp. Space Corp conducts research and development on aircraft, spacecraft, and other related technologies. It is a private organization that receives contracting from public and private customers. Space Corp employs roughly 20,000 individuals in the United States. In addition to conducting research and development, Space Corp also operates a small constellation of satellites and infrastructure to support these spacecraft.[15] The objective is to calculate a complete risk curve so that different security safeguards can be prioritized.

The full risk curve is made up of different regimes, namely a portion that consists of frequent, low-impact incidents, and a portion consisting of larger but rarer incidents. While some organizations may have large amounts of historical data, these past statistics are limited to incidents that have already occurred and may not address incidents that are either rare, or new types of

[15] Again, Space Corp is a realistic but fictitious organization designed to illustrate interesting cyber security considerations for an aerospace organization. Any similarities to real aerospace organizations are coincidental and unintentional.

incidents that have not emerged yet. Therefore, historical data can be a useful starting point, but should be augmented with an analysis of different classes of scenarios.

To assess the total risk, the risk curve is broken into three different regimes.

1. **Data-driven model**: derived from historical incidents if the data exist and the data are stable over time.
2. **Scenario-based model**: used to model incidents that have not yet occurred (typically large impact incidents).
3. **Overlap regime**: combines the data-driven model and the scenario-based model by overlapping the risk curves to avoid the double counting of incidents.

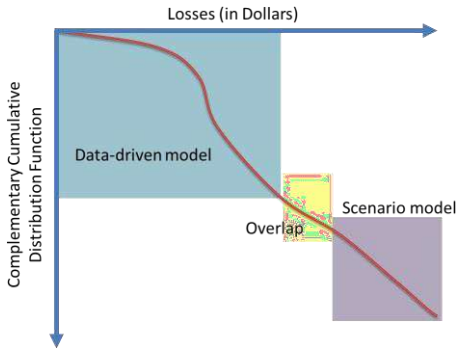


Figure 4: The three regimes of a risk curve.

The actual risk curve is generated via a Monte Carlo simulation and a probabilistic model that is initialized with information gathered from data, experts, and a Bayesian analysis of certain scenarios. These parts fit together into a general model for assessing cyber risk at an organization.

Modeling cyber security necessitates a systems analysis that incorporates information about the attackers, defenses, the system, and the impacts. Information about attackers is useful to determine how to best defend the organization, meaning that analysts can determine who they are, what they know, and what they want. Additionally, many organizations have limited data about how their systems operate, and what requires protection (i.e. what the crown jewels are). A cost model is formulated to quantify impacts of cyber incidents, including investigation time, reputation damage, business interruption, and others. This is done in dollar terms, since monetary impacts are needed to compare cyber risk to other types of business risk. Finally, the cost-effectiveness of different security investments is assessed by modeling the risk reduction associated with each safeguard.

While cyber security is relevant to many stakeholders at many different levels, the focus of this dissertation is to assess cyber security at the organizational level, not at the individual or

government level. The model presented in this paper is not well suited for governments, who face strategic cyber security decisions that interact with a complex economic and political landscape.[16] The method presented here is best suited for organizations or even smaller scales, including risk management at a project level.

1.3 Dissertation Overview

The remainder of the dissertation is as follows. Chapter 2 covers the background of cyber security in organizations, as well as how qualitative risk analysis and quantitative risk analysis tools have been developed and implemented. Chapter 3 describes a quantitative risk model for assessing cyber risk at a specific organization, and Chapter 4 provides an overview of how to analyze cyber security incident data, using data from a large US-based organization as a case study. Chapter 5 provides case studies on evaluating the effectiveness of different cyber security safeguards. Chapter 6 concludes and recommends future work.

[16] Several parts of the model would still apply to this frame, but additional components would be required to capture the geopolitical interactions of nation states.

2 Background and Related Works

Information technology (IT) has revolutionized how organizations operate. However, the convenience of IT is at odds with security. Early computer networks were designed to enable connectivity and communication. As IT became more prevalent, more systems were connected to networks. With this additional connectivity and convenience came an increase in the number of vulnerabilities. Early attacks were mostly proof-of-concept and developed by computer experts, or the result of experiments gone awry (Orman, 2003). Over time, the attackers and techniques changed. The Internet is now crowded with script kiddies[17], criminal organizations, hacktivists, and nation states attempting to compromise machines at different levels.

Attacks against organizations now include vandalism (website defacement and denial of service attacks), extortion (ransomware), organized crime (stolen credit card information), and targeted attacks to steal trade secrets. Despite these threats, some organizations have not recognized cyber security as an important concern until recently. The person in charge of cyber security varies from organization to organization and can include a chief risk officer, chief security officer, chief information officer, chief technology officer, or some other variant. Despite the C-suite notation, a chief information officer often lacks significant authority and is left with the difficult job of securing an organization with limited resources while communicating risk to a CEO and board members who may not understand cyber security.

Obtaining sufficient resources for securing an organization has always been difficult because security involves an investment to avoid a loss, but never results in a profit. Justifying security expenditures is challenging when competing with other business opportunities that have a positive return on investment. Therefore, a common pattern of events often occurs at many organizations:

1. An organization has poor information security and resources are not allocated efficiently.
2. The organization is hacked.
3. Large investments are made in security to fix the problem.

The preceding scenario is suboptimal because decisions are reactionary and made based on past outcomes, not risk. In 2014, JPMorgan Chase announced that their budget for cyber security was $250 million per year (Glazer, 2015). Several months later they were hacked, leading to another

[17] *Script kiddies* is a general term referring to amateur hackers. This class of attackers often lacks sophisticated technical abilities, but will run code written by others. Several hacking tools exist (e.g. Metasploit) with intuitive user interfaces that allow adversaries to easily launch attacks without having to identify vulnerabilities or write exploits.

announcement that the security budget would now be $500 million per year. However, it is not clear what this figure is based on—this is, why the investment is $500 million, and not, say, $450 million or $550 million. Ad hoc as these first efforts might seem, organizations now recognize that security must be managed and are seeking out tools to assess cyber risk.

2.1 Quantitative Cyber Risk

The work in this dissertation is deeply rooted in decision analysis, risk analysis, statistics, and probability theory. Decision analysis provides a normative method for approaching decision making under uncertainty (Howard, 1968; Howard & Abbas, 2015). Risk analysis provides a rich toolset for modeling a system, including its attackers, defenders, and vulnerabilities. For example, Paté-Cornell and Guikema modeled the threat of terrorist attacks and used systems analysis to rank different countermeasures (Paté-Cornell & Guikema, 2002). Other systems require models that incorporate management failures and human errors, as shown in failure analysis of offshore oil platforms (Paté-Cornell, 1993). Other domains have required precise mathematical tools to address phenomena with extremely heavy tails, such as asteroid impacts (Reinhardt, Chen, Liu, Manchev, & Paté-Cornell, 2015). Cyber security is a unique field given that it bridges many of the typical risk analysis application areas by incorporating interconnected systems, adaptive adversaries, management and human errors, and heavy-tailed distributions.

Cyber security is a broad field, and a large body of work is devoted to researching other aspects of cyber systems including international cyber law, cyber conflict, and strong encryption (Korzak, 2014; Junio, 2013; Abelson et al., 2015). While some work exists to quantify risk in limited domains (e.g., Miura-Ko & Bambos, 2007), there are fewer examples that assess cyber risk at an organizational level. For example, much work has been devoted to applying machine learning techniques to intrusion detection systems (IDS) to improve the detection of malicious activity in an organization, but this is a fundamentally different scope than quantifying cyber risk on a larger scale (Tsai, Hsu, Lin, & Lin, 2009; Mukkamala, Janoski, & Sung, 2002). However, many of these excellent technical analyses can be used as inputs to a general risk model. For example, Shostack's book on threat modeling is an excellent resource for analysts who are modeling adversaries and vulnerabilities in a system (Shostack, 2014). Attack trees and kill chains may be a useful way to model functional connections in a computer network (Schneier, 1999; Mauw, & Oostdijk, 2005; Hutchins, Cloppert, & Amin, 2011).[18] Many useful tools exist, but organizations also need a way to leverage these rigorous methods into organizational-level risk assessments.

[18] Attack trees are a representation of attacks and countermeasures. Attack trees are similar to fault trees, a well-established tool commonly used in risk analysis, but attack trees usually include actions, whereas fault trees usually include component failures. Kill chains are similarly related and involve modeling the

Researchers and IT security analysts have worked toward developing rigorous tools to assess cyber risk in organizations for over a decade. However, data poverty in the public domain has prevented the published validation of cyber risk models, which has limited the amount of work done in academia. Private consultants may actually be advancing the field the most by using incident data, Monte Carlo simulations, and sensitivity analysis to quantify risk in companies, but those companies typically do not publish details of their work for obvious reasons.[19]

Early attempts to assess cyber risk used an expectation-based approach to assess the probability of an event and its consequences. For example, one of the earliest metrics for measuring cyber risk was the annual loss expectancy (ALE), where O_i is outcome i, $I(O_i)$ is the impact of outcome i, and P_i is the frequency of event i.

$$ALE = \sum_{i=1}^{N} I(O_i)\, P_i$$

However, risk is not equal to the expectation, and over the years, more sophisticated methods were proposed. In 2000, Soo Hoo's dissertation used decision analysis and probabilistic risk analysis to assess loss distributions associated with different cyber security safeguards (Soo Hoo, 2000). Looking back, it is remarkable that such early work contained so many useful techniques, many of which have still yet to be adopted. Soo Hoo's method involved assessing attack frequencies at an organization, the probability of success of those attacks, and the monetary losses (e.g., information theft, system downtime, and information loss).[20] He then obtained assessments for the effectiveness of different security safeguards to quantitatively model the cost-effectiveness of each safeguard. Several powerful techniques are included in the model, including sensitivity analysis, value of information calculations, and probabilistic dominance. Despite the high quality of work and the demonstrated usefulness of his approach, Soo Hoo's method (while widely cited) was not incorporated into many other cyber risk models that are publicly known. The lack of data likely prevented the widespread adoption of these useful techniques.

Given the lack of public data, much of the research throughout the 2000s was focused on economic models of cyber security or empirical research designed to establish ground truths. In 2002, Gordon and Loeb published an economic model of optimal security investment (Gordon &

---

sequence of steps an adversary takes to achieve an objective so that defenders can intervene in the appropriate places to disrupt the adversary's operation.

[19] For example, see Risk Lens, which builds on the FAIR framework.

[20] Soo Hoo's model uses probability distributions over different consequences, which is a critical feature that the cyber risk community has largely not adopted. Again, it is remarkable that this insight was proposed 15 years ago, but has yet to become widespread. Significantly more data exist now (15 years later) so that the consequence distributions can be greatly improved, although Soo Hoo's work correctly emphasized the method and not the input distributions (which are trivially changed to reflect better data).

Loeb, 2002). The model centers on estimates of threats, vulnerabilities, and impacts and uses the assumption that security investments have a decreasing marginal return to calculate an upper bound on security investments. Gordon and Loeb found that decision makers should not invest more than 37% of the expected loss of a security breach.[21] The Gordon-Loeb model sparked many more papers aimed at quantifying cyber risk. Bojanc and Jerman-Blažič compare different metrics including net present value, return on investment, and internal rate of return (Bojanc & Jerman-Blažič, 2008). In 2013, the pair authored another paper with a more detailed and cost structure (Bojanc & Jerman-Blažič, 2013).

Many others have recognized the need for better cyber risk models. Thomas et al. developed a branching activity model that described different attack and impact scenarios (Thomas, Antkiewicz, Florer, Widup, & Woodyard, 2013). In this paper, the authors provide an overview of the difficulties of assessing cyber risk impacts, including disincentives for organizations to disclose data, massive uncertainty, and human biases. Thomas et al. also notes that the heavy-tailed nature of certain losses requires probabilistic methods to be used. Sonnenreich et al. also described challenges associated with heavy-tailed distributions even earlier, but their analysis of the return on security investment is limited by a lack of high quality data (Sonnenreich, Albanese, & Stout, 2005). The US Department of Defense uses another risk model called the Network Risk Assessment Tool (NRAT), which was developed to assess operational risk for DoD networks (Whiteman & Winter, 2013).

One of the most popular methods for assessing cyber risk in organizations uses Factor Analysis of Information Risk (FAIR) (Freund, & Jones, 2014). FAIR decomposes cyber risk by assessing the loss event frequency and the loss magnitude. Both of these factors can be divided further, resulting in manageable assessments that a decision maker or an analyst can perform more easily. Figure 5 shows how the different assessments are functionally combined into overall risk. Once each component is assessed, it is rolled up into the loss frequency and the loss magnitude, which combine into the overall risk. The FAIR method is now widely taught and is an excellent resource for modeling, assessing, and managing cyber risk.

[21] The 37% rule is a result of the assumption that security investments are effective at a marginally decreasing rate.

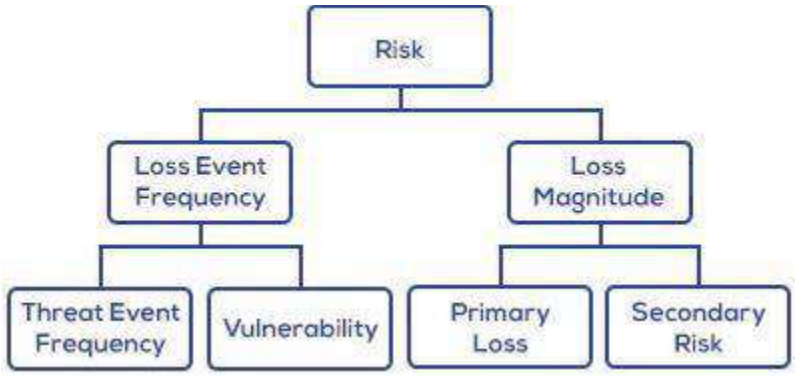


Figure 5: The FAIR approach to cyber risk. [22]

Much of the work discussed thus far occurred in the growing field of Economics of Information Security. Researchers in the field of decision analysis and risk analysis have largely not ventured into cyber security, with a few exceptions.[23] The Mission Oriented Risk and Decision Analysis (MORDA) tool was developed for the Department of Defense and uses attack trees, adversary models, and decision analysis to assess cyber risk (Buckshaw et al. 2005). More recently, Parnell et al. published a model for cyber security investment decisions for the US Air Force (Parnell, Butler, Wichmann, Tedeschi, & Merritt, 2015). The model uses probability trees to calculate the chance an attacker will be successful and the defender's ability to prevent, detect, and respond to attacks.

2.2 Qualitative Cyber Risk

Despite a growing body of literature on quantitative risk analysis in cyber systems, qualitative assessments are still the predominant method used in organizations. This is in part due to the cyber security standards published by the National Institute of Standards and Technology (NIST, 2012).[24] NIST has published many documents as part of a special publication series (SP-800) which range from very detailed and specific standards for encryption to general frameworks for risk management in cyber systems. These publications have acted as a critical resource for decision makers, but are an insufficient guide for analysts to conduct quantitative risk assessments.[25] The original version of the NIST Guide for Conducting Risk Assessments (SP 800-30) was published

[22] Taken from www.fairinstitute.org.
[23] INFORMS, the flagship conference for decision analysis and risk analysis, had only six papers with the word *cyber* in the title in 2014, out of roughly 5,000 papers.
[24] The International Organization for Standardization (ISO) also publishes cyber security standards, although for practical purposes they are very similar to NIST.
[25] It is important to note that NIST never intended for these documents to be used as a rigorous cyber risk guide. Further, NIST has been constrained by industry partners and other forces that have limited its ability to publish more detailed standards on cyber risk assessments. Thus, decision makers have largely been left without a guide for managing cyber risk.

in 2002 (Stoneburner, Goguen, & Feringa, 2002). While the value of the document was significant at the time, the document actually argued against quantification, going so far as to state:

> "Other impacts (e.g., loss of public confidence, loss of credibility, damage to an organization's interest) cannot be measured in specific units but can be qualified or described in terms of high, medium, and low impacts."

In 2012, a new version (Revision 1) was released to replace the 2002 version (NIST, 2012). Unfortunately, quantitative techniques are again given little discussion, and some of the risk assessment recommendations pass over relevant tools and techniques. While the document does not explicitly advocate for qualitative methods over quantitative methods, only qualitative and semi-quantitative examples are given. Further, there are issues with the examples given in the appendix (see figures 6 and 7).

TABLE G-3: ASSESSMENT SCALE – LIKELIHOOD OF THREAT EVENT OCCURRENCE (NON-ADVERSARIAL)

Qualitative Values	Semi-Quantitative Values		Description
Very High	96-100	10	Error, accident, or act of nature is **almost certain** to occur; or occurs **more than 100 times a year.**
High	80-95	8	Error, accident, or act of nature is **highly likely** to occur; or occurs **between 10-100 times a year.**
Moderate	21-79	5	Error, accident, or act of nature is **somewhat likely** to occur; or occurs **between 1-10 times a year.**
Low	5-20	2	Error, accident, or act of nature is **unlikely** to occur; or occurs **less than once a year,** but **more than once every 10 years.**
Very Low	0-4	0	Error, accident, or act of nature is **highly unlikely** to occur; or occurs **less than once every 10 years.**

Figure 6: NIST assessment scale. Taken from (NIST, 2012). Note the lack of a quantitative scale, even though probability would provide the most natural assessment of likelihood. Also note that the semi-quantitative scale decreases linearly, while the frequency decreases logarithmically. This subtlety is not highlighted, but implies that a decrease from 10 to 8 is worth much more than a decrease from 4 to 2. If an analyst fails to recognize this, poor decisions will be made.[26]

[26] While qualitative scales are generally considered ordinal, semi-quantitative scales can be cardinal or ordinal. Valuing a magnitude decrease in frequency linearly assumes that a decision maker would pay the same amount to reduce the frequency of a hack from 100 times per year to 10 times per year as they would pay to reduce the frequency of a hack from 10 times per year to 1 time per year. This is clearly not true in cyber security, since a decision maker would pay much more to eliminate 90 attacks compared to 9 attacks.

TABLE I-2: ASSESSMENT SCALE – LEVEL OF RISK (COMBINATION OF LIKELIHOOD AND IMPACT)

Likelihood (Threat Event Occurs and Results in Adverse Impact)	Level of Impact				
	Very Low	Low	Moderate	High	Very High
Very High	Very Low	Low	Moderate	High	Very High
High	Very Low	Low	Moderate	High	Very High
Moderate	Very Low	Low	Moderate	Moderate	High
Low	Very Low	Low	Low	Low	Moderate
Very Low	Very Low	Very Low	Very Low	Low	Low

Figure 7: NIST risk table. One of the most confusing examples in the document is the level of risk calculation from the likelihood and the impact. A table is constructed to combine qualitative assessments to determine the overall level of risk. For example, if the impact is low and the likelihood is very high, the overall risk is low. No discussion about this determination is included.

In 2014, NIST released a new document titled "A Framework for Improving Critical Infrastructure Cybersecurity" (NIST, 2014). The framework lists five functions for achieving cyber security outcomes; identify, protect, detect, respond, and recover. Each function has categories and sub-categories that specify controls, processes, and management that support the function. The document also discusses the risk analysis cycle. The NIST cybersecurity framework has been championed by the US government and other organizations. In February 2015 at the White House Cybersecurity Summit at Stanford University, the president of the United States addressed industry leaders, the government, academia, and the public about the importance of cyber security. At the summit, several business leaders announced the adoption of the NIST framework, including Apple, Bank of America, U.S. Bank, Pacific Gas & Electric, AIG, Walgreens, and Kaiser Permanente (White House, 2015).

The primary value of many of the NIST cyber risk documents is their ability to stress the importance of cyber risk in organizations. While there are shortcomings in the specific methods that are proposed in these documents, the framework is likely to add significant value by starting the conversation. Also, it is important to acknowledge that the NIST documents are constrained in their technical sophistication. Organizations would have difficulty adopting a framework with detailed mathematical formulations; thus an emphasis on simplicity is important. However, the NIST documents should contain a richer discussion about quantitative methods and reference the sophisticated methods that exist.[27]

[27] It is interesting to compare the NIST cyber security documents to NIST documents published on nuclear power plant safety. The nuclear power documents are significantly more rigorous and detailed, containing technical details and sophisticated modeling techniques. However, nuclear power risk analysis is much

Other qualitative frameworks for assessing risk in cyber systems exist as well. Katsikas presents a comprehensive review of risk management methods (Katsikas, 2009). COBIT (Control Objectives for Information and Related Technology) is a governance framework that is sometimes used in IT.[28] OCTAVE (Operationally Critical Threat, Asset, and Vulnerability Evaluation) is another method, developed with US-CERT (United States Computer Emergency Readiness Team) at Carnegie Mellon (Alberts & Dorofee, 2002).[29] All of these methods are limited in their applicability due to their qualitative nature.

2.3 Data Analytics of Cyber Security Incidents

Public data scarcity has been a major limiting factor in the general study of cyber risk. Organizations may have advanced techniques that are in-house, but these are not in the public record. However, empirical research is slowly accumulating, leading to important insights for decision makers and analysts (Moore & Anderson, 2011).

More academic research is needed in cyber security. Vendors routinely publish whitepapers with misleading or incorrect conclusions. In 2014, the Ponemon Institute published a report on the global cost of cyber crime containing errors and sometimes misleading results (Ponemon Institute, 2014).[30] Other publications present misleading conclusions, for example, using the mean to describe highly skewed data (Greisiger, 2013). Data are also routinely based on small sample sizes from voluntary surveys. Florêncio and Herley wrote an entertaining paper on the well-known issues with cyber crime surveys (Florêncio & Herley, 2013). Unfortunately, many researchers use these reports because other data are not available, which gives industry surveys much more credibility than they deserve (Ryan & Jefferson, 2003). A small number of industry-sponsored reports are useful (for example, Verizon's annual Data Breach Investigation Report).

Academic research is important to firmly establish ground truths and to prevent misconceptions of cyber security. McAfee, an antivirus software provider, has been heavily criticized for proposing that cyber crime costs the global economy $1 trillion per year, a number that has virtually no support (Maass & Rajagopalan, 2012).[31] Researchers debunked this estimate

---

more mature than cyber risk analysis. NIST cyber risk documents will likely evolve to become more sophisticated as the field develops.

[28] http://www.isaca.org/cobit/

[29] http://www.cert.org/resilience/products-services/octave/

[30] From page 8 "We also distinguish viruses from malware. Viruses reside on the endpoint and as yet have not infiltrated the network but malware has infiltrated the network. Malicious code attacks the application layer and includes SQL attack." This description is at odds with standard definitions of viruses and malware.

[31] The trillion-dollar figure was based on a single survey that was used to extrapolate to the global economy.

in a paper studying the cost of cyber crime (Anderson et al., 2013). Biener, Eling, and Wirfs studied cyber losses from an operational risk database and discussed implications for insurance markets, offering more data-driven estimates of actual cyber impacts (Biener, Eling, & Wirfs, 2015).

Another area of considerable debate is the effect of cyber breaches on an organization's reputation. Research shows differences in what is said about a data breach (for example on social media) and the actual stock price impact (Sinanaj, Muntermann, & Cziesla, 2015). Unsurprisingly, consumers can be very vocal about a data breach, but how this translates into material impact is uncertain. From a business perspective, reputation is valuable in order to secure future profits. Therefore, many researchers have used stock price as a proxy for reputation damage, and several papers have analyzed stock market returns after a data breach is announced (Campbell, Gordon, Loeb, & Zhou, 2003; Cavusoglu, Mishra, & Raghunathan, 2004; Kannan, Rees, & Sridhar, 2007). Research has consistently shown a very weak effect between data breaches and stock prices; for example, both Campbell et al. and Cavusoglu et al. found evidence of a decrease in stock price for two days after a breach announcement, but not longer.[32] Gordon, Loeb, and Zhou found evidence that the negative effect of data breaches is decreasing over time, suggesting that investors are becoming conditioned to data breaches (Gordon, Loeb, & Zhou, 2011). In fact, business interruption may be a much costlier risk to organizations than data breaches (Goldstein, Chernobai, & Benaroch, 2011).

Romanosky, Hoffman, and Acquisti study which data breaches are litigated and which are settled (Romanosky, Hoffman, & Acquisti, 2014). They find that only 4% of breaches involve litigation, but 50% of litigated breaches are settled. However, it is unclear whether these statistics are representative. Also, in some cases the threat of litigation may be enough to prompt a settlement, meaning legal risk is actually much higher for organizations. In cases where data is not available, experts may be consulted instead. While considerable evidence exists that expert probability elicitation can be faulty, it can also be very effective if biases can be controlled (Kahneman, 2011). Herrmann studied quantitative assessments of IT risk probabilities, demonstrating that IT professionals have unique biases, notably including pessimism (Herrmann, 2013). Data is still preferred in most cases if it exists.

Publicly available cyber databases are becoming more common, but currently exist only in narrow regimes. For example, data on botnet infections, spam emails, and malicious pieces of code are compiled and maintained by organizations and researchers. However, cyber security incident data is very limited. Privacy Rights Clearinghouse maintains an open, crowdsourced repository of

[32] Some inherent difficulties exist in this type of analysis, given how quickly the error grows. Still, it is interesting that evidence for a stronger link was not found.

data breaches.[33] These incidents are skewed toward events that receive media attention, but the database is still a useful resource. Several papers have analyzed the rate and severity of data breaches using this data source (Wheatley, Maillart, & Sornette, 2016; Maillart & Sornette, 2010; Edwards, Hofmeyr, & Forrest, 2015). VERIS (Vocabulary for Event Recording and Incident Sharing) is another well-documented framework for reporting incidents.[34] The VERIS data contains victim demographics, incident descriptions, discovery and response, impact assessment, and many other fields. One of the drawbacks to VERIS is that the large number of fields can create data fatigue for investigators who repeatedly enter the same information for low-impact incidents. Overall, however, these open-source databases are extremely valuable, providing a picture of the cyber security landscape for a range of companies. Very few studies have analyzed all cyber security incidents that have occurred at a single organization. Condon, He, and Cukier published an analysis of cyber security incidents at the University of Maryland in 2008, although the data consisted primarily of malware incidents (Condon, He, & Cukier, 2008). In this dissertation, incident data from the US Department of Energy (DOE) are analyzed, along with 60,000 cyber security incidents that occurred at another large, US-based organization (see Chapter 4).

Although cyber incident data is rarely made public, many organizations actually have high-quality data in-house that can be used to analyze cyber risk over time. Organizations often record cyber security incidents to track employee workload, satisfy auditors, fulfill reporting requirements, or analyze cyber risk. While in-house security incident databases are often neglected, they contain invaluable information that can be leveraged to assess the threats, vulnerabilities, and impacts of cyber attacks, providing a detailed view of cyber risk in an organization.

Cyber security documentation is common in large organizations. The Department of Defense uses the Joint Incident Management System, and private companies with government contracts report cyber incidents on federal information systems to US-CERT (Cyber Incident Handling Program, 2014; US-CERT).[35] NASA uses an Incident Management System that is run by its Security Operation Center (Martin, 2012; Martin, 2013). Additionally, several third-party software solutions are available for cyber security incident tracking.[36] Many standards and best practices exist that outline how to create and maintain incident tracking (Killcrece, Kossakowski, Ruefle, & Zajicek, 2003). Recently, a working group at the Department of Homeland Security (DHS) published guidelines for incident recording via an excellent whitepaper (DHS, 2015).

[33] See pricacyrights.org
[34] See veriscommunity.net
[35] See https://www.us-cert.gov/incident-notification-guidelines
[36] See RSA Archer, Demisto, and Resilient Systems (formally C03), now acquired by IBM.

Overall, incident management systems are poised to become a critical part of every organization's cyber security. Improved situational awareness, trend tracking, and risk management are all improved by recording cyber incidents.

2.4 Scenario Analysis

Probabilistic risk analysis and the analysis of classes of scenarios has been used to assess high-impact risk in other domains. Specific scenarios have also been analyzed in cyber. A research group from the University of Cambridge conducted a detailed analysis of an attack on the US power grid (Ruffle et al., 2015). The researchers studied how a malware strain could cause a power grid blackout and quantified the economic impact to businesses, the government, and insureds by constructing quantitative models informed by real-life incidents. This type of analysis enables decision makers to have high quality discussions around the assumptions and outputs of a scenario model.

Another challenge in modeling cyber security becomes how to integrate a scenario analysis with historical data. Fortunately, this problem has been addressed with several techniques across a number of domains. One formulation consists of parameterizing the probability assessments obtained from experts and treating it as a prior distribution, which can then be updated using the historical data (Karam and Planchet, 2015). A similar problem is encountered in climate science and financial modeling. In both, historical data exist, but may be insufficient because the underlying system is changing. In climate science, increased CO2 levels require combining historical data with detailed models of atmospheric chemistry and physics.

In this dissertation, scenario models are integrated with data-driven models by developing an overlap region of the risk curve. Historical incidents that have occurred rarely (but at a sufficient level so that they can be analyzed) can be used to benchmark the scenario model output and explore different possible futures.

2.5 Summary

The work in this dissertation builds on decades of other high-quality research. Many of the ideas have been proposed before, but the solutions were often ahead of their time. Often, public data did not exist or some managers had not yet recognized the importance of cyber security. The model presented in this dissertation is an aggregation of many ideas, methods, and models that have been published before. However, a detailed manual for conducting a rigorous quantitative risk analysis of cyber security is urgently needed and this dissertation seeks to fill that gap.

3 The Model

Cyber risk is made up of common, small impact incidents combined with rarer, more impactful incidents. The entire risk curve needs to be assessed so that decision makers can understand both types of risk; historical incidents that have already happened as well as new scenarios that may not have occurred. The assessment of these two types of incidents may require different modeling methods. For example, some organizations may have internal databases that can be used to assess the probability and impact of cyber incidents. In fact, some classes of cyber incidents can be shown to have been very stable in the past. For example, section 5.6 shows how website attacks at one large organization have historically occurred at an annual rate of about 50 per year. However, external events could change the frequency of website attacks meaning that this rate could change over time. Additionally, certain large-impact incidents may simply not have occurred yet because not enough time has passed. Therefore, statistics are not enough and historical data need to be augmented with additional modeling techniques. In the case of cyber security, analysts may use expert opinions or models of scenarios that have not occurred yet to come to a more complete assessment of the entire risk curve. Using this framework, cyber risk can be modeled in three regimes:

4. **Data-driven model**: derived from historical incidents if the data exist and the data are stable over time.
5. **Scenario-based model**: used to model incidents that have not yet occurred (typically large impact incidents).
6. **Overlap regime**: combines the data-driven model and the scenario-based model by overlapping the risk curves to avoid the double counting of incidents.

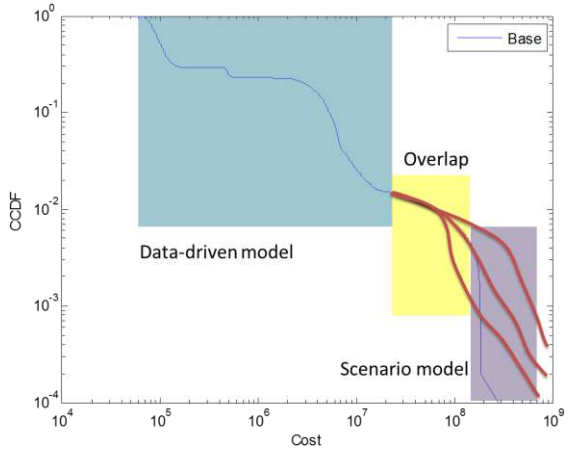


Figure 8: The three regimes of the risk curve; the data-driven model, the scenario-based model, and the overlap regime.

Once these three regimes are combined, a decision maker has a full characterization of the cyber risk facing an organization.

The rest of this chapter is structured as follows. First, a specific case of assessing the full risk curve at an organization is given by analyzing the fictitious organization, Space Corp. The adversaries, threats, and cost of cyber incidents at Space Corp are modeled by using the three regimes discussed above, namely the data-drive portion, rarer events modeled with scenarios, and an overlap region. To generate the full risk curve, a Monte Carlo simulation is used to obtain the risk curves and to compare the value of different security safeguards. The chapter concludes with a discussion of how to generalize the model to quantify risk at any organization.

3.1 Data-Driven Model

Some organizations may have historical cyber incident data in-house that have been recorded over time. While statistics alone are not enough to assess cyber risk, some insight may be obtained by analyzing these data and using them to inform a data-driven model of cyber risk. Many organizations have more data than they think. As discussed earlier, incident management systems that record cyber security incidents are excellent resources. Organizations that do not have a formal incident tracking system in place may still obtain data from other sources. The rate of lost devices is likely recorded by the procurement office, given that new purchases need to be made to replace lost equipment. Common malware infections are often recorded in a helpdesk ticketing system. The legal team at an organization often records data disclosures since these entail certain legal requirements. Information from penetration tests, third party audits, and the security team can provide data indicating the rate of vulnerabilities and patching. The combination of these information sources is often enough to obtain rough assessments of the rate and consequences of different attacks. Further, using probability distributions instead of point estimates allows a decision maker to incorporate uncertainty over these inputs.

Once data are obtained, they need to be analyzed in a way that enables the calculation of different risk curves at an organization. Choosing the appropriate level of detail for a cyber security model is very important. A large number of taxonomies and frameworks exist for cyber security that focus on attackers, defenders, vectors, exploits, or victims (Meyers, Powers, & Faissol, 2009; Kjaerland, 2006; Harrison & White, 2011; Simmons, Ellis, Shiva, Dasgupta, & Wu, 2009; Chapman, Leblanc, & Partington, 2011). The most important feature of a model is to choose a characterization that is useful to the decision maker. Certain organizations may perform a detailed analysis of their adversaries, while others will focus on entry points to their organizations. None of these choices is right or wrong; each is simply a different frame for modeling the problem.

In this dissertation, an incident-driven model is used to assess cyber security at an organization. The probability that different adversaries attempt to gain access and the likelihood of a vulnerability are encoded in the model, but the focus on incidents allows for a more intuitive assessment of different safeguards and the impact of cyber incidents.

3.1.1 Cyber Security Incidents

In this model, the monetary consequences of cyber security at an organization are calculated as the aggregation of the impacts of many individual cyber security incidents. Cyber security incidents are defined in this context as events that are documented by a security operations center analyst. Note that these incidents are not necessarily cyber attacks. Misplaced laptops and data spillage incidents are not caused by malicious actors, but are categorized as cyber security incidents. This definition also helps to clarify the ambiguous nature of port scans, which could stem from network reconnaissance by an adversary, or legitimate research projects that map the Internet.[37] Further, the cyber security incidents discussed here are not machine data in the form of logs, email filters, or intrusion detection system alerts. Organizations often state that their networks are attacked millions of times per day, but each "attack" is often a benign connection attempt that is blocked at the border. In order to make the analysis more meaningful, the model presented here is driven by cyber security incidents that involve some form of human interaction by the defender. Cyber security incidents are typically recorded in *incident management systems* (IMS). IMS can be formally designed systems like RSA Archer or Demisto, homegrown databases, or even incidents casually recorded in a person's email inbox.

Cyber security incidents typically involve a minimum investigation time of one hour. Incident management systems track events above a minimum impact level, making the analysis much more useful than noisy log data. For example, a lost device will be reported and an analyst will interview the employee to determine how the device was lost, ensure data was encrypted, or initiate a data disclosure notification process. Malicious email incidents are reported to a security analyst who deletes the email, adjusts email filters to prevent the same email from infecting other machines, and takes remediation actions in the form of removing malware from a user's device or resetting passwords.

The incidents recorded in IMS usually require some form of investigation or remediation. For example, a malicious email that is blocked by the email filter may not trigger an incident to be

[37] Port scans involve an actor sending a "ping" to an IP address on a certain port. The pinged device's response can often include identifying information, including the type of device, software version, and serial number. Port scans are used as network reconnaissance by adversaries, but also by academic researchers.

created, while an attempted website defacement might involve identifying the attacker's IP address and adding it to a blacklist. Determining which events are recorded as an incident is subjective, but over a long period of time the recorded incidents are an excellent indication of how the security operations center spends its time.

3.1.2 Adversaries

Organizations face many different threats, but often cannot distinguish among different adversaries. For example, it is difficult to tell if a port scan is from a nation state or a group of academic researchers. Other times, organizations may be able to infer information about an attacker. Unique malware that has never been observed in the wild is the hallmark of a targeted attacker with sophisticated resources, while a phishing email attempting to steal webmail credentials often comes from common spammers. Some organizations might find it useful to begin an analysis by listing possible attackers. For example, organizations engaged in government activity may be more likely to be attacked by hacktivists.

3.1.3 Attack Scenarios (Incident Categorizations)

In this dissertation, cyber security is modeled using an attack scenario (incident category) and impact-driven approach, primarily due to the form of the data used in the analysis. Adversaries, vulnerabilities, and exploits are modeled only in certain cases where a scenario analysis is useful. Attack vectors are used in a top-down modeling approach. A useful analogy to consider is home burglaries. A *vector* is a way that an adversary might get into the house, such as the front door, a window, the garage, or a chimney. For each vector, there are several *exploits* that an adversary might use. For example, a burglar attacking the front door could use a lock pick, a sledgehammer, or a deliveryman costume to compromise the home through the front door vector. In cyber systems, many attack vectors exist. An adversary might gain access by sending a malicious email to a user, attacking the organization's websites, compromising a user who visits a compromised website with an outdated web browser, or obtaining physical access.

Cyber security is a very complicated problem with many technical challenges. One of the reasons that quantitative cyber security models have not been adopted more prevalently thus far may be because the models are often too detailed. For example, organizations' websites are often under attack by adversaries trying to hack a server, obtain information, or deface the website. An analyst could quickly be overwhelmed by modeling the technical exploits against websites,

including directory traversal attacks[38], SQL injections[39], cross-site scripting[40], or brute force attempts[41]. This level of detail often works against the analyst, since uncertainty increases very quickly. For example, the rate of website attacks or compromises is generally known, but the rate of SQL injections versus directory traversals may not be. Using vectors simplifies the model, so detailed exploits need to be included only if they provide sufficient value. For example, in a home, some safeguards protect against one exploit (e.g., reinforcing the door protects against a sledgehammer, but not a deliveryman), while others might reduce several (e.g., posting an armed guard could eliminate both lock picking and a sledgehammer attack). Further, the modeling detail may be totally unnecessary in certain situations. If a safeguard involves moving websites to a segmented part of the network that isolates lateral movement (a DMZ[42]), the impact of all website exploits is limited. Considering individual exploits is important only if the safeguards affect exploits; if the safeguard affects the vector, then the additional modeling detail is irrelevant.

Figure 9 shows the different vectors used in this analysis. Note that the model includes incidents that are not attacks, but that generally fall under the responsibility of a security operations center (e.g., data spillage).

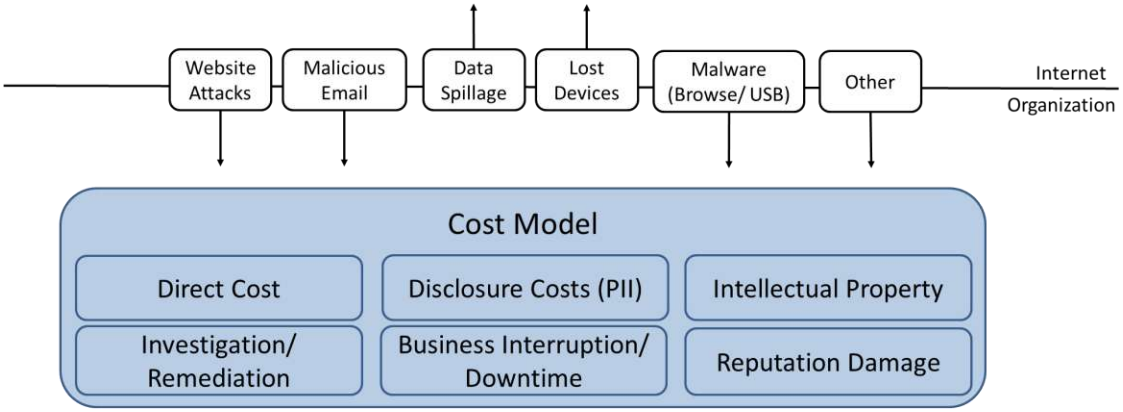


Figure 9: Cyber security incident categories. Adversaries may gain access to a network through website attacks, malicious emails, malware incidents, and other attack vectors. Data spillage and lost devices may release data into the wild, although attackers do not gain access to an organization through these vectors. Each vector can cause losses via six cost types, which are each modeled using historical data and expert opinion.

In this dissertation, six attack vectors are modeled:

[38] Directory traversals involve bypassing the intended file hierarchy structure of a system to gain unauthorized access.
[39] SQL injections use specially crafted queries for a database to obtain unauthorized access or information.
[40] Cross-site scripting involves injecting malicious code into a web application.
[41] Brute force involves trying all password combinations.
[42] DMZ stands for de-militarized zone, and is a segmented part of a network designed to limit lateral movement through a network by an adversary.

Data spillage:[43] Incidents that could possibly disclose information to unauthorized individuals are categorized as data spillage. For example, an employee could forget to encrypt an email that contains social security numbers.

Malicious emails: Any intrusion or attempted attack that originates through email is classified as an email incident. For example, criminals may try to extract a user's email credentials to use their account for sending spam ("phishing"), or attach malicious files to an email to infect a user's machine with malware.

Lost or stolen devices: Laptops, tablets, phones, and other hardware can be lost or stolen. These incidents typically require different levels of investigation depending on the type of device and the encryption level.

Website incidents: Any attack that exploits websites operated by the organization is classified as a website incident, including website defacements, SQL injections, and server compromises.

Web browsing and USB incidents: Malware that does not originate via email or through a website is categorized as a web browsing/USB incident. For example, a user may inadvertently download malware while visiting a compromised website. Users may also spread malware via USB devices.

Other: While other types of incidents take place, many do not occur with enough volume to merit the creation of a specific category. For example, denial of service attacks and insider attacks occurred very rarely at the organization studied here. False alarms and near misses also occur at a low rate. Incidents classified as tasks are also relatively low-volume. Examples of tasks include documenting network exercises, wide-scale patching, or investigations (such as pulling log files) meant to aid an audit or an inquiry (e.g., pulling an employee's emails after allegations of harassment). All of these incidents are bundled together into an "other" category. This classification limits the conclusions that may be drawn from an analysis of this category, but incident subtypes can be extracted if more detail is required.

Note that the model above is not exhaustive. Several other vectors could be identified, such as physical access or supply chain attacks. However, in many organizations, the vectors above are

[43] Data spillage incidents are sometimes categorized as "insider attacks" by other sources. The term *insider attacks* is not used because the majority of incidents are accidental and not malicious.

responsible for over 95% of past security incidents. In subsequent modeling iterations, more vectors can be included.

3.1.4 Impacts

Measuring impacts of cyber security incidents has been a major challenge historically. In this model, I distinguish between the incident severity (measured in man-hours of investigation) and the monetary costs of an incident. The dataset presented here has a severity measure for each incident, describing the number of hours spent on incident investigation and remediation. For example, a malware infection investigation would involve a security operation center analyst identifying the infected machine, imaging the device, removing the malware, and restoring any backups. Website attacks can involve much more time since investigators need to analyze logs to identify how an attacker gained access to a system and ensure that all malicious code has been removed.

In this model, indicator functions are used so that incidents may accrue additional monetary costs if a certain investigation time threshold is exceeded. This assumes that incidents with a low severity do not cause significant monetary costs, which is consistent with observations of the organization's operation. While it is possible that a simple malware strain that takes five hours to investigate could lead to millions of dollars of losses, this scenario does not occur in the data. Incidents that cause monetary damage tend to receive more investigation, sometimes for legal or auditing documentation or out of a general consensus that costly incidents deserve more attention.

The monetary impacts in the organization are assessed via six cost types. Historically, the cost to an organization has been assessed using the confidentiality, integrity, and availability (CIA) of information. However, translating CIA into a monetary cost is challenging. Modeling the monetary costs up front aids in communication and allows for comparisons between cyber risk and other types of risk, for example business risk or liability. The six types of monetary losses used here are:

Investigation Cost: Each security incident requires time to investigate and remediate. For example, lost laptops investigations involve the victim being interviewed to determine the circumstances surrounding the loss. The investigator will also determine the state of the laptop (sleep mode, hibernation mode, etc.) and confirm whether any personally identifiable or sensitive information was on the machine. Data spillage incidents involve time spent documenting how information was spilled and removing spilled data from unauthorized machines. While sometimes overlooked, the investigation of security incidents at an organization is a major driver of the overall cost. One CISO

at a large tech firm expressed difficulty in communicating to upper management that "even small incidents cost money."

Direct Cost: Any hardware replacement or direct monetary losses are classified as direct costs. For example, lost devices need to be replaced when they are stolen. Replacement of other types of hardware, for example generators or manufacturing equipment, may become more common in the future as cyber attacks have more frequent physical impacts (Zetter, 2015). Direct money transfers due to fraud and ransomware are also included here.[44]

Business Interruption: In certain situations, cyber incidents may lead to the disruption of operations or unavailability of services. Internal productivity may also suffer such as in the case where a stolen laptop leaves an employee unproductive until a new machine can be obtained. Denial of service attacks can prevent banking customers from accessing their accounts, or shoppers from buying merchandise online. It is important to note that some revenue may be recovered if a system is down for a short period of time, while other businesses operate on much smaller tolerances (for example, booking a hotel room at a different website if the first website is down).

Reputation Damage: Cyber breaches affect the reputation of an organization. They may cause a short-term decline in the stock price of a publicly traded company (Cavusoglu, Mishra, & Raghunathan, 2004), reduce a customer base, or reduce the organization's ability to secure future business deals. Laptop losses that release personally identifiable and sensitive information have periodically received significant news attention, although it is difficult to connect the bad press to a specific economic loss.

Credit Monitoring/Breach Notification: Credit monitoring and breach notification is becoming a common cost for organizations. When personally identifiable information is stolen, the organization is often required by law to make a disclosure. Organizations often offer free credit monitoring as an act of good faith, although the effectiveness of this service has been questioned (Krebs, 2015c). Costs can accumulate quickly based on the huge volume of individuals that are impacted (for example, a breach of 50 million records would generate a cost of almost $25 million in postage alone).

[44] For example, an adversary could impersonate the CFO and request that a finance employee transfer a large amount of money to an offshore account.

Loss of Intellectual Property: Cyber attacks can sometimes lead to the loss of intellectual property (IP), trade secrets, and other sensitive information. The value of this information is often unclear and most cyber insurance policies will not cover losses stemming from IP or trade secret disclosures. The complex interaction of market forces makes quantifying IP losses to an organization very difficult. Microsoft routinely has its software pirated, but still maintains a solid market share in China (Mozur & Wingfield, 2016). On the other hand, US solar panel manufacturer SolarWind has accused China of stealing trade secrets for manufacturing solar panels and has publicly expressed concerns that this information will allow Chinese companies to undercut their product's competitiveness (Cardwell, 2014). The valuation of IP is difficult, but some methods exist (Andrijcic & Horowitz, 2006).

Organizations that have unique costs due to cyber attacks may choose to model more impacts, which simply involves creating a new impact model. For example, another category that quantifies the impact of losing strategic national secrets could be added. The loss of the blueprints of a new aircraft fighter would impact an organization via this cost vector. The impact of the loss of national secrets is outside the scope of this dissertation.

For each of these impact types, a sub-model is created to calculate the corresponding impact of an incident. Every incident always accrues cost due to time spent investigating the incident. Other impacts (reputation damage, business interruption, etc.) may occur depending on the type of incident and the severity (meaning the total hours of investigation).

3.1.5 Monte Carlo Simulation

Once the distribution of different attack types, their frequencies, and their impacts are modeled, the data-driven risk curve can be calculated using a Monte Carlo simulation. Note that the distributions of the inputs may come from historical data (in the case of the data-driven regime) or scenarios (in the case of larger incidents). Therefore, the incidents that are simulated may have occurred in the past, but can also be new scenarios that have not yet been observed. In other words, the Monte Carlo model generalizes into a simulation method even for incidents that have not yet occurred. Incidents are simulated at the rate specified and each incident has a corresponding impact drawn from a distribution. The cost due to each incident is calculated in dollar terms. The model simulates one year of cyber incidents, although this can be easily changed. At the end of the simulation, all of the costs are summed to obtain a total yearly cost. From a large number of runs (n= 10,000), the complementary cumulative distribution function is calculated.

More formally, define the following terms to use in the simulation:

Symbol	Meaning
C	Cost
i	Denoting an incident
j	Denoting the type of incident (website, email, malware, etc.)
H	Hours
V	Indicator function
P_i	Probability of loss for incident type i
DC_i	Direct Cost for incident type i
PI_i	Privacy Information loss for incident of type i
RD_i	Reputation Damage loss for incident of type i
IP_i	Intellectual Property loss for incident of type i
BI_i	Business Interruption loss for incident of type i

The total cost of cyber incidents at an organization over a year is simply the sum of the cost of each incident that occurs.

$$Yearly\ Cost = \sum_i C_i$$

The total cost of each incident is obtained by summing over each impact category, namely investigation costs, direct costs, business interruption, reputation damage, credit monitoring, and loss of intellectual property. The assumption is made that each hour of investigation costs $100, so the number of hours spent investigating the incident is multiplied by 100. In this model, the simplifying assumption is made that the costs are conditionally irrelevant given the hours of investigation. Certain situations could occur where this assumption breaks down, such as the case where an intellectual property loss increases the probability of reputation damage. The total cost is given as

$$C_i = H_i * \$100 + DC_i + BI_i + RD_i + PI_i + IP_i$$

The impact categories (business interruption, reputation damage, etc.) can be conditioned on the type j of incident (email, website, data spillage, etc.). The cost of an incident i of type j is a function of the hours of investigation of incident i, impact distributions, indicator functions, and conditional probabilities. For example, the cost of an incident i of type ‘data spillage’ can be written as

$$C_i = H_i * \$100 + RD_i + PI_i$$

Only investigation time, reputation damage, and privacy information disclosures can occur from a data spillage incident (see Chapter 5.4.2 for more detail). The costs can further be broken down as

$$C_i = H_i * \$100 + V(H_i) * (RD_i + PI_i)$$

$$V(H_i) = \{\begin{matrix} 1 & if\ H_i \geq 500 \\ 0 & else \end{matrix}\}$$

$$PI_i = \{\ Uniform(60k\ to\ 5M)\ \}$$

$$RD_i = \{\begin{matrix} Uniform(1M\ to\ 2M) & with\ probability\ 0.45 \\ Beta\ Dist\ (100M\ to\ 160) & with\ probability\ 0.05 \\ 0 & with\ probability\ 0.5 \end{matrix}\}$$

In other words, reputation damage and privacy information costs only occur if the hours of investigation exceed 500, hence the indicator function $V(H_i)$. If this occurs, a random variable is drawn from a distribution to obtain the loss level of the incident. Privacy information losses are drawn from a uniform distribution, while reputation damage is drawn from a piecewise distribution representing three outcome types (see Chapter 5 for more detail). Once the cost for each incident is drawn from the distributions and calculated, the risk curve can be obtained by simulating many times and forming the cumulative distribution function, or the complementary cumulative distribution in this case.

Figure 10 shows a decision diagram of cyber security at an organization, and how the rate, impact, and incident type are relevant to the total cost. Using an incident based model is convenient because some uncertainties become conditionally irrelevant to the cost given information about the rate, type, and impacts of attacks. For example, analysts can build out inferences about the attacker, but these data are not required to calculate the return on security investment of safeguards. The Monte Carlo is used to simulate the decision diagram shown in Figure 10.

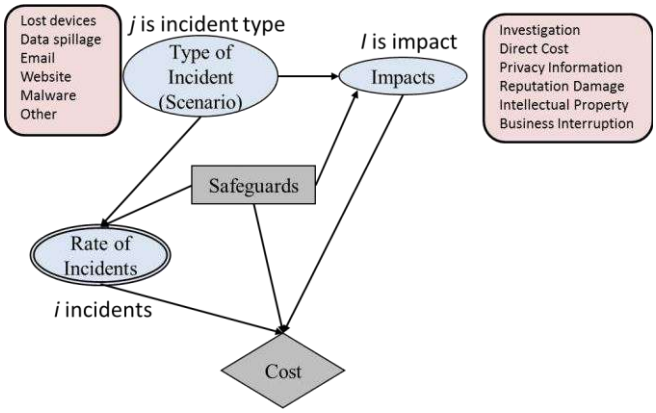


Figure 10: A Monte Carlo simulation is used to calculate the risk curves.

3.1.6 Modeling Choices

It is important to emphasize that the data-driven model is not limited to using the historical rate or impact data. In fact, artifacts often exist in the data that need to be identified, cleaned, and removed. For example, one organization observed a large spike in lost devices during one month, which skewed the expected number of lost devices for the next year. After further investigation, it was determined that the spike in lost devices was due to an audit that identified a large number of devices that had gone missing over an extended period of time but had all been recorded as being lost simultaneously. Given this information, the analyst was able to remove these outliers and produce a more accurate assessment of the rate of lost devices. Similarly, the same organization observed a consistent decrease in the number of malware incidents year after year. After discussing this with the CISO, the analyst modeled a 10% decrease in malware incidents. The analyst also included ransomware attacks, although very few of these attacks had ever been recorded at the organization. Based on industry reports and conversations with other companies, the CISO believed that ransomware attacks were on the rise and would occur more frequently in the next year.

The incorporation of outside knowledge may be subjective, but can still be very rigorous. Bayesian models can be constructed to carefully encode new threats, new signals, and evidence into a simulation. Decision makers are left with a wide range of alternatives for data sources, allowing the construction of a detailed model that is informed by both statistical data and subjective information sources.

Much of our ability to model cyber security is a direct result of the stability of cyber security incidents, which generally (in the datasets analyzed here) occur at a constant rate over time and have a severity distribution that evolves on a timescale of years (see Chapter 5 for more discussion). Because of this, many aspects of cyber systems can be modeled fairly easily. For example, simple statistical tools can be used to verify that the rate of data spillage incidents at one organization have been fairly stable over the past six years. It is certainly possible for this to change

(for example, something underlying about the system could cause an increase), but starting from the base rate of historical incidents may be appropriate for certain types of incidents.

Once data are identified, they need to be incorporated into the model. There are two general approaches that an analyst could use. Analysts can either use historical data to obtain a parameterized distribution or sample directly from historical incidents (bootstrapping). Each technique has benefits and disadvantages, but both can be useful depending on the specifics of the model. In the model presented here, I use a combination of parametric and bootstrapping approaches. For example, the rate of incidents is analyzed parametrically while incident severity is generally simulated via bootstrapping.

Parameterize and simulate approach

One useful technique for obtaining probabilistic inputs involves fitting the data with a mathematical function, which is used as a probabilistic input to the model. For example, the distribution of hours of investigation of certain vectors is well modeled as a power law distribution. Therefore, an analyst can obtain the best fit and draw from that distribution. Similarly, the rate of incidents can be modeled as a Poisson process and be characterized by a single parameter (lambda).

The *parameterize and simulate* approach is useful because it allows a decision maker to model different security safeguards at the process level. For example, the rate of malicious email incidents may decrease with a better email filtering system, thereby adjusting the rate of incidents (given by λ) to $\lambda(1 - r)$ where r is the additional proportion of incidents that are detected. Similarly, the severity distribution may change from a power law distribution with an α of 1.2 to an α of 1.4 if the organization implements improved detection capabilities. It also allows an analyst to sample from incidents that are bigger than the largest incidents that have been observed. This is useful to supplement the scenario analysis that is used to study large or rare incidents that have not yet occurred.

The parameterization works very well if the phenomenon is relatively simple, but can break down in certain cases. For example, figure 11c shows the distribution of investigation time for malware incidents. Note that a power law is not a good fit for the data, and different regimes seem to exist. In this case, a parametric form of investigation time is not easy to obtain. Even for cases where the data are highly patterned, several different regimes exist. There may be a regime for small incidents, large incidents (modeled as a power law), and outliers.

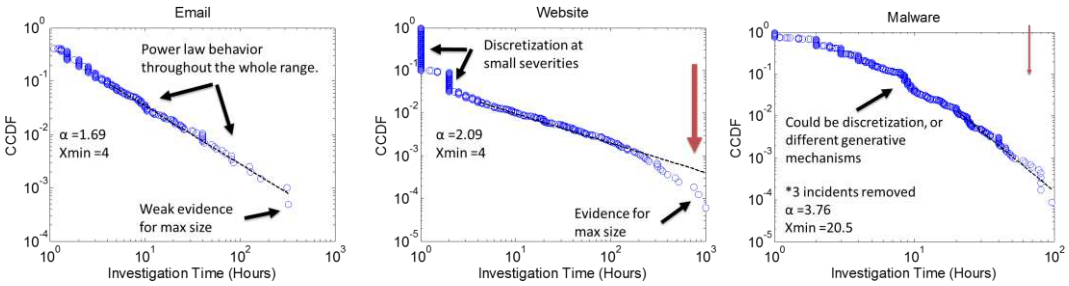


Figure 11: Curve fitting cyber security data. Note that certain incident types follow a clear trend (a), while some incident types require more complicated modeling (b). Other data may result in a poor fit because of multiple generative mechanisms that cannot be distinguished (c). Graphs (a) and (c) have outliers removed to improve the overall model fits.

Overall, parameterizing the inputs gives a smoother sample distribution, allows for more sophisticated safeguard models, and allows extrapolation into incidents that have not yet occurred, but can also significantly complicate the model.

Bootstrapping approach

A second approach involves bootstrapping, or directly sampling from historical incidents. This method is considerably easier to use because no curve fitting is required. The historical data can still be adjusted using a simple rule (e.g., the impact is reduced by a factor of 2), or using a "what if?" analysis.

A downside to the bootstrapping approach is that some discretization effects can take place, especially with high severity incidents. Only incidents that have occurred historically can be sampled in the simulation, and severe incidents are often sparse at high severities.

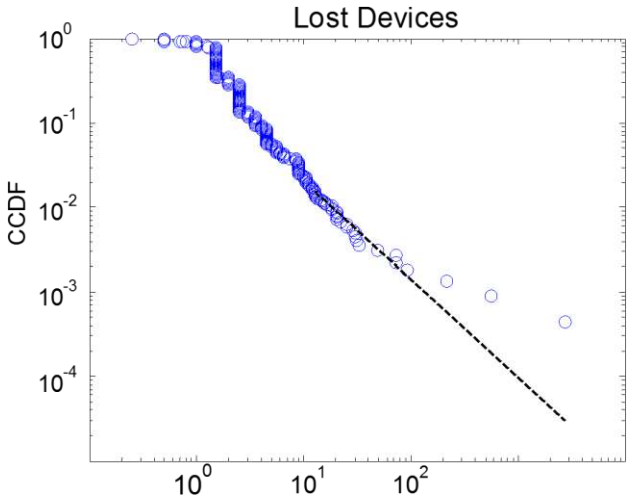


Figure 12: Bootstrapping cyber incident data. Bootstrapping avoids more complicated models caused by a parametric approach, but results in sparse samples of large-severity incidents.

Overall, bootstrapping is a very simple method and may be more accurate if the investigation distribution is complicated. However, it is limited in the sense that it can consider only historically observed incidents.

Other modeling choices

The incorporation of the cost model is another important task. Costs can simply accrue proportionally to the hours of investigation, or include an indicator function to include other factors. Costs can thus be coupled or decoupled, in the sense that they could be restricted to include other factors like reputation damage and business interruption costs, or include only one or the other.

In either case, it makes sense to separate incidents into a low regime where investigation time is the sole monetary cost, and a high regime where additional monetary costs may occur. It is desirable to have additional costs correspond to high investigation times, because this is what is empirically observed. It would be unexpected to have an incident require only 10 hours to investigate but cost millions. Therefore, an indicator function is preferable. For example, a decision maker could choose to include reputation damage with an indicator function $\boldsymbol{I}$ as a function of the investigation time $\boldsymbol{V}$ such that

$$I(V) = \begin{cases} 1 & if\ V \geq 200 \\ 0 & else \end{cases},$$

meaning that reputation damage occurs only if the investigation time reaches 200 hours.

3.2 Scenario-Based Model

Historical data are not sufficient to assess the entire risk curve because certain incidents may not have occurred yet, either because not enough time has passed or because cyber security is an evolving domain with an adaptive adversary. In these cases, different classes of scenarios can be modeled and assessed using expert probability elicitations, engineering data, near misses, and other sources of information. The purpose of this section is to describe how the rate and consequences can be derived for each scenario, which are then used as inputs to the Monte Carlo simulation to obtain the scenario-based regime of the overall risk curve. In some cases, the inputs may be entirely data driven (akin to actuarial tables for car insurance where many statistics exist) while scenarios are solely used in other cases where no data exist. The details of the scenario model also are dependent on what scenario is considered and different modeling tools may be substituted in depending on how the scenario is assessed. In this section, a scenario model is presented to address

the risk of nation state actors infiltrating an organization and exfiltrating intellectual property or disrupting operations.

3.2.1 Scenario Model Outline

Each scenario incurs an effect on the organization. These effects fall into the same impact categories as described in the data-driven model, but may require additional special treatment. Some of the scenarios could fall outside of the normal cyber effects, meaning that a careful consideration of severe impacts should be conducted. Another important consideration is the time-based nature of the discovery of the intrusion. Many network intrusions are not discovered until weeks or months after the adversary has gained access into the system. This well-known statistic highlights the importance of quick detection since organizations can sometimes take quick action that limits the severity of an incident.

The overall model structure for scenarios consists of three parts; a probability, loss function, and a discounting rate based on detection time. The probability is calculated by modeling each scenario as a sequence of discrete steps that need to occur for an adversary to achieve some objective. Each scenario begins with an initiating event such as a malicious email, a website attack, or a click on a malicious webpage. The attack may progress, becoming more severe if the defender is unable to block, detect, or remediate the malicious action. Each step in the attack progression has conditional probabilities associated with the attacker's path which could include success or detection. Therefore, the probability of each scenario *i* can be written as:

$$p(outcome_i) = P(IE_i) * P(A_i \mid IE_i) * P(B_i \mid A_i, IE_i) \ldots * P(Z_i \mid A_i, B_i, \ldots Y_i, IE_i)$$

Where IE_i is the initiating event of scenario i, and A_i, B_i, ... Z_i denote additional steps that an adversary needs to accomplish to achieve its objective.

The loss function of each scenario is the same form as the loss function used in the Monte Carlo simulation of the data-driven model and consists of investigation costs, direct costs, business interruption, reputation damage, credit monitoring, and loss of intellectual property. Using the same notation as in section 3.1.5,

$$C_i = H_i * \$100 + DC_i + BI_i + RD_i + PI_i + IP_i$$

Here, the cost functions are again assumed to be conditionally independent given the hours of investigation, although in certain situations, there could be dependencies. For example, business disruption could increase the probability of reputation damage.

The detection of the adversary also needs to be modeled, since quick detection can limit the impact of certain scenarios. To this end, we introduce a detection coefficient function $D_i(t)$ given as

$$D_i(t) = (1 - e^{-t})$$

The detection coefficient denotes when an attack is detected and is multiplied by the cost of an incident to reduce its impact. If it is detected immediately, then the costs become 0 since $D_i(t = 0) = 0$. If the attack is not detected for years, then the full cost occurs since $D_i(t = \infty) = 1$. In practice, more complex functions could be used for the detection coefficient, but the exponential argument demonstrates the method for this case.

Combining the loss function and the detection coefficient results in the general equation for the losses associated with scenario i.

$$C_i(t) = (H_i * \$100 + DC_i + BI_i + RD_i + PI_i + IP_i) * D_i(t)$$

At this point, the probability and cost of scenario i can be input into the Monte Carlo simulation to obtain the scenario-based regime of the risk curve.

3.2.2 Scenario Model Illustration

Chapter 5 describes a case study for Space Corp, a large research and development facility. Space Corp is concerned with several attack scenarios, including a nation state compromising its network and either obtaining intellectual property (which has occurred before) or sabotaging the space communications network resulting in a system wide failure of spacecraft a spacecraft (which has not occurred before).

First, the system needs to be modeled. Based on expert assessments, nation state attacks are most likely to originate via an email attack or a network attack. Other initiating events could be modeled separately (i.e. supply chain attacks or malicious insider attacks), and some initiating events cannot lead to certain scenario outcomes (e.g. lost devices cannot lead to ransomware incidents). After the initiating events are defined, the system needs to be modeled to determine the sequence of steps that an adversary needs to accomplish to obtain their objective. This includes modeling the system safeguards that in are in place. For example, Space Corp utilizes a segmented network infrastructure to guard against network penetration. Websites are isolated on a DMZ (demilitarized zone) which is functionally separated from the rest of the network. If a website is compromised, lateral movement into the core network is much more difficult.

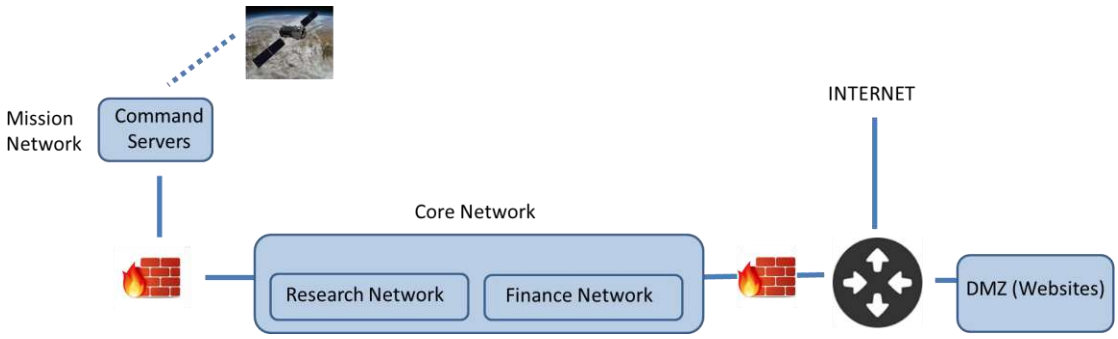


Figure 13: A high level schematic of the Space Corp network.

Once an adversary gains access to the core network, they may either look for intellectual property, or attempt to gain access to the mission infrastructure. Lateral movement occurs through a process whereby the adversary attempts to learn about the internal network structure, including how devices are connected and which have exploitable vulnerabilities. An adversary may conduct a scan to see what devices are exposed, or query directories to get a list of accessible files. These actions involve a risk of detection, depending on the sophistication of the defender. The adversary may face stagnation as well. In other words, it is possible that the adversary is unable to proceed deeper into the system, but has not yet been detected. In these cases, they may continue to persist in the network and reattempt an intrusion at a future date, or they may simply abandon their hacking attempt.

While sensitive data may reside on the core network, it is much harder to gain access to the communications servers that connect to the spacecraft. Credentials would likely have to be stolen for the adversary to be able to gain access to this part of the network. Once on these machines, the adversary would have to establish an uplink to the spacecraft, issue commands, and have those commands successfully received and acted upon. Most spaceflight operations use tailored command dictionaries to communicate with a spacecraft. Specific rules must be followed for commands to be correctly transmitted. These rules are not published and are considered extremely sensitive, meaning that naïve attempts at satellite communications would likely fail. If an adversary had previously been able to either eavesdrop on communications or obtain a copy of the command dictionary, then the attack would have a higher chance at success.

The steps required to gain access to the Space Corp network and either exfiltrate intellectual property or establish a connection with a spacecraft are modeled using a sequence of steps with different probabilities. These probabilities come from data and expert opinion, and take into account prevention and detection capabilities. For example, a system that monitors connections to the mission network can detect attempts at lateral movement and appear in the model by reducing the conditional probability of mission network access given admin credentials have been obtained. In some cases, similar historical events may have occurred, or certain ‘near misses’ can be used to

learn about the likelihood of success. Near misses are particularly useful and involve historical incidents where some but not all of the steps were achieved.

Figure 14 shows the series of steps that an adversary must complete to achieve different outcomes. The probabilities shown in the event tree are realistic numbers, but do not come from expert sources or data. Note that this attack sequence presents several opportunities to validate the model. For example, the calculated rate of the intermediate step of a user or website compromise can be compared to historical data. These observations can be used to calibrate the overall attack probability, since several steps along the way have been achieved, even if the final steps have not. In other words, the output of this model gives not just the probability that intellectual property theft or a satellite network failure occurs, but also gives the probability of near miss incidents, like the probability that mission network access is obtained per unit time (in this case, per week).

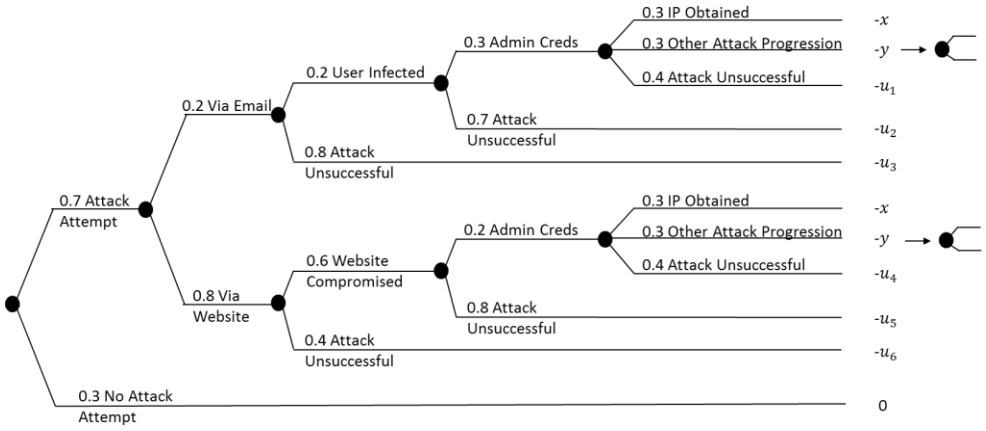


Figure 14: Attack sequence for obtaining intellectual property from Space Corp.

Based on Figure 14, the probability of each scenario can be calculated. For example, the probability of intellectual property theft is

$$P(IP\ Obtained) = P(Attack\ Attempt) * [P(Via\ Email) * P(User\ Infected) * P(Admin\ Creds \mid Email) * P(IP\ Obtained\ ) + P(Via\ Website) * P(Website\ Compromised) * P(Admin\ Creds \mid Website) * P(IP\ Obtained)] = 0.0227$$

The probability of IP theft per week is 0.0227 per week, which roughly translates to one IP theft incident per year. This is a slightly higher frequency than what has been observed historically at Space Corp (roughly one IP theft incident every six years). Figure 15 shows an influence diagram for calculating the probability of IP theft.

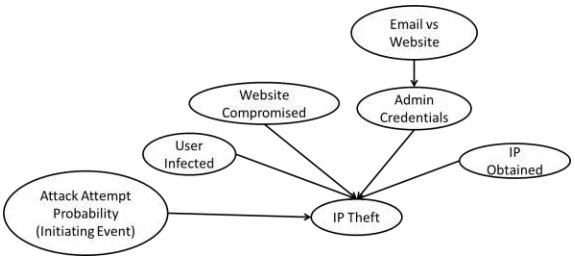


Figure 15: An influence diagram for scenarios at Space Corp.

Other scenarios could be modeled as well. Figure 16 shows the attack progression that could evolve if an adversary continues lateral movement after obtaining IP while not being detected.

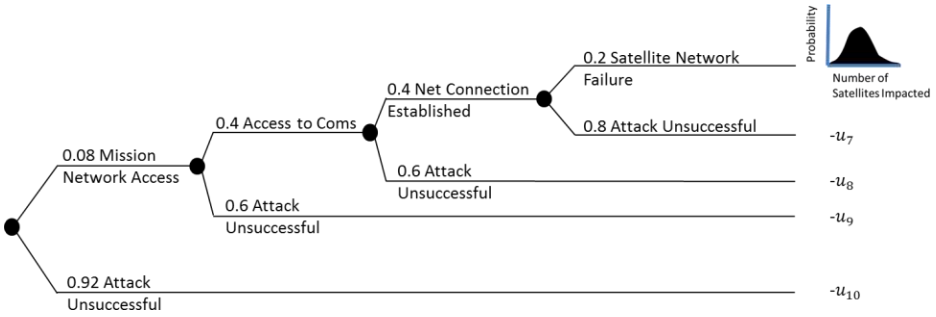


Figure 16: Attack sequence for a satellite network failure attack

The probability of satellite network failure is calculated as 0.000058, or roughly once every three hundred years. Analyzing the attack progression, it is clear that the crux step occurs when moving to the mission network, given that admin credentials have been obtained. This low probability corresponds to the security controls that Space Corp has in place, which make it difficult to move from one network to the other. Additional security monitoring (a system safeguard) put in place at the mission network entry point could further reduce the probability of a successful attack, meaning that the attack progressions rely on both the adversary's skill and the system's responses.

New information could change these conditional probabilities, leading to a higher probability of a scenario. For example, if several near misses occur where adversaries gain access to the mission network after compromising admin credentials, then the model inputs would need to be updated to reflect this new information.

Other scenarios may be relevant to different organizations. Concerns exist about the increasingly complicated and vulnerable supply chains that could be used to inject a compromised piece of equipment into an organization. Insider attacks could also be analyzed. These tend to be rare, but potentially very impactful at an organization. A treatment of insider attacks involves considering the vetting process for hiring new employees, what those employees have access to, and oversight mechanisms in the organization that could detect suspicious activity. Many organizations have to balance oversight safeguards due to privacy and trust concerns. Further, some

members of the organization inevitably need to be given more privilege and access than others, meaning that they may hold more keys that can be abused.

3.3 Combining the Data-Driven Model with the Scenario-Based Model

The basic process for combining the data driven model with the scenario based model involves overlapping both of the impact curves in an intermediate region. This essentially calibrates the scenario based model by linking the frequency of incidents that have been detected at an organization with the output of the scenario analysis, and provides a benchmark for the large impact incidents. In the overlapping region, the models can be adjusted to obtain the best fit.

In the analysis presented here, only one scenario (intellectual property theft) is analyzed. If multiple scenarios are considered, then the dependencies between them need to be modeled. If an adversary is able to obtain intellectual property, the probability may be higher that the adversary can access other parts of the network and therefore cause even greater damages.

Intellectual property theft is an interesting case for Space Corp, given that historical cases of this scenario exist. These incidents constitute some of the largest attacks that have occurred in the past. Further, many incidents have occurred historically where adversaries began the attack progression but were detected and stopped before it was complete (near misses). These historical incidents can act as benchmarks to assess the accuracy of the scenario model. Space Corp is also concerned with catastrophic intellectual property theft incidents that have not occurred yet.

Figure 17 shows the risk curves for malicious email attacks, and a notional illustration of the risk derived from the intellectual property theft scenario. Note that the scenario model curve shows a higher risk in the tail because the historical data do not have catastrophic intellectual property theft incidents that are addressed by the scenario model. Also note that the scenario model underestimates the risk curve in the overlap region, because the email risk curve includes other incidents besides those targeting intellectual property. The challenge therefore becomes how to balance these forces to get a good overlap.

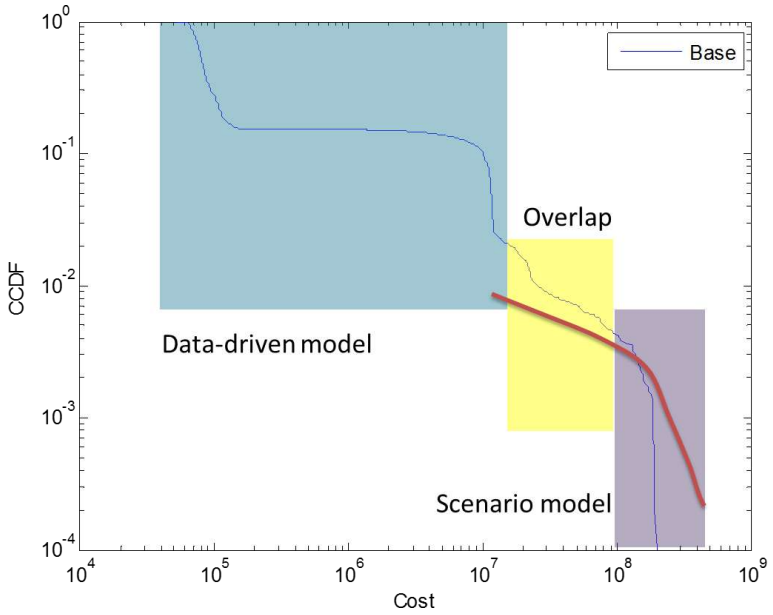


Figure 17: The overlap method to combine the two model types.

One method for overlapping these models involves carefully separating historical IP theft incidents and comparing only those occurrences to the scenario model. In that case, the two curves can be properly overlapped and a least squares calculation can be used to obtain the best fit. This may involve shifting the scenario model curve, which can be interpreted as applying a probability correction term that normalizes the probability of large impact incidents.

Once the two models overlap, an arbitrary cutoff point can be chosen to transition between the two models. A small change in the transition point will not materially affect the risk distribution, since a small change in the transition point will result in the same behavior, given that the models overlap. Figure 18 shows the final risk curve for email incidents at Space Corp, using the notional scenario model discussed above.

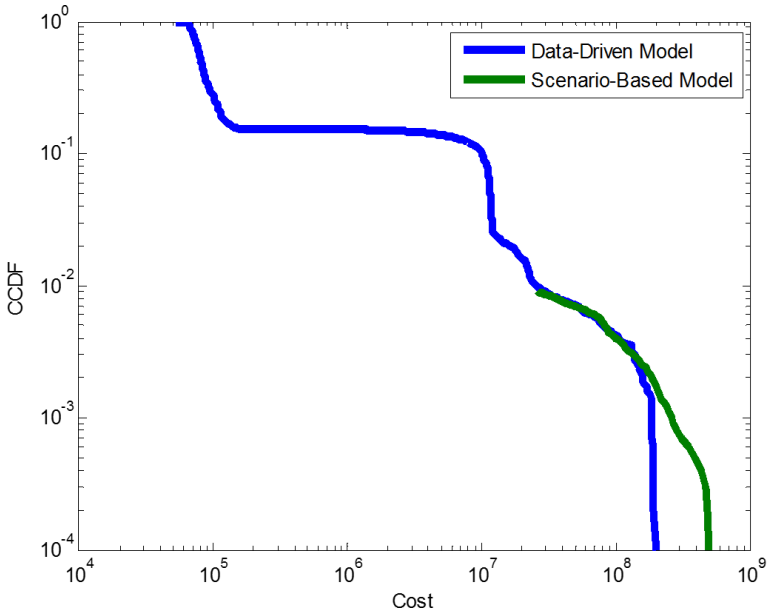


Figure 18: The final risk curve for malicious email incidents.

3.4 The Total Risk Curve

Once each type of incident is analyzed, the outcomes of each simulated year can be compared. The figure below shows a typical output, which displays the yearly costs due to several different vectors. The yearly costs are sorted via a complementary cumulative distribution function, so that a decision maker can explicitly see how often losses of a certain size occur.

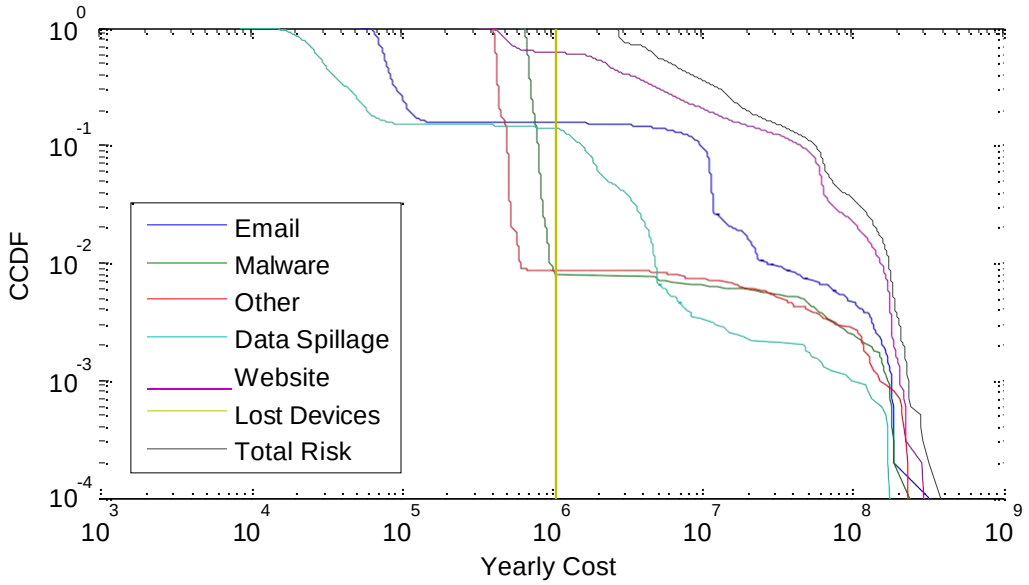


Figure 19: A typical cyber risk curve. Note that lost devices cost about the same amount each year, while other incident types have large ranges in losses. Websites are the largest risk to the organization and contribute the most risk to the curve.

There are several methods for comparing alternatives. Probabilistic dominance is one of the simplest. For example, if the maximum loss due to lost devices with full disk encryption is less than the minimum loss due to lost devices without full disk encryption, then the full disk encryption alternative is stochastically dominant.[45] First-order dominance results when the CCDF of one vector is always smaller than the other (e.g., the data spillage CCDF is always less than the website CCDF in figure 19).[46]

To obtain the actual value of each alternative, the risk attitude of the decision maker needs to be elicited. The risk attitude represents the tradeoff between different losses and uncertainty, meaning that a risk averse decision maker would value an uncertain deal with positive gains at less than its expected value.[47] Since many of the outcome distributions in cyber security are heavy-tailed, decision makers need to take special care to use a utility curve that accurately expresses their preferences at high losses (meaning that exponential utility functions may not necessarily be an appropriate choice). Since risk attitudes are a well-developed topic, they are not emphasized in this

[45] Note that in the case study in section 5, the full disk encryption alternative does not satisfy stochastic dominance; the example is used here to improve clarity.
[46] Second-order probabilistic dominance is also useful, given that it allows an analyst to state that any decision maker who is risk averse and prefers more money to less will prefer the alternative with 2nd order dominance.
[47] For example, a risk neutral decision maker values a 50-50 chance at $0 and $100 at $50. A risk averse decision maker would value the deal at less than $50.

dissertation. Instead, this dissertation focuses on intuitive insights that can be derived from risk curves (the CCDFs).

Once the risk curves and values are obtained, analysts can choose to communicate the results in a number of ways. Some organizations will prefer to put results back into a qualitative format (low, medium, or high risk). Others might use the median, or the 10^{th}, 50^{th}, and 90^{th} percentiles. In any case, calculating and presenting the probability distribution over yearly losses is important because it represents the most comprehensive view of losses. If decision makers want to simplify the insights, it is trivial to go from a risk curve to the average, median, or some other measure.

Analysts should exercise caution to ensure that the simplification of the risk curve preserves the integrity of the results. One possible simplification is the use of value at risk (VaR). Value at risk can be misleading because its name suggests a different meaning than it actually has, and it is poorly equipped to fully represent heavy tails. Value at risk is commonly mistaken to mean the maximum loss. For example, a VaR of $10 million is interpreted to mean that $10 million is at risk and can be lost, but no more. In fact, VaR specifies the losses at a defined percentile (e.g., 95^{th} percentile). A value at risk of $10 million means that with a probability of 0.95, the annual losses will be less than $10 million. However, VaR says nothing about what the probability of the losses could be when they exceed $10 million. Decision makers are often surprised to hear that a mutual fund lost $30 million when the VaR was $10 million. Since cyber impacts are heavy tailed, a significant portion of the risk resides in the top 5% of incidents.

Other metrics have attempted to alleviate this problem by using tail value at risk (also known as conditional value at risk). This metric reports the expected losses given that the VaR threshold (e.g., 95%) is exceeded. Organizations should evaluate these different metrics and choose the best option for their needs, but also take care to accurately communicate the meaning and results of the analysis.

3.5 Overarching Bayesian Network Model for Cyber Security Risk

The model presented for Space Corp is extendable into a general model for assessing cyber risk at any organization. Figure 20 shows a decision diagram for cyber security at an organization. The specific threats, vulnerabilities, and consequences will change depending on which organization is being analyzed, but the overall modeling process will be the same.

The diagram contains five groups of related nodes that assess different aspects of organizational cyber security. The first (green) relate to the adversaries that an organization faces and the details of defensive measures at an organization. The second group (red and light blue) summarizes different classes of scenarios for how attacks or incidents may occur at an organization.

For example, the attack process is modeled as penetration into the system, navigation through the system, extraction of the information, and use of the data (or action taken if no data is extracted). Countermeasures are also considered (purple), given that organizations may take steps to limit the success of an incident. Finally, the consequences (orange) translate an attack effect into monetary damages of different types.

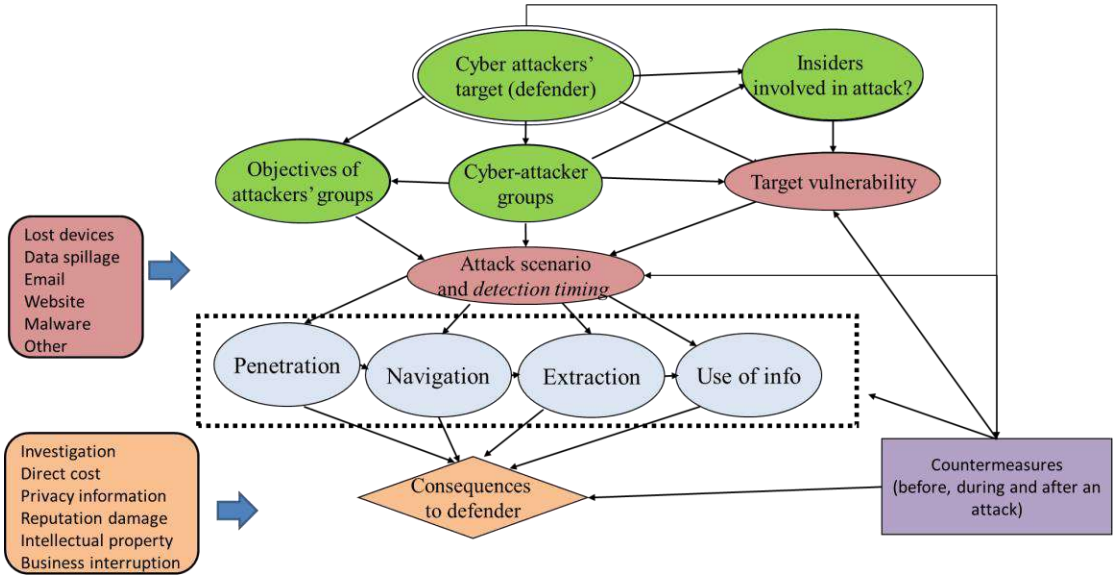


Figure 20: Decision diagram for cyber security. Modified from Paté-Cornell, presented at the International Risk Governance Council's Expert Workshop for Cyber Security Risk Governance, Zurich 2015.

The analysis is conducted from the perspective of an organization (the defender). Many parts of the model are conditioned on the specifics of which organization is being analyzed. The industry, size, and location of the defender will change the likelihood that it will become a target. The model also needs to be tailored to the technical specifications of the organization. For example, the internal network structure, the perimeter network exposures, and the detection capabilities all influence the probability that a cyber attack will be successful. Some of these details are inherently uncertain or unknown to the organization. The exact number of exposures may be uncertain, due to the constant growth of software vulnerabilities. Organizations may also have exposures that they are unaware of, as is often the case on legacy networks. Organizational churn can result in misconfigurations that cause network exposures as well, which are difficult to detect. One of the first steps in modeling cyber risk is to gather information about these uncertainties.

Organizations also need to learn about their attackers. Large retail organizations face threats from criminals and lone hackers, but are rarely targeted by nation state level adversaries. Manufacturers on the other hand may face attacks aimed at stealing their intellectual property. The wide range of attacker sophistication means that learning about what threats an organization faces is critical to understanding its overall risk and which defenses are most cost-effective.

Next, organizations need to understand the possible attack scenarios. Attackers typically follow a standard high-level process to gain access to a system which includes network reconnaissance, penetration, lateral movement, privilege escalation, data exfiltration, and the usage of data. Organizations need to understand these attack progressions and the safeguards in place that can detect or disrupt them. The uncertainties around these actions are captured by the attack scenario node.

Ultimately, a decision maker is largely concerned with the monetary losses associated with cyber security incidents. These losses include all impacts caused by the attack including business interruption, reputation damage, repair or replacement of equipment, and others. To determine the final impact, the decision maker needs to know what incidents have occurred and how much each of them costs. In the model, this is captured by an impacts node and a rate node.

If certain types of data exist, the decision diagram above can be simplified further. Figure 21 shows a modified diagram for the case where some historical data exist. In this situation the rate of incidents and impact distributions may be observed, meaning that the attacker and organization information are irrelevant to the cost, given that the rate and impact are known.

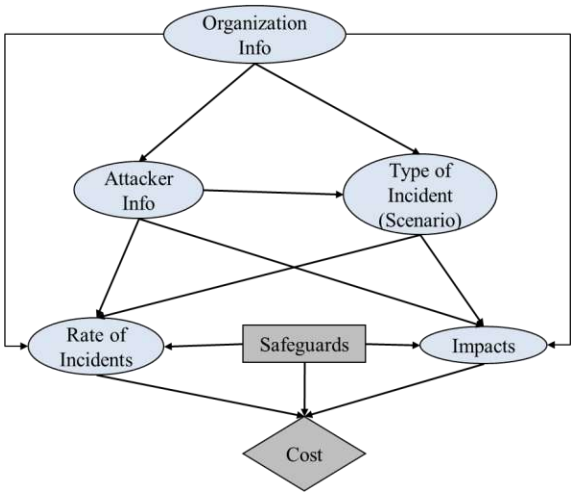


Figure 21: Modified decision diagram, where the rate and impact of incidents are observed.

3.6 Recommendations and Implications

Quantitative risk models require more input and effort than other risk assessment techniques, but are significantly more powerful. Analysts can test their assumptions via sensitivity analysis, which can determine what variables are most important while assessing the robustness of the final decision. Value of information calculations can prevent organizations from spending too much on costly penetration tests that may not deliver any value. At the level of CISOs, it enables optimal resource allocation under uncertainty. Quantitative cyber risk also improves risk communication at

the boardroom level. Cyber risk is immediately more compelling to decision makers when it can be communicated in dollar terms. Many boardrooms have treated cyber security qualitatively different than other business risks, using special tools and qualitative scales that prevent accurate comparisons across different types of business risks. Translating cyber risk into monetary outcomes enhances the conversation about cyber risk and can ensure that the subject of cyber risk is addressed. Finally, quantitative cyber risk is already becoming a critical tool for cyber insurers. In 2015, cyber insurance was already a large market and is forecast to grow rapidly over the next several years (Reuters, 2016). However, while traditional insurance markets are well modeled and have well-established actuarial tables, cyber insurance is still very much an unknown domain. Underwriters have already begun to invest heavily in cyber risk assessment tools and will likely drive much of the cyber risk quantification in the near future. PRA will be a critical tool to accurately assess cyber risk.

4 The Data

Currently, there are very few sources of publicly available cyber security incident data. Historically, this has been one of the major limiting factors that has prevented more work on quantitative cyber risk. However, organizations often have more data than they think, and their data are frequently more useful than they think. In this dissertation, I present a case study on the statistical analysis of cyber security incident data at a large, US-based organization. This exercise aims to discover useful insights as well as obtain probabilistic inputs for the risk model described in Chapter 3. Organizations may need to use special techniques to assess cyber risk in the cases where data do not exist. For example, large incidents occur rarely, meaning that organizations may not have historical incidents or data that can be used to inform a model. In these cases, scenario analysis needs to be used to assess the risk of large impact incidents.

The majority of the data used in this chapter come from a large, US-based organization.[48] Cyber security incidents were manually created by security operations center investigators over a six-year time span. Each incident contains basic information about when it was opened and closed, the investigator, systems impacted, an incident description, and an estimation of the total number of hours that it took to investigate and remediate the incident. Each incident was categorized using a certain taxonomy that, unfortunately, turned out to be poorly suited for risk analysis purposes. Therefore, the data were cleaned and partitioned into six different categories using a combination of incident tags, keyword searches, and manual categorization.

Sixty thousands incidents over a six year time period represents a substantial amount of data, which makes the use of statistical tools and techniques very convenient. Further, it is important to note that while some organizations may claim to record more data, they are often referring to a different type of dataset when they do. For example, executives are fond of saying that their organization gets attacked ‘millions of times per day’. This statistic refers to the number of connection attempts to their network, and is distinctly different from the creation of an investigation ticket.

Different data visualization techniques can be used to explore the data and determine what interesting analysis might be performed. Exploring the data is a crucial step that can sometimes be overlooked; many organizations probably have large amounts of cyber security incident data that have not been analyzed. Figure 22 shows over 60,000 cyber security incidents recorded over six years at the illustrative Space Corp. Note that when the hours of investigation are plotted on a log scale, it is easy to see that most incidents are resolved in less than ten hours. However, large

[48] Due to security concerns, further identifying characteristics about this organization may be obfuscated.

incidents have occurred (although more rarely) that took thousands of man-hours to investigate and accrue many additional costs.

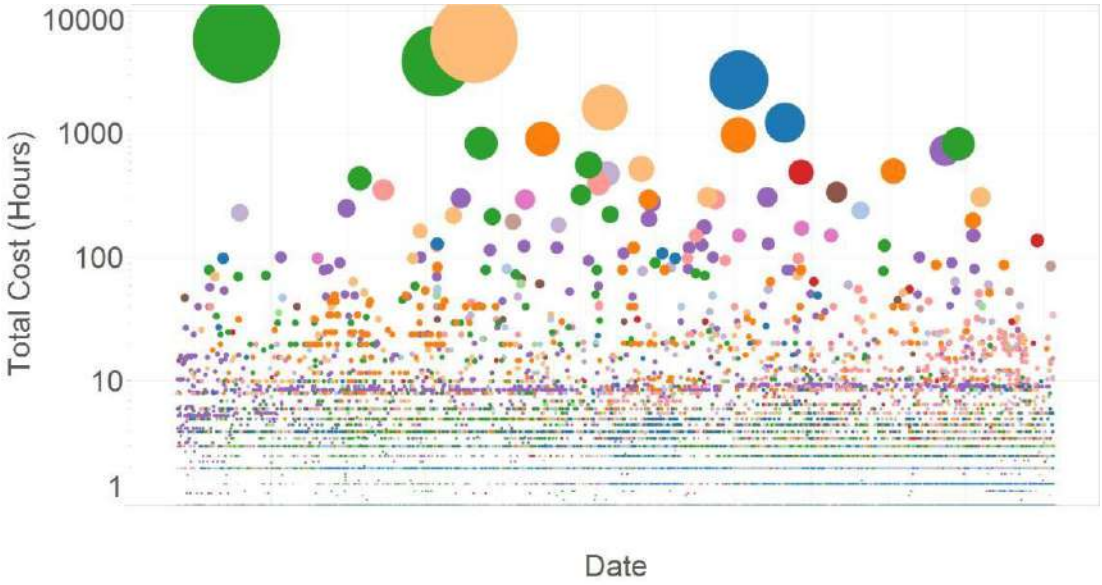


Figure 22: 60,000 cyber security incidents at a large organization.

Zooming into these data, some interesting trends emerge. Figure 23 shows the number of shellshock attacks against the organization over time. The shellshock vulnerability allows an attacker to execute arbitrary commands against servers, which can lead to unauthorized access. Shellshock was publicly announced on September 24th, 2014. Within five hours, the first shellshock attack was detected against this organization, illustrating how quickly adversaries can use vulnerabilities. There is also a clear pattern in the day of the week that attacks were most likely to occur, with large increases on Thursdays and Fridays. The shellshock attacks at this organization also correspond to a European workday more so than a US workday, which could suggest where some of the attackers resided. Finally, shellshock attacks continued for many months, demonstrating the importance of reaching 100% patching levels, since unpatched machines are likely to be found after enough time.

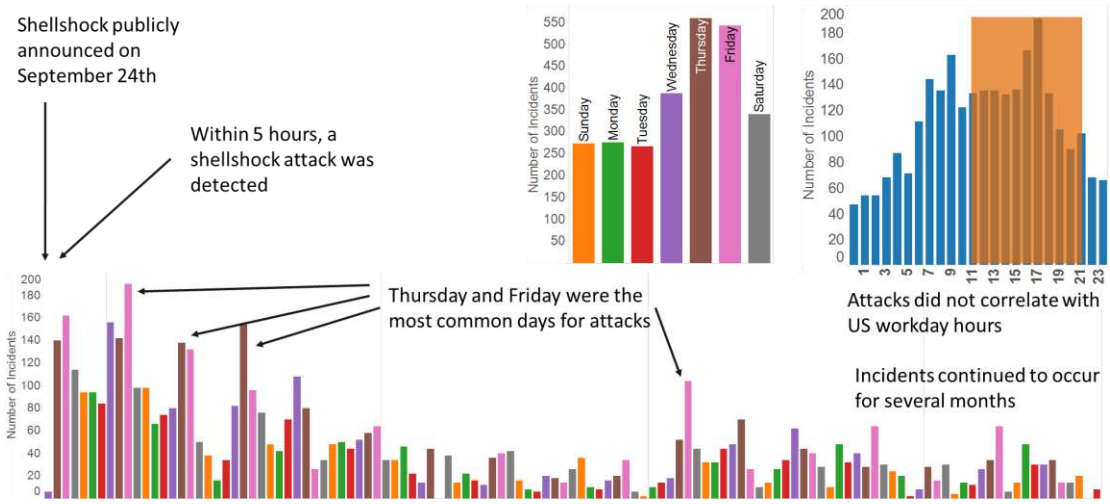


Figure 23: Shellshock attacks at a large organization.

Trends in specific attack vectors can be analyzed as well. Figure 24 shows the cumulative number of lost devices over time. First note that there is a large change that takes place around 2012. This change in the slope (reflecting a change in the rate of lost devices) occurred because of a change in reporting guidelines. Before 2012, only lost laptops were generally recorded, while cellphones and security tokens were recorded after 2012. Another anomaly occurs in 2014, which can be attributed to an audit that recorded many lost devices at once. Besides these minor events (which must be investigated to ensure the correct conclusions are being reached), the rate of lost devices at this organization is remarkably constant over time so far. This is useful for risk modeling purposes.

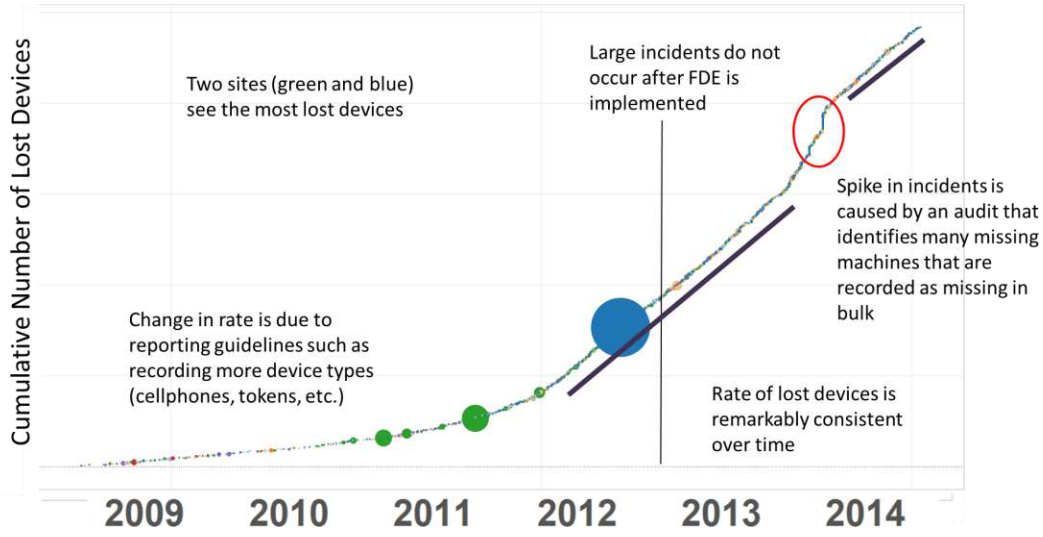


Figure 24: Lost devices at a large organization.

Several other tricks can be used to assess cyber security incidents. Changes in reporting guidelines or incident definitions may lead to discontinuities that need to be identified. Figure 25 shows the

number of incidents recorded each month over the six years. Several lines are plotted representing all incidents, and incidents that were larger than 2, 5, 10, and 20 hours. The linear regression for each line is also plotted. Note that there is a large increase in the total number of incidents (blue line). However, the increase is much smaller for incidents greater than 2 hours and nonexistent for incidents great than 5 hours. This shows that the rapid increase in the total number of recorded incidents was driven by small incidents. After speaking with the CISO of this organization, it became clear that the increase was due to a change in reporting guidelines for website incidents and not indicative of any changes in how often the organization was attacked.

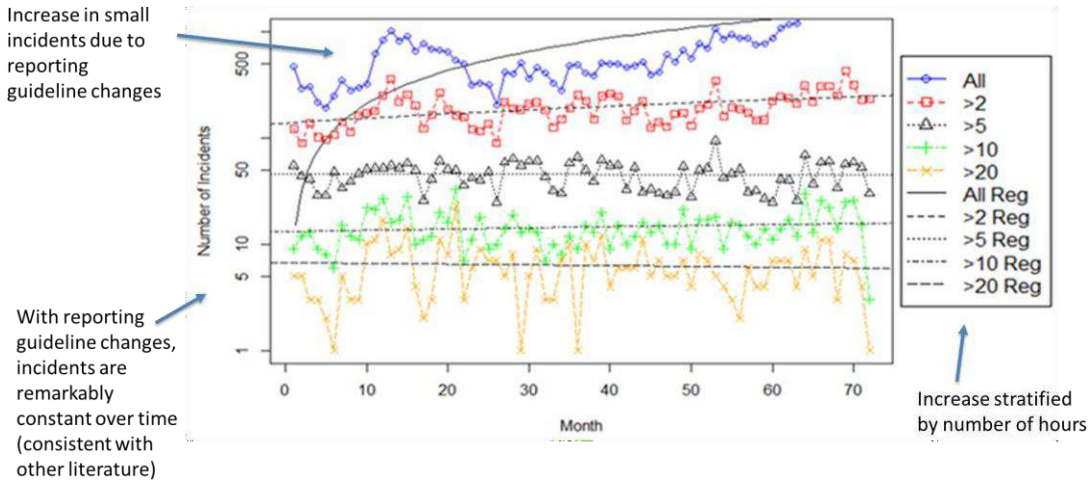


Figure 25: Cyber security incidents at a large organization by hours of investigation. The number of incidents is shown for each month, along with a linear regression. Note that because the *y*-axis is on a log-scale, the linear regression may appear curved.

Cyber security incident data can be incredibly useful, but analysts need to understand the context of the data. Working with the IT staff and the CISO is critical so that the correct conclusions are reached. Otherwise, an innocuous change in reporting guidelines could lead to incorrect conclusions about the increase in the number of cyber attacks.

4.1 Analysis of Sparse Data

The minimum amount of data that is useful can be quite small. On April 16, 2015, the US Department of Energy (DOE) released data on 1,131 cyber security incidents through a Freedom of Information Act Request issued by Stephen Reilly, an investigative journalist from *USA Today* (Reilly, 2015). The data come from the Joint Cybersecurity Coordination Center (JC3) for the Department of Energy, which receives cyber security incident reports from DOE sites. The data contain an ID number, date, category, site (redacted), program office, summary (redacted), and status (closed for all).

ID	Date Created	Category	Site	Program Office	Summary
648220	10/4/2010 5:39	Malicious Code	Exemption b(7)(E)	HQ	Exemption b(7)(C)
648240	10/4/2010 10:45	Malicious Code	Exemption b(7)(E)	HQ	Exemption b(7)(C)
648279	10/5/2010 8:49	Malicious Code	Exemption b(7)(E)	HQ	Exemption b(7)(C)
648312	10/5/2010 13:44	Malicious Code	Exemption b(7)(E)	EM	Exemption b(7)(C)
648313	10/5/2010 14:33	Malicious Code	Exemption b(7)(E)	EM	Exemption b(7)(C)
648314	10/5/2010 14:33	Malicious Code	Exemption b(7)(E)	EM	Exemption b(7)(C)
648315	10/5/2010 14:33	Malicious Code	Exemption b(7)(E)	EM	Exemption b(7)(C)
648338	10/5/2010 16:21	Malicious Code	Exemption b(7)(E)	HQ	Exemption b(7)(C)
648388	10/6/2010 13:10	Malicious Code	Exemption b(7)(E)	SC	Exemption b(7)(C)
648400	10/6/2010 14:40	Compromise - User (Intrusion Successful)	Exemption b(7)(E)	SC	Exemption b(7)(C)

Figure 26: A sample of data from US DOE.

Only six types of incidents were listed in the data, namely malicious code, successful distributed denial of service (DDoS), unsuccessful DDoS, compromise (user), compromise (root), web defacement, and unauthorized use. The dataset contains little information, but there are still many interesting insights that can be gained from analyzing the data. Figure 27 shows the cumulative number of cyber security incidents over time, color-coded by incident type. It is immediately apparent that the majority of incidents involve malicious code and that the rate of incidents is relatively constant over time.

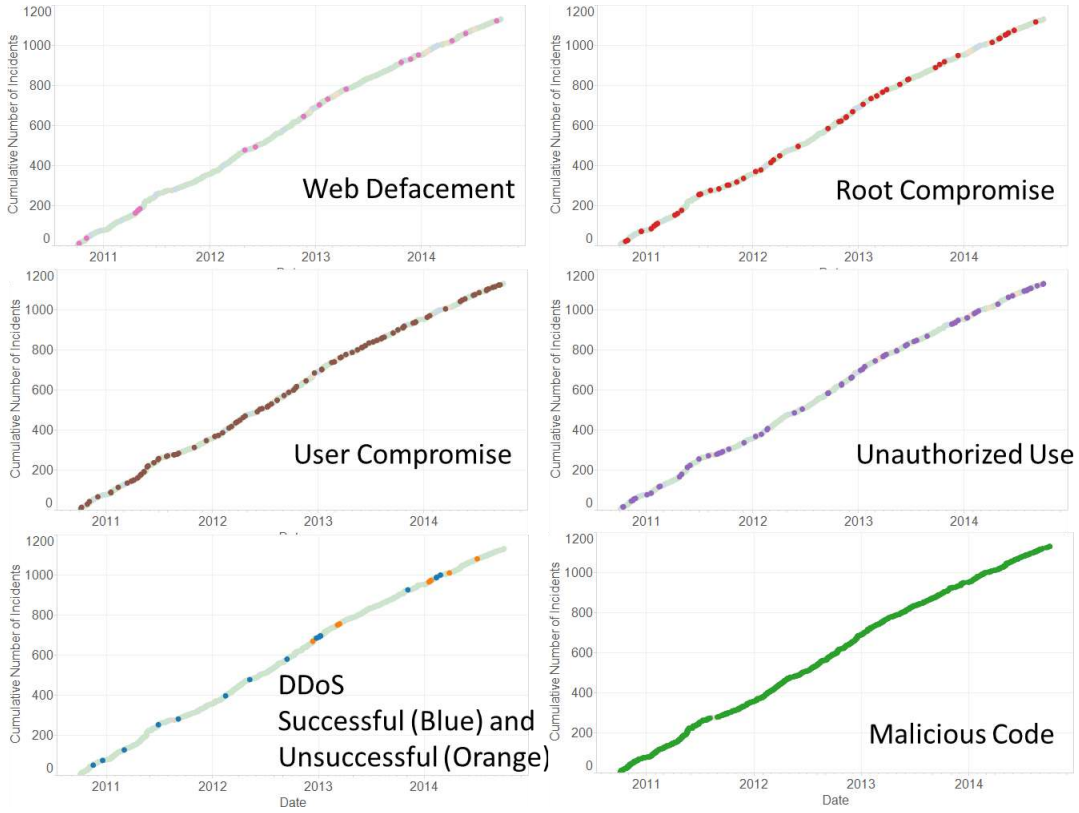


Figure 27: Cyber security incidents at US DOE. Note that malicious code incidents occur most frequently.

The data span from October 4, 2010 to October 3, 2014, meaning that every month has been recorded four times.[49] Figure 28 shows that the summer months typically experience fewer incidents. It is clear from the data that security incidents are created from Monday through Friday during normal working hours.[50]

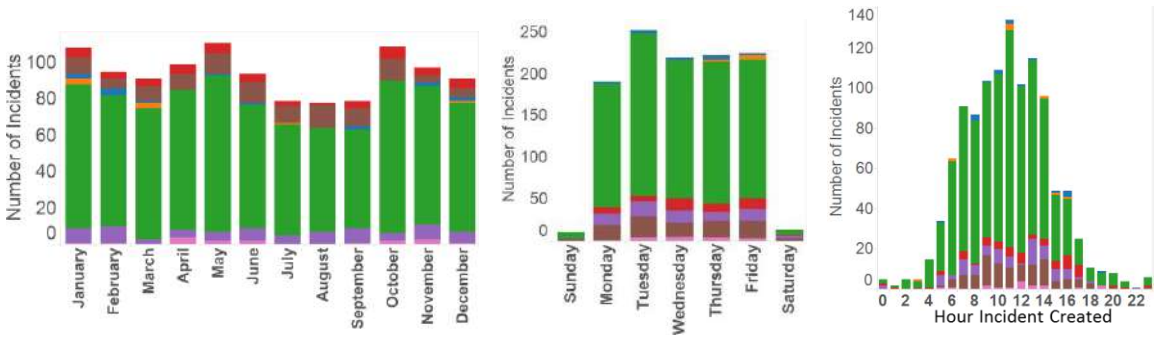


Figure 28: DOE incidents by month, day, and hour.

Many other conclusions can be drawn from the data, including the fact that incidents appear to be created manually and the security operations center is likely not physically staffed 24/7 (Kuypers & Paté-Cornell, 2016). Figure 29 shows the times that DOE incidents are recorded, the majority of which occur between 5 a.m. and 4 p.m. This pattern could be the result of the attacks (who might only attack from 5AM to 4PM) or the security operations workers (who are on the job 5 a.m. to 4 p.m.). To gain insight into which is correct, the distribution of incidents can be compared to the illustrative organization studied here that recorded 60,000 incidents (Figure 29b). Note that while more incidents are created during normal working hours, a steady volume of incidents occurs outside of the standard workday, indicating that the hourly pattern in the DOE data is an artifact of their work schedule.[51]

Other interesting patterns can be observed as well. Website incidents occur at a relatively constant rate throughout the day and across weekdays (figure 29c), while lost devices are reported most often on Mondays and early in the morning (figure 29d), since workers report devices that are lost on weekends and evenings first thing the next working day.

[49] Note that the FOIA request was from October 1, **2011**, but data was delivered from 2010 onwards.
[50] Interestingly, JC3 is listed as a 24/7/365. https://www.first.org/members/teams/jc3-circ
[51] Note that the working hours are offset because they are recorded in UTC.

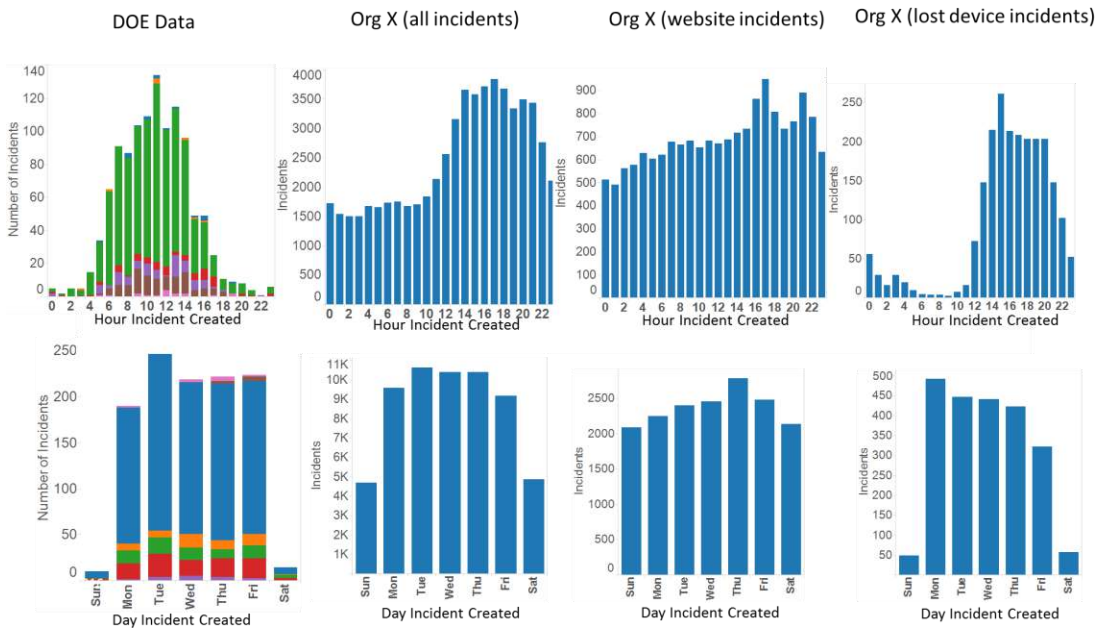


Figure 29: A comparison of two security operations centers.

While the data to not contain any information on the severity of different incidents, the rate at which incidents occur can be studied. Figure 30 shows the number of incidents in each month. The rate of incidents is decreasing over time, but that this is mostly driven by a decrease in malicious code incidents. The rate of root compromises, user compromises, DDoS attacks, web defacements, and unauthorized use incidents has been either constant or decreasing in the considered time interval.

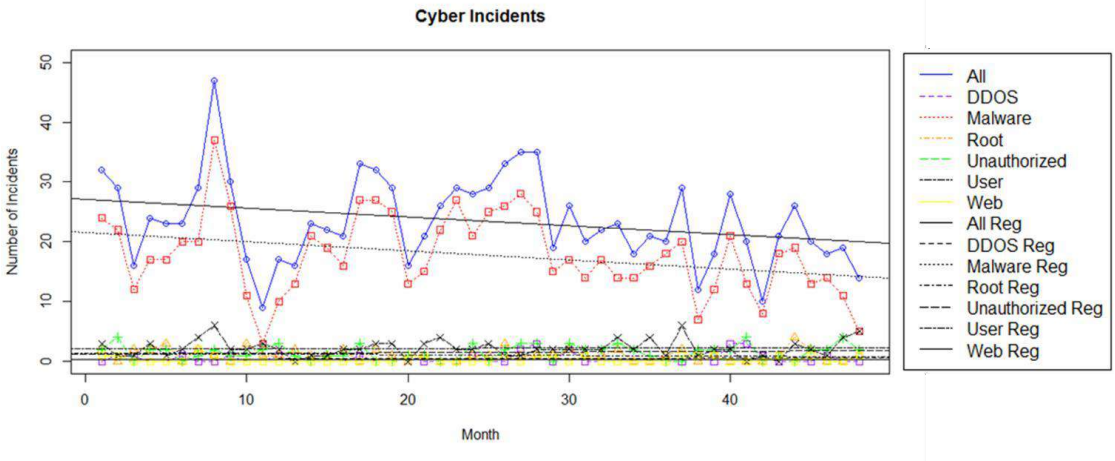


Figure 30: The number of incidents per month at US DOE by type. Black lines are a linear regression.

The arrival process of security incidents can be studied as well. A Poisson process is a reasonable guess for the arrival of incidents, given that security incidents arise from a process where many endpoints (servers, workstations, etc.) each have a small probability of being compromised. To gain

more insight, the interval between incidents can be calculated. Only intervals less than ten hours are used, because the effect of the workday can be seen in data and distort the results.

Figure 31: Arrival times of cyber security incidents at US DOE from 2010 to 2014. Note that the time decreases exponentially (shown in figure 32), but additional peaks occur that correspond to roughly one work day between incidents.

The time between arrivals in a Poisson process follow an exponential distribution. Testing several distributions, the data are found to be best explained by an exponential distribution with $\frac{1}{\lambda} = 2.5869$, meaning that the expected time between incidents was roughly 2 hours and 35 minutes. The arrival process was tested against many other distributions, including gamma, logistic, normal, Weibull, negative binomial, and many others to ensure a good fit.[52]

[52] The matlab function "allfitdist" by Mike Sheppard is used, available here: http://www.mathworks.com/matlabcentral/fileexchange/34943-fit-all-valid-parametric-probability-distributions-to-data

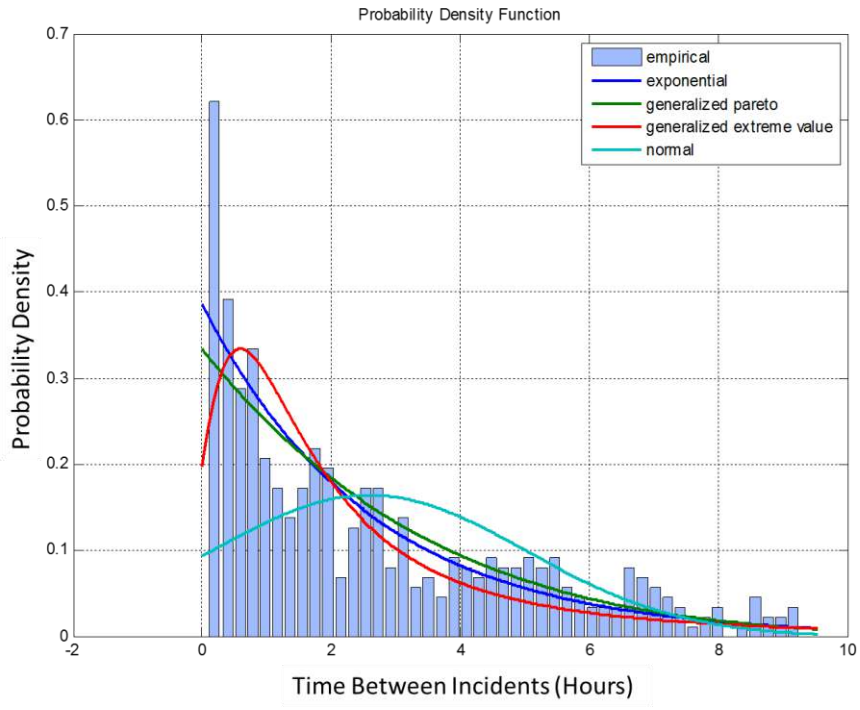


Figure 32: Probability density function for the time between incidents, along with several fits. The exponential distribution gives the best fit, suggesting the arrivals follow a Poisson process.

4.2 Conclusions

In many ways, cyber security is still in its infancy. Rigorous methods for assessing cyber risk are just beginning to emerge, and basic ground truths about cyber security incidents are still largely unknown. This chapter demonstrates how even heavily redacted and sparse statistical data can be used to determine interesting insights about past cyber security incidents at a large organization. Analyzing past cyber security incidents can lead to actionable insights as well as high-quality probabilistic assessments for quantitative cyber risk modeling.

5 Applications

To demonstrate how the model works, a case study tailored to a large hypothetical aerospace organization that we call Space Corp was presented as discussed in chapter 1. The data used in this case study were obtained from an anonymous organization to show how an analysis of historical cyber incidents can inform the model.[53] This chapter focuses on a case study involving a decision maker at an aerospace organization who would like to assess cyber risk and is considering several possible safeguard investments with the aim of quantifying which strategy would be most effective in reducing cyber risk.

Aerospace organizations face unique challenges in securing themselves against cyber threats. In addition to common attacks such as malware infections and website defacements, persistent attackers may also attempt to compromise intellectual property, engineering data, or other sensitive information. Many aerospace organizations also exist in a complicated collaborative environment that requires balancing security with research partnerships and public engagement. As a result, the common qualitative cyber risk analysis techniques discussed in Chapter 2 may be insufficient for providing decision support to CISOs faced with implementing cyber security defenses.

Quantifying the impacts of cyber security incidents presents major challenges as well. Some costs are well understood, such as credit monitoring for privacy information disclosures. However, spacecraft failure due to a network intrusion, sensitive information loss, espionage of engineering methods, and reputation damage are much more difficult to quantify. The high-risk nature of spaceflight missions makes quantitative risk analysis a critical tool for decision makers.

5.1 Model Setup

The following sections present several case studies where cyber risk is analyzed for the fictitious aerospace organization that we called Space Corp. As mentioned earlier, while Space Corp is not a real organization, the data used in these case studies come from real organizations. Data used here on cyber security incidents come from a large US-based organization that recorded over 60,000 cyber security incidents over a six-year time period. While there are significant differences between the actual data source and Space Corp, the data are used in this case study to illustrate the risk management process from start to finish. Further, throughout the case studies, publicly available data are used, along with expert assessments. These expert assessments were obtained using two elicitation scenarios. Some experts were given a description of Space Corp and asked to give

[53] Note that while the attacks experienced by an aerospace organization would be different from the data presented here, the analysis and method are identical.

assessments under the assumption that they actually worked for Space Corp. Other experts were asked to give assessments for their own organization. In the second case, the assessments may be slightly modified or obfuscated before being used as examples for Space Corp (for example, a loss assessment might be converted from an actual organization to Space Corp using the number of employees as the scaling factor). The result of these assessments is a realistic modeling exercise that mirrors some real-life examples that have been performed with other organizations (but cannot be publicly released due to security concerns).

In each case study, historical data are combined with other information sources to obtain incident frequency and impact distributions inputs for a probabilistic model. The model emphasizes monetary impacts and quantifies costs by combining the historical data of a large organization, non–cyber-related aerospace failures, and willingness-to-pay. Monte Carlo simulations are used to compare the cost-effectiveness of different security safeguard investments, such as two-factor authentication[54] for webmail and data loss prevention.

5.2 Cyber Incidents at Aerospace Organizations

Many aerospace companies have already experienced large cyber security incidents. Lockheed Martin was the victim of an intrusion by the Chinese government, resulting in the loss of a significant amount of information about the F-35 Joint Strike Fighter (Schwartz, 2011). The same Chinese hacking group has targeted several other aerospace companies, and documents show that radar designs, engine blueprints, and export controlled information have been compromised at several organizations. The hacking of satellites has also occurred, as detailed in a congressional report that describes four instances where an attacker obtained unauthorized access to Landsat-7 and Terra EOS spacecraft in 2007 and 2008 (US–China Economic and Security Review Commission, 2011). A hacker in Romania gained access to the Jet Propulsion Lab's network in 2012 (Martin, 2013). Many other cyber attacks against NASA have been documented as well (Martin, 2013). Several talks on hacking satellites have even been given at various hacking conferences including Defcon and Black Hat.

Aerospace companies make attractive targets due to their high rates of sensitive technological information and must operate in a secure environment. However, companies need to make risk-based decisions to ensure that cyber security is balanced with risk management of other typical engineering failures. Striking the right balance is critical.

[54] Two-factor authentication (TFA) requires that a user enter a password and a token for authentication. For example, a user will keep a device that generates six-digit numbers every minute and enter the number and their password when logging in.

5.3 Model Scope

The framing and setup of the model in this chapter closely parallels the model detailed in Chapter 3. Space Corp aims to quantify cyber risk using a mix of historical information, public data, and expert knowledge. In particular, the organization divides cyber security incidents into several attack scenarios. Further, the value of four core security safeguards is modeled and evaluated, namely data loss prevention, two-factor authentication, a demilitarized zone for websites, and full disk encryption.

Data Loss Prevention – Data loss prevention (DLP) is a type of software that monitors data on a network. Sensitive data can be identified by searching for files marked with certain classifications, or certain data formats can be monitored (e.g., nine-digit numbers that would suggest a social security number). DLP can be an effective safeguard against unintentional data disclosure either by alerting a user that something appears to be sensitive (and therefore needs to be encrypted), or by fully preventing a user from distributing a file that is found to be sensitive.

Unfortunately, there are several issues with DLP technology, including benign files that might be marked as sensitive and sensitive files marked as benign. Further, an attacker can still exfiltrate information relatively easily, for example by encrypting the information before it is transmitted outside the organization.

Two-Factor Authentication – Currently, Space Corp is vulnerable to attackers who phish credentials from users. Research shows that phishing success rates are relatively high (10–20%), meaning that an attacker can obtain credentials fairly easily. The credentials can then be used to access email or other applications which can lead to spamming, theft of confidential information, or moving laterally to other systems in the organization. Two-factor authentication improves security by requiring both a password and a one-time security token for a user to gain access to an application. Persistent attackers can still use clever attacks to obtain access to systems; however, TFA increases the level of difficulty substantially.

DMZ for Websites – Websites attacks occur often and at all severity levels, making them a high risk at Space Corp. Legacy equipment is common throughout the network and independent research groups have created independent servers, resulting in a heterogeneous environment. Further, many web applications are not maintained, since many projects have no funding after the project is completed. Moving all websites to a DMZ would consolidate the website resources to a central location, making patch management and monitoring much easier. Further, the DMZ resides on a

segmented portion of the network, so that an adversary cannot easily move laterally to other parts of the network if a web server is compromised.

Full Disk Encryption – One of the most common incident scenarios is the loss of unencrypted data on laptops. When a laptop is lost or stolen, it is simple for an attacker to recover unencrypted data from the device. Full disk encryption secures the information on the device by requiring a passcode before the computer will boot. Space Corp has already implemented full disk encryption, but would like to assess its effectiveness. The benefit of other lost device safeguards, including asset recovery software, theft awareness training, and laptop loaner programs, will also be evaluated.

In order to determine the value of each of the safeguards above, the frequency and impact of cyber attacks must be carefully modeled, along with the effectiveness of each cyber security safeguard. The following sections describe this process for each attack type.

5.4 Data Spillage

Cyber losses often occur through user error. In particular, information may be unintentionally disclosed, resulting in credit monitoring costs or reputation damage. Organizations categorize these incidents as data spillage[55], which is defined as the unauthorized (and accidental) release of information. The type of information released can include personally identifiable information or sensitive information, such as classified information or intellectual property. There are several causes of data spillage.

***Privilege error*:** Users can have unauthorized access to materials due to misconfigurations or privilege errors. For example, research groups might store information on a SharePoint (a collaboration application) site, but the controls might be improperly configured so that other researchers have access to materials.

Ignorance: Users may share sensitive materials with unauthorized individuals because they are unaware that the material is sensitive or that a user is not authorized to handle that material. For example, a human resources worker trying to collect names, addresses, and birthdates of employees

[55] The term *data leakage* is also sometimes used. In this dissertation, data leakage is classified as a type of data spillage that occurs slowly over time. For example, applications that anonymize users incorrectly may "leak" information that can be used for deanonymization over time.

might email an Excel document out to a large group asking for people to proofread their information.

***Mistakes*:** Even if users are well informed about regulations and the systems are well configured, data spillage may still occur. Users forget to pick up documents from printers, or a user might forget to click the "encrypt" button when sending an email of social security numbers.

Data spillage incidents have a large variation in size, ranging from a single individual being mailed the incorrect tax form to an attachment containing 10,000 social security numbers that is sent to the wrong mailing list. The costs associated with data spillage incidents can also range across many potential impacts. Incidents need to be investigated to determine the scope of the spillage. Incident remediation can involve scrubbing data from systems where information was leaked, changing access controls to prevent similar errors from occurring in the future, and notifying all of the stakeholders of the information that was leaked. In some cases where the data spillage is especially severe, reputation damage or breach notification procedures may be implemented. However, other impacts are rare: business interruption rarely occurs because data disclosure is not likely to impact operations, intellectual property costs do not occur because information spillage is typically limited in scope and distribution, and direct costs do not occur since equipment is not damaged.

Data spillage incidents can be very intricate, consisting of several events that would be difficult to model as a chain of events. For example, in 2010, Stanford Hospital released medical records to a subcontractor for processing. The subcontractor mistakenly released the data to a job applicant as part of a test designed to test the applicant's ability to process medical data. The job applicant then took the medical records and posted them to an online forum, along with a description of the task and a request for help in processing the data (Ouellette, 2014). Roughly 20,000 medical records were exposed. Despite the unlikely chain of events that led to this incident, data spillage can be modeled simply by modeling the three different mechanisms of data spillage, which each occur regularly over time.

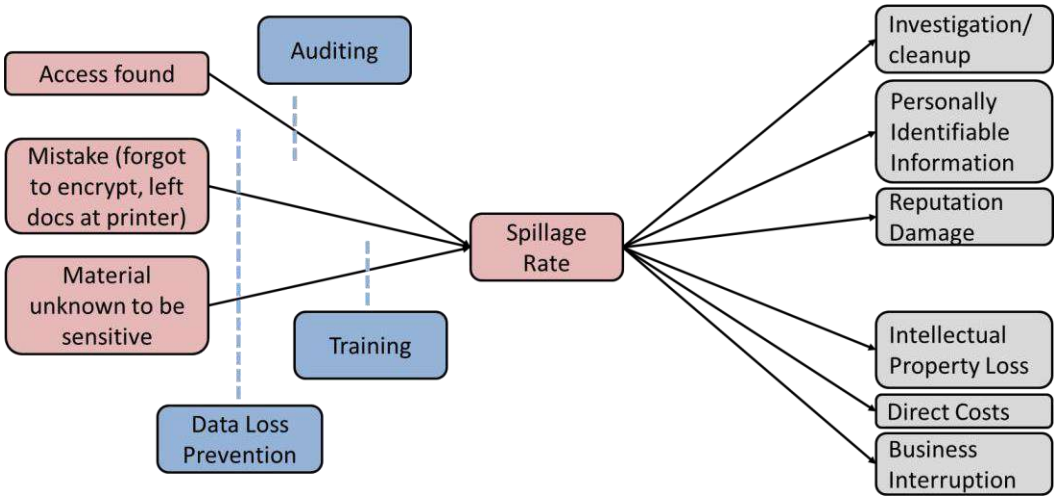


Figure 33: Illustration of the mechanisms of data spillage, along with the costs and safeguards. Note that IP loss, direct costs, and business interruption are in gray to signify that they do not occur often at Space Corp.

5.4.1 Data Spillage Frequency and Severity

Data spillage incidents can be broadly categorized into three different regimes:

A. Small incidents consisting solely of investigation costs.

B. Medium incidents, where most of the time corresponds to investigation costs, but some portion of the time spent may relate to notifying stakeholders or documenting small-scale personally identifiable information (PII) disclosures.

C. Large incidents, where additional costs relating to PII disclosure may occur and are not included in the investigation time.

Figure 34 shows the distribution of investigation times for data spillage incidents. The division into three regimes is notional, but is useful to broadly characterize the different impacts associated with data spillage incidents.

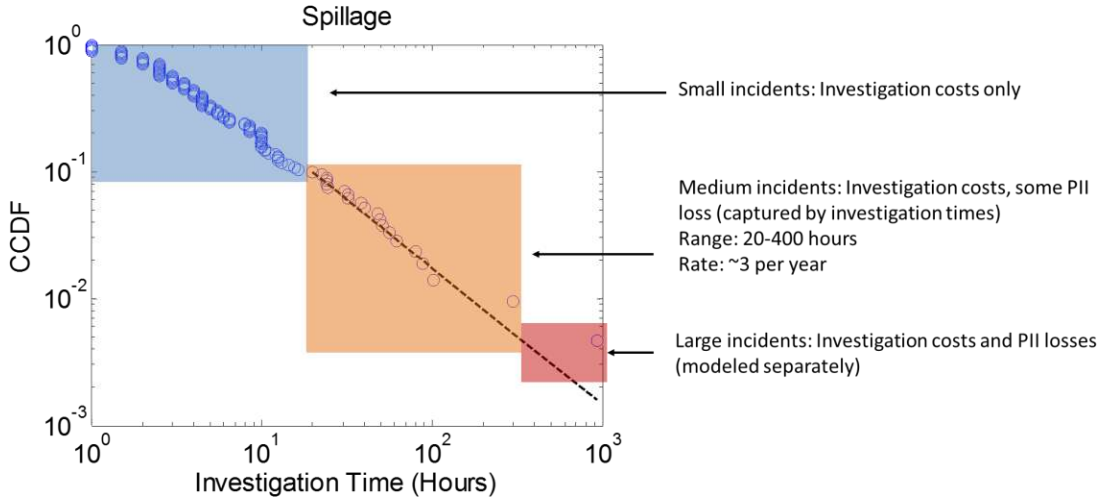


Figure 34: The complementary cumulative distribution function of investigation times for data spillage incidents. Note that the largest incidents may involve additional costs that are modeled separately.

Figure 35 shows the CCDF for the hours of investigation for data spillage by year. The distribution has remained approximately constant over time, indicating that the distribution of past incidents is an excellent approximation of the severity of future incidents.

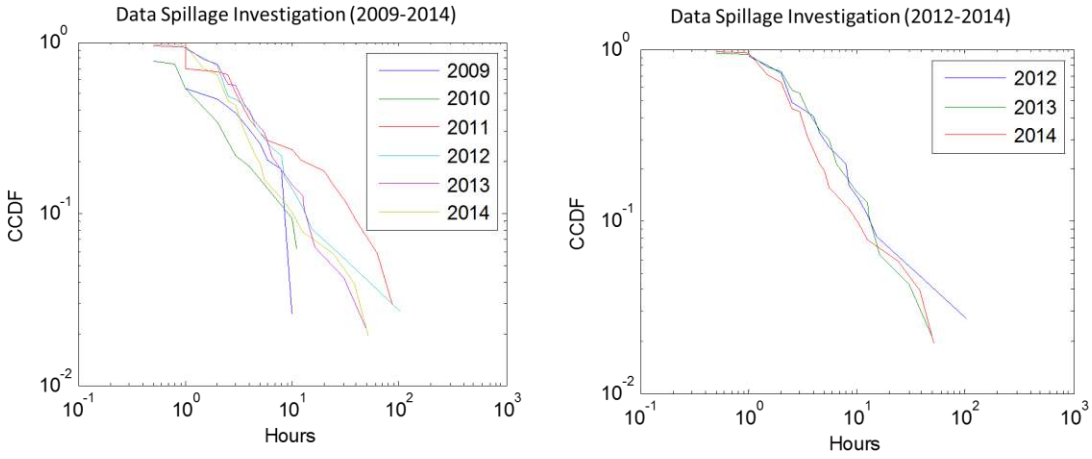


Figure 35: Data spillage impact distribution over time. Note that the distribution varies each year (left figure), but the distribution has remained very similar the past three years (right figure).

Analysis of Data Spillage Rates

Data spillage incidents are largely caused by human error which results in a steady stream of incidents that occur over time. Compared to other incident categories such as malware incidents or website attacks, data spillage incidents occur at a relatively low rate of roughly 40 incidents per year at Space Corp. Larger incidents that take more than 20 hours to investigate occur more rarely, at the rate of about seven per year. Figure 36 shows the rate of data spillage incidents of different sizes over time.

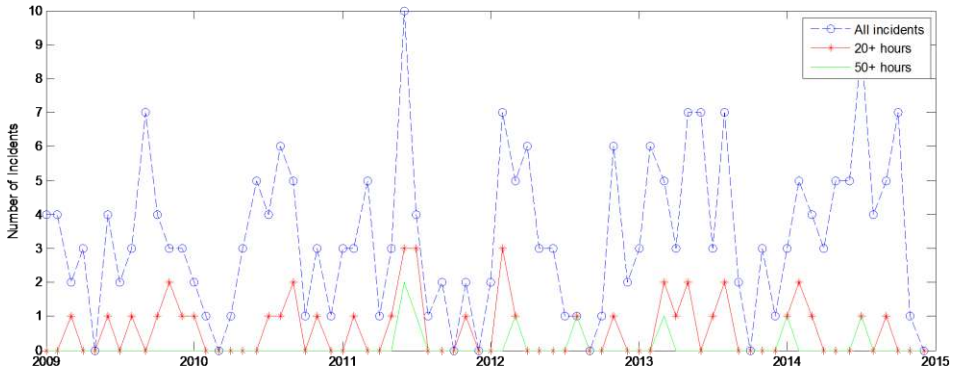


Figure 36: Data spillage incidents over time.

Taken together, the analysis of the rate and severity of data spillage incidents indicates that historical events may be good indications of future data spillage incidents. Therefore, the model utilizes a rate equal to the number of incidents that occurred each year on average. The severity of

each incident is calculated through bootstrapping, meaning historical incidents are randomly sampled to obtain the severity of an individual incident (see Chapter 3).

5.4.2 Data Spillage Impacts

Data spillage incidents can result in a variety of costs. For Space Corp, these costs occur via investigation time, costs associated with personally identifiable information loss, and reputation damage. Each impact is carefully modeled through the use of historical data, open source information, and expert elicitation.

Investigation costs

As discussed above, the distribution of investigation times is derived from historical data (see figure 34). Each hour of investigation is costly to the organization, given that a security operations investigator has to spend time investigating and remediating the incidents. The Space Corp finance department assigns a time cost of $100 per hour of investigation.

Reputation costs

If a data breach is the result of gross negligence or especially severe, reputation damage may result. In this case, Space Corp assesses that the threshold for reputation damage occurs when the investigation time reaches 500 hours or more, meaning that reputation damage can occur only for especially severe incidents.[56] Experts assess that for incidents that exceed this investigation time, there is a 50% chance that the large incident does not attract attention, meaning that there are no additional costs. There is a 45% probability that the incident causes additional oversight requirements by regulators, which includes audits, new compliance standards, and other reviews. In the past, external audits and requirements have cost Space Corp between $1 million and $2 million. Auditing costs are therefore modeled as a uniform distribution between $1 million and $2 million. Finally, there is a 5% chance that the data spillage incident results in severe reputation damage where future contracting will be reduced. This cost is modeled using a beta distribution (see figure 37).

[56] Note that while it is possible that an incident that takes a small number of hours to investigate results in large reputation damage, there are no examples of this in the dataset.

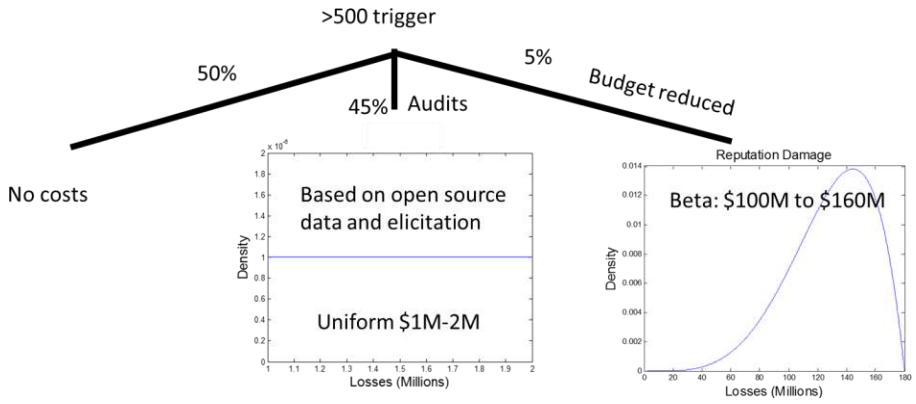


Figure 37: Model of reputation damage.

Personally Identifiable Information (PII) costs

Data spillage incidents can often result in the release of names, addresses, social security numbers, or other PII. If a PII disclosure involves a sufficient number of records, large costs become realized in the form of contacting impacted individuals and offering credit monitoring services. Space Corp does not handle PII of any customers or third parties besides its internal employees. Therefore, the number of records that could be exposed is capped at 2,000 individuals, plus 5,000 employees who no longer work with the organization. Based on several instances of PII disclosures at Space Corp and expert elicitation, the lost distribution due to PII disclosure is assessed to be uniform between $60,000 and $5 million.

Other organizations may have significantly different distributions. For example, different customer databases might exist that contain thousands, millions, or tens of millions of records, leading to PII losses that should be modeled as a heavy-tailed distribution.

Other costs

While it may be possible for data spillage to result in business interruption, direct monetary costs, or intellectual property loss, these types of losses have been rare at Space Corp so far. Other organizations may assess these risks differently. For example, in certain highly sensitive business environments, any disclosure of information could lead to a halt in operations. Space Corp assesses the chances that a data spillage incident leads to these loss vectors to be sufficiently low that they are not included in the cost function.

5.4.3 Risk Curve

Once the rate of data spillage incidents and the impact distributions are assessed, a Monte Carlo simulation is run to obtain a risk curve for data spillage, which is shown in figure 38. Analyzing the distribution, several insights are apparent. First, data spillage is comparatively low risk, with

typical losses (90th percentile) ranging from $10,000 to $100,000 per year. These costs are driven by incident investigation, which does not vary significantly across years. Additional losses due to auditing become apparent in the $2–$4 million range and occur roughly 10% of the time. Finally, an extreme risk region due to reputation damage exists toward the tail of the distribution. Note that in the tail region, it is important to carefully consider the accuracy of the model. For example, large incidents may cause a breakdown in the assumption of independence across events, meaning that one large incident may change the probability of additional large incidents.[57] Further, these losses often correspond to a possible existential risk to the organization, which is not modeled from historical data.

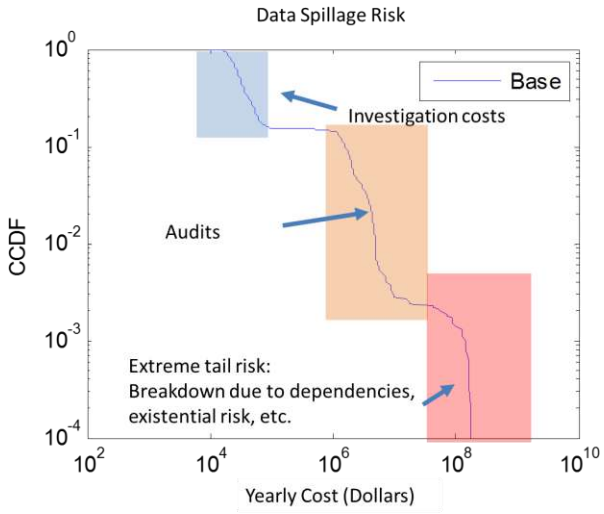


Figure 38: A risk curve for data spillage incidents.

5.4.4 Data Spillage Safeguards

Data spillage will likely continue to be a challenge for many organizations due to the human error that is involved. The question becomes how to manage the risk due to data spillage events effectively. Several safeguards could reduce the risk, including auditing to reduce privilege errors, training to reduce the rate of incidents caused by lack of awareness of data sensitivity labels, and data loss prevention (DLP) software. DLP software is designed to detect sensitive information and protect its use, transmission, and storage. A variety of techniques are used to identify sensitive data, including explicit labeling, pattern recognition (for example, looking for nine digits that indicate a social security number), and keyword searches. Currently, challenges still exist in eliminating false positives and capturing all of the sensitive information to be labeled. Further, DLP still has

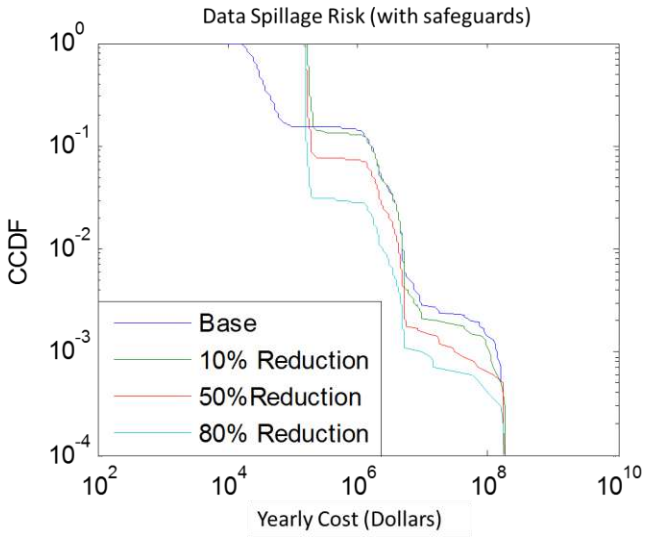
[57] Whether the conditional probability would go up or down is uncertain; a large incident may signal significant problems, meaning more large incidents are likely. However, the large incident may also cause additional resources and attention to be focused on the problem, resulting in a decrease in the risk.

limitations, namely interfering with functionality or missing certain types of data spillage. For example, DLP cannot prevent screenshots or photographs from copying sensitive information.[58]

Enrollment Alternatives

Technologies are often more effective when they target the users who are most likely to cause errors. At Space Corp, personally identifiable information resides on a relatively small number of endpoints, the majority of which are in the human resources (HR) and finance departments. A limited deployment of data loss prevention for these endpoints would cost approximately $150,000 per year in software and licensing. Additionally, an extended deployment ($500,000 per year) could place DLP software on many more endpoints in an effort to control data spillage across the organization in general.

Based on industry information and trials of DLP at other institutions, Space Corp assesses that a limited enrollment of DLP software will reduce the rate of data spillage incidents by 50%. Using this information, the Monte Carlo simulation is used to generate a new family of risk curves that compare the losses associated with DLP with and without DLP, which is shown in figure 39. A low estimate (10% rate reduction) and a high estimate (80% rate reduction) are shown to develop intuition about the overall effectiveness of data loss prevention.



Scenario	Expected Value
Base	$645K
10% Reduction	$685K
50% Reduction	$458K
80% Reduction	$295K

Figure 39: Data spillage risk curves for a limited DLP initiative. The lines correspond to the base risk, as well as a 10%, 50%, and 80% reduction in the rate of data spillage incidents. The table reports the expected value for each case.

[58] Taking pictures of sensitive information is a real threat. In one organization, a username and password for a shared terminal was written on a Post-it note. A picture of the office was taken and posted on a public website, with the login credentials clearly visible.

Figure 39 shows that the yearly costs due to data spillage will be lower more than 85% of the time if DLP is NOT implemented. This is because data spillage is a low-cost vector most years. Therefore, a particularly severe year must occur for DLP to be cost-effective. A 50% reduction in the rate of incidents translates to a lower expected value of losses, primarily through a small reduction in the large impacts due to auditing and reputation damage. Note that although the expected value is smaller for the case with a 50% reduction in incidents, the losses observed will be larger than the base case the majority of the time. This is an important consideration for management.

Analyzing the results of the Monte Carlo simulation, a limitation with using the expected value can be observed as well.[59] Each scenario involves a small chance of a very high loss, so the expected losses are much higher (in some cases 2 orders of magnitude) than the losses that will be observed 90% of the time. Therefore, the ability to analyze the full risk curve is useful to a decision maker because low, frequent losses can be explicitly compared with severe but infrequent losses.

Overall, DLP will cost an organization more money than it saves most of the time. Sensitivity analysis can be used to determine that a limited DLP rollout would be cost-effective if it reduces the rate of incidents by 20% or more. Given that the current estimate involves a 50% decrease in incidents, the limited deployment is a good investment. However, this recommendation is very sensitive to the cost of DLP software, meaning that every $50,000 increase in the price of DLP requires an additional 10% improvement in the overall effectiveness for a risk neutral decision maker. Therefore, the extended DLP enrollment is not cost-effective.

Other considerations

A large number of other factors should be considered when considering DLP software. For example, care should be taken to avoid DLP implementations that are severely disruptive to work flows. DLP might flag documents with nine-digit numbers as social security data, although many zip codes are nine digits as well and should not be blocked. Users have also been known to take shortcuts if security controls are too burdensome, and may not implement DLP effectively. Finally, the ease with which DLP can be defeated by a malicious insider or an adversary should be studied. For example, criminals will routinely encrypt data before exfiltration it so that security systems cannot detect what is leaving the network. Given these considerations, DLP software is close to a break-even implementation, meaning that a more detailed model may be required.

[59] Adding a risk tolerance would make the prospect valuations more accurate, but does not alleviate discrepancy between frequently observed losses and the value of the alternative, since obtaining a single value for an uncertain prospect with heavy-tailed outcomes is inherently difficult.

5.4.5 Regulatory Changes

Nearly all aspects of the cyber security landscape are rapidly changing, and compliance requirements are no different. Space Corp is considering the scenario where its main customer mandates a full-scale DLP software requirement. Space Corp has already calculated that the extended deployment is not cost-effective, but would like to study how this security requirement could impact its decision. Initially, reputation damage was assessed as a 50% chance with no additional costs, a 45% chance of auditing costs, and a 5% chance of lost future business. However, the new compliance requirements change these prospects, since it becomes much more likely that any data spillage incidents will lead to auditing and reputation damage. Figure 40 shows the newly calculated risk curve under a mandate for DLP software, which corresponds to a 100% chance of reputation damage given that a large data spillage incident occurs.

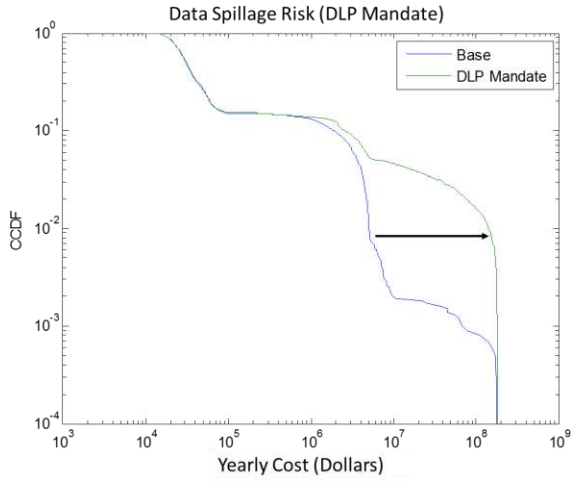


Figure 40: Risk curve with under compliance mandates. The base risk of data spillage can be compared to losses when DLP is a compliance requirement, meaning costs will be higher if a severe incident occurs.

Analyzing the risk curve, decision makers can see exactly where the additional risk occurs, which happens to be in the tail of the distribution. Data spillage incidents are rare, meaning that the losses will be the same most years. However, if a large data spillage occurs and regulators are lying in wait, large losses can occur.

5.4.6 Insights and Conclusions

Constructing a quantitative model is insightful for Space Corp. The data clearly show that while data spillage is a relatively low risk overall, large losses can occur due to auditing and reputation damage. Further, data loss prevention is effective at reducing the number of mistakes, but is not very effective at preventing malicious insiders or some advance attackers from exfiltrating information. A full deployment of DLP is not cost-effective for Space Corp, although a limited deployment becomes cost-effective if the software can reduce the rate of data spillage incidents by

more than 20%. Finally, evolving compliance requirements may change the optimal decision, since additional losses will result if Space Corp is not in compliance with their required standards.

Data spillage risk is relatively straightforward to analyze, given the stochastic nature of incidents and the simple cost function. Other attack scenarios are more complicated, given that adversaries may be adaptive. The next section explores malicious email attacks.

5.5 Malicious Email

Some of the most impactful data breaches have occurred via malicious emails, including the Target hack[60], University hacks[61], and the RSA hack[62]. Three common types of malicious emails include worms (self-propagating programs that infect and spread through email accounts), malware delivery (where malicious software is downloaded to a user's computer once a malicious attachment or malicious link is opened), and credential theft (where the adversary attempts to obtain a user's credentials).

Adversaries obtain access to a user's machine for several reasons. Some criminals may use a compromised computer as part of a botnet, a network of computers used for denial of service activities, malware distribution, or bitcoin mining. Other adversaries target access to networks to search for credit card information or personal data. Many malicious emails that an organization receives may not involve the organization specifically. For example, certain programs are designed to crawl the Internet and identify email addresses on webpages. Companies that list the individual contact information for all of their employees will likely receive more malicious emails. Further, many criminals who compromise users are not targeting an actual business but rather its users. For example, malicious emails designed to phish a user's banking credentials may be sent to a user's work email.

Worm incidents are becoming less frequent over time as attackers pivot to other forms of cyber attacks. Malware delivery via email, on the other hand, is very common and seems to be the favored method of entry for many criminals and persistent attackers. Attackers typically have a high success rate in tricking users to click on malicious attachments or malicious links. For example, attackers send users emails about a FedEx package with an attached invoice that contains malicious software.

---

[60] A third-party contractor that ran the HVAC system was phished via email. Those stolen credentials were then used to compromise Target's network.

[61] American universities have been the target of many phishing attacks from nation states.

[62] A malicious attachment was sent to a user. The attachment was flagged as spam, but the user went into the spam folder and opened it, resulting in the infection.

Credential theft is also a common occurrence, although it is often used by lower-level actors who hijack email accounts to distribute spam. Spammers use very basic techniques, such as notifying a user that the storage limit on their email inbox has reached a limit and that the user must log into a system to approve more space. The prospect of losing email services consistently scares a handful of users to authenticate via the bogus link, allowing spammers to log into the email service and use the account to launch more spam. Phished credentials could also be used to log into an account to obtain information from the inbox, although this scenario is much more rare. Another concern is that other services will be accessed with the stolen credentials, which illustrates the importance of good visibility into the authentication mechanisms across an organization. Spammers are relatively easy to detect because a user will either get notifications that they are sending unwanted emails, or the volume of emails that is sent will trigger alarms. Other malicious email attacks are much more difficult to detect.

Malicious emails are launched by many types of adversaries, but can primarily be traced to persistent attackers, criminals, and spammers. Figure 41 shows the CCDF for email incident investigation time and divides incidents into several attacker divisions. For example, nearly 80% of incidents occur at very low severities and accrue less than two hours of investigation time. These incidents correspond to unsuccessful attacks. Often, a user will recognize that an email is malicious and will forward it to the security operations center. An investigator will delete the email, check the sender information, investigate whether any other employees have a similar email in their inbox, and possibly create a filter for other emails that contain the same content.

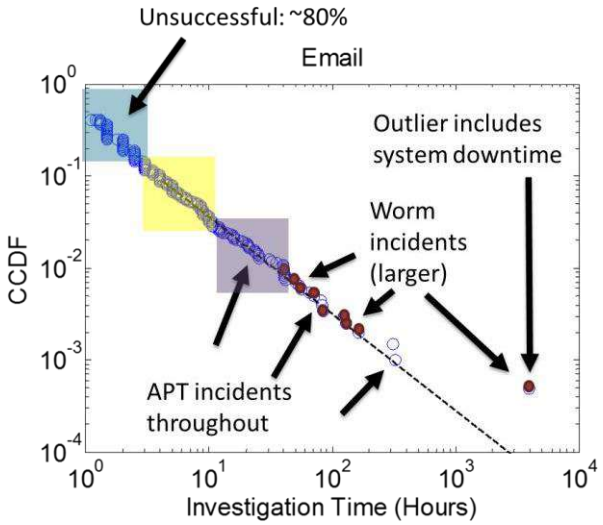


Figure 41: The CCDF for malicious email incident investigation times.

The remaining 20% of incidents range across spammers and criminals (typically moderate severity), persistent attackers, and worm incidents. Persistent attacker incidents span across all severity levels, from small to severe.

5.5.1 Analysis of Malicious Email Attacks

Space Corp recorded thousands of malicious email incidents over the six-year period. These incidents need to be cleaned and categorized to obtain a clearer picture of attacks that occur. First, unsuccessful malicious email attempts are separated. Of all of the malicious email attempts that were recorded, 81.4% took 1 hour or less to investigate, suggesting that the email was flagged as malicious and investigated, but that the email was not successful. This agrees very well with the heuristic that between 10 and 20% of malicious email attempts are successful.

Next, worms are removed. Worm incidents are a very small proportion of the overall number of incidents, but occur almost exclusively at the very high end of severities. This is because email worm incidents spread very quickly and involve a large amount of time devoted to incident remediation.

Persistent adversaries are modeled next. These attacks range in sophistication, but can deliver exceptionally targeted attacks that are tailored to a specific target.[63] Persistent adversary attacks can range in severity as well, depending on whether an attack is successful or not. Roughly 10% of successful malicious emails are attributed to persistent adversaries. Different attacker groups have different strategies. For example, some attack campaigns may use targeted emails with malicious attachments, while others may focus on gaining a user's credentials. Given the uncertainty involved with the usefulness of obtaining a user's credentials, advanced attackers are observed to favor malicious software in their attacks.

Finally, the remaining emails can be distinguished by the type of attack and often the attacker. Credential theft, where a user's password is elicited, occurs 44% of the time, while the remaining 56% of the time involves malicious software use, most often by criminals. Credential theft is typically used by criminals who use a compromised email account to send out spam to other inboxes. These attacks are easily identified by a security team because a compromised email inbox often begins sending out thousands of emails per minute. Therefore, the detection is straightforward from a reactionary perspective, and the remediation involves cleaning up the user's inbox, resetting credentials, and identifying the original compromise vector.

[63] These targeted attacks are often referred to as "spearphishing".

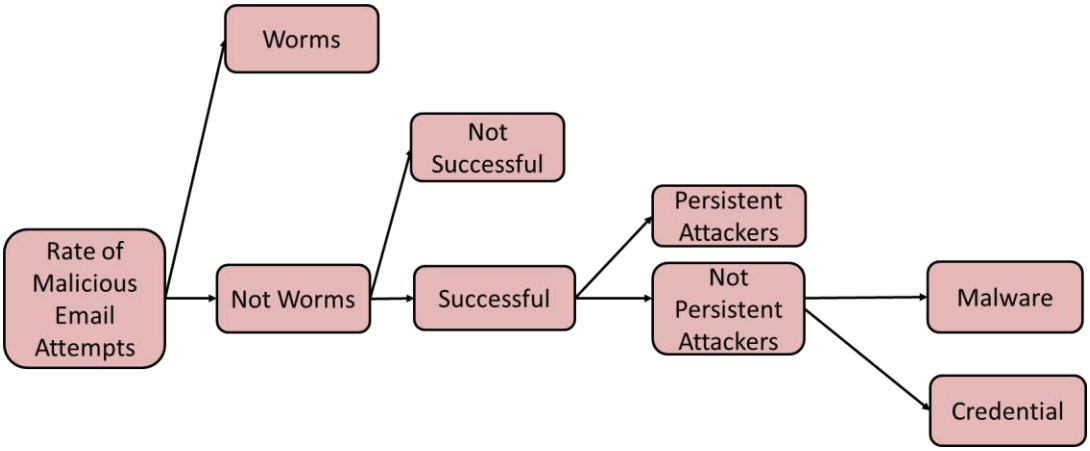


Figure 42: An illustration of email incident categorization.

Emails involving malicious software (delivered via an attachment or a malicious link that takes a user to a compromised website) tend to be more difficult for a security team to detect. Typically, malicious software will be downloaded onto a device once a user opens a malicious attachment, or clicks on a malicious link that takes the victim to a compromised website. If the user's browser has a vulnerability, malicious software can be downloaded onto their machine. Once this occurs, the malicious code may lie dormant for a period of time before beaconing out to a command and control server. This establishes a backdoor to the end device within a network, allowing an adversary to gain higher access privileges while gaining access to more parts of the network. Privilege escalation and lateral movement in a network is a complicated process. At a high level, an adversary will iterate through a series of tools and techniques, some of which will be effective and some of which will not. Throughout this process, there is a constant struggle between exploitation and being detected.

Rate of attacks

While incident investigation times for malicious emails at Space Corp have remained constant over the past several years, the rate of incidents has increased dramatically (Figure 43). From 2009 to 2015, the number of malicious emails rose from roughly 10 per month to nearly 150 per month. However, the vast majority of these additional incidents are not severe and take less than 3 hours to investigate.

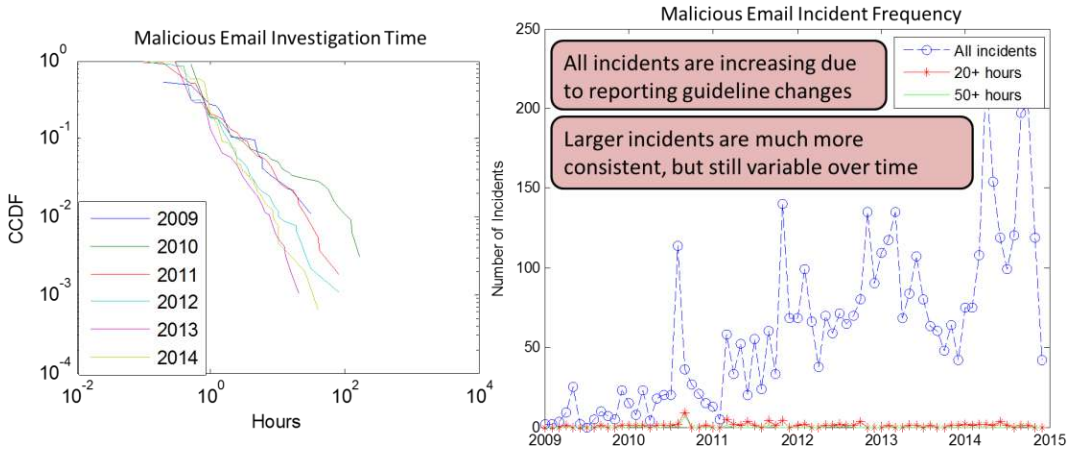


Figure 43: The distribution of email incidents. While the exact distribution varies year to year, malicious email attacks retain similar properties in their severity. The rate of low-impact malicious email incidents is increasing, but large incidents still occur at a constant rate over time.

Figure 44 shows the rate of large malicious email incidents (defined as five or more hours of investigation). Large incidents occur at a remarkably stable rate over time. This nuanced view of cyber security incidents illustrates the need to carefully assess trends. Many industry surveys report that phishing incidents are rapidly increasing and may be the most common attack vector (Verizon, 2014). While the total number of malicious email incidents may be increasing, it is important to look at the corresponding severity of those incidents as well.

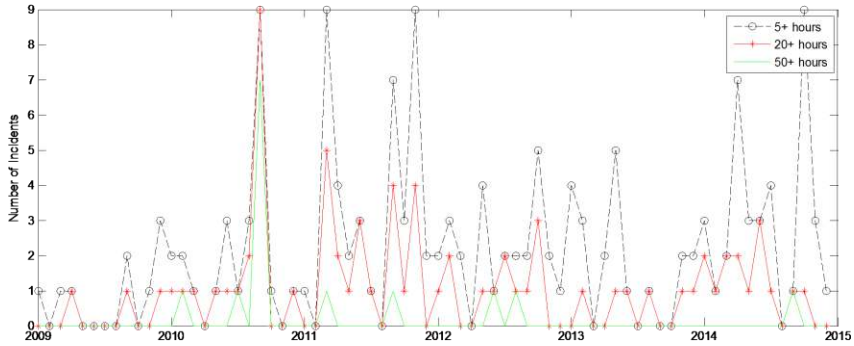


Figure 44: The rate of large incidents over time.

Malicious email compromises can be very complicated, involving large uncertainties in the probability of attack, the probability of success, the existence of further vulnerabilities that allow a user to compromise a machine, and the impacts to an organization. In situations with extreme uncertainty, it becomes even more critical to quantitatively model the process instead of regressing to qualitative methods. Incorporating the uncertainties leads to a richer understanding of the relative importance of different variables through tools such as sensitivity analysis. For example, it is difficult to assess the importance of the rate of malicious emails sent into an organization versus

the probability of success. Quantitative modeling allows a decision maker to explicitly compare how losses change in either case.

For the purposes of modeling malicious emails, several assumptions are made. First, worms are considered to be rare and infrequent, since adversaries seem to be moving to other attack vectors. Therefore, the model assumes that worm incidents will not occur over the next year (to simplify the results). Further, the rate of malicious emails will increase slowly (~5%) over the next year, while the fraction of malicious emails caused by persistent attackers, spammers (credential theft), and criminals (malware) will remain proportional to the current observations of 10%, 35%, and 45%. Threat intelligence can augment the model by tailoring the prospective evolution of malicious email attacks. For example, 2016 saw an increase in malicious email attacks against healthcare institutions, designed to encrypt healthcare data and extract ransom money. Organizations in the healthcare industry should incorporate this information into their assessments for more accurate projections.

5.5.2 Malicious Email Impacts

Malicious emails can lead to a variety of costs, including investigation time, reputation damage, business interruption, and intellectual property loss. Direct costs and privacy information losses rarely occur through malicious emails at Space Corp, but may be more common for other organizations. For example, specially targeted phishing emails have attempted to trigger large invoice wire transfers to international criminals (Krebs, 2015b). These attacks are becoming more common against small businesses and can present substantial risks.

The investigation times are well known due to historical data, and the same reputation damage model used for data spillage is used for malicious emails. Each large malicious email incident may result in audits (uniform distribution), lost business (beta distribution), or no costs. Based on historical data, the threshold for the additional costs is much lower for malicious email incidents (200 hours) than for data spillage incidents.

Malicious email attacks involve two new cost vectors that need to be modeled, namely business interruption and intellectual property loss. Business interruption is highly dependent on the organization type. For example, some organizations are service providers and lose revenue when content is not delivered. Netflix, cloud providers, and media outlets all operate on this cost model. Other organizations may have more uncertain costs associated with business interruption events. Financial institutions that offer online banking need high availability, since outages will damage customer confidence and could translate into lost sales if a consumer chooses to take their business elsewhere. Even here, the true cost is highly dependent on the sector: for example, changing banks is labor intensive and one short-term outage is unlikely to result in lost customers.

However, shorter service lifecycles with low switching costs are more vulnerable to outages, such as in the case of hotel reservations or transportation services (i.e., Uber or Lyft).

Other types of organizations operate on an entirely different product lifecycle, with unique business interruption challenges. Email worms or large-volume phishing campaigns may require the suspension of email account services company-wide. Malware can lead to system downtime or the unavailability of services and cyber incidents can even disrupt core operations. For example, malware has disabled control systems on several deep sea oil rigs, causing the rigs to drift off their wells and bringing oil extraction activities to a halt (Zukfeldt, 2015).

Obtaining a distribution for business interruption costs requires careful input from decision makers, incorporating observations of historical incidents and the involvement of other domain experts. For Space Corp, business interruption takes the form of core services being disrupted, leading to employee downtime. Space Corp assesses that there is a small probability of a delay that causes a deadline to be missed, in which case costs would be on the order of millions of dollars. If these costs are realized, there would likely be further reputation damages, which are modeled separately. A beta distribution is used to model the business interruption costs.

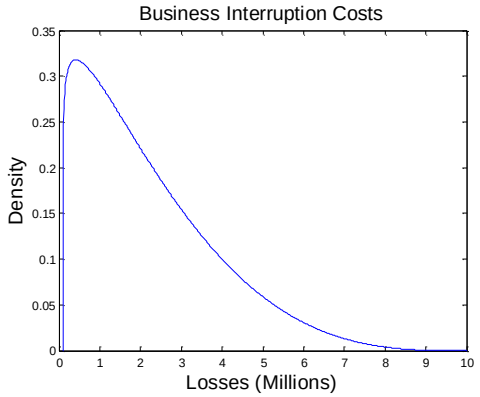


Figure 45: Business interruption costs for Space Corp.

Intellectual property (IP) losses must also be modeled. Currently, there is much disagreement about the best way to model IP losses. IP losses are actually focused on trade secrets or manufacturing details that are not otherwise public. Patented technology and copyright infringement is a different issue, given that patents are public knowledge.

Malicious emails can lead to the loss of intellectual property. The valuation of IP is a challenging prospect in itself. It is clear the IP should not be valued based on the investment that was required for development, since the actual worth may be more or less. Different organizations will value IP differently. Some sources point to hackers' theft of IP from SolarWorld as the cause of its financial difficulties (Cardwell, 2014). However, Cisco code has been found line-for-line in

Huawei routers, demonstrating a clear case of IP infringement, but Cisco continues to retain a heavy market share.

Because of the difficulties in valuing IP at an organization, a willingness-to-pay model can be used. Here, the decision maker's preferences are directly elicited. If an organization considers their information to be especially proprietary, they will assign a higher value to it and thus put more security controls in place. However, organizations that do not place a high value on IP will likewise spend less to safeguard this information.

Eliciting IP losses can be difficult, because the losses may range across orders of magnitude. Analysts must be prepared to elicit damages on a linear-linear chart, or log-log, depending on the nature of the losses. For Space Corp, IP losses are assessed on a log-log plot based on industry data and IP theft observed at the organization. Figure 46 shows the elicitation chart, along with some questions that can be used to guide the responses.

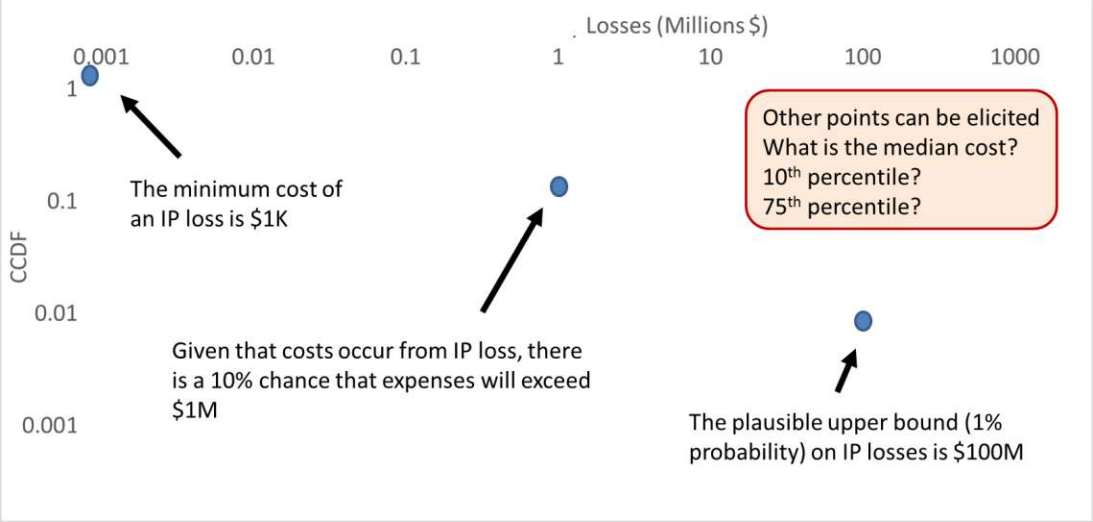


Figure 46: An example of eliciting the monetary losses associated with IP loss. Note that the elicitation takes place across orders of magnitude. These elicitations are used to create a family of loss curves that can be sampled from.

5.5.3 Simulating Malicious Email Risk

The assessments pertaining to the frequency of malicious email incidents and cost models is used to initialize another Monte Carlo simulation (figure 47). The rate of incidents is projected by experts to increase 5% over the next year. The investigation times are drawn from the observed distribution of historical incidents. Other costs occur with an investigation time trigger of 200 hours (assessed from the historical data). Reputation damage is modeled with the same parameters found in the data spillage case study (50% no additional cost, 45% audit cost from uniform $1–2 million, and 5% reputation damage from a beta distribution ranging from $100 to $160 million). Business interruption is modeled with a 200-hour trigger and results in a 50% chance of losses drawn from

a beta distribution between $100,000 and $10 million. Intellectual property losses are drawn from a power law elicited from decision makers with a maximum loss of $10 million.

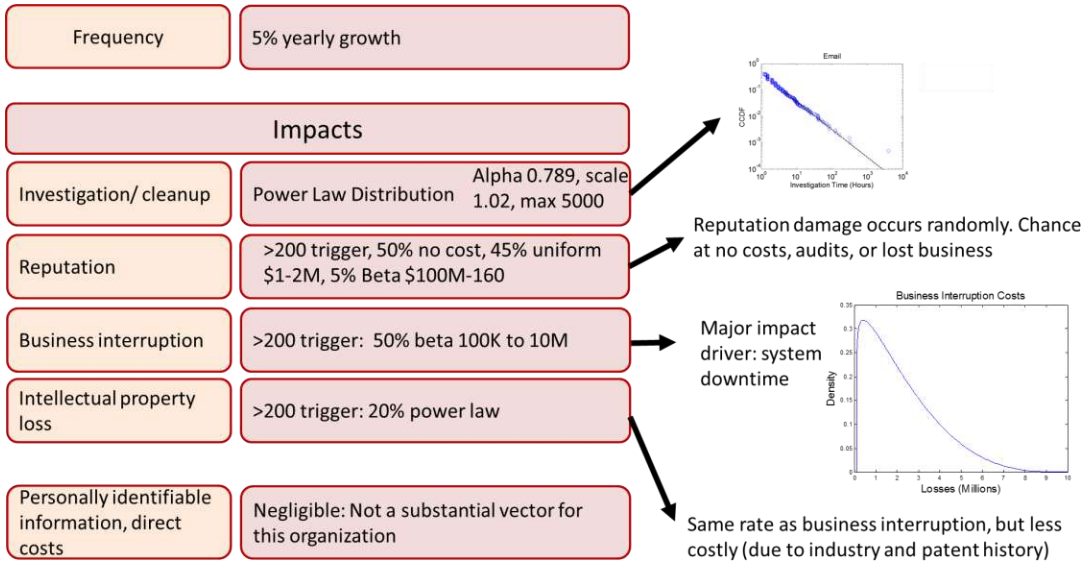


Figure 47: Model inputs for malicious email risk model.

Risk Curves

Figure 48a shows the risk curve for malicious emails. It is immediately clear that malicious emails are significantly more costly than data spillage incidents. While only 20% of data spillage incidents cause more than $20,000 in losses, the top 20% of malicious email incidents cause $3 million or more in damage. This is in part due to a rapid acceleration of costs caused by reputation damage and business interruption. Figure 48b shows how the results are different when worm incidents are added back into the simulation; investigation costs would rise, but there is little change in the tail risk.

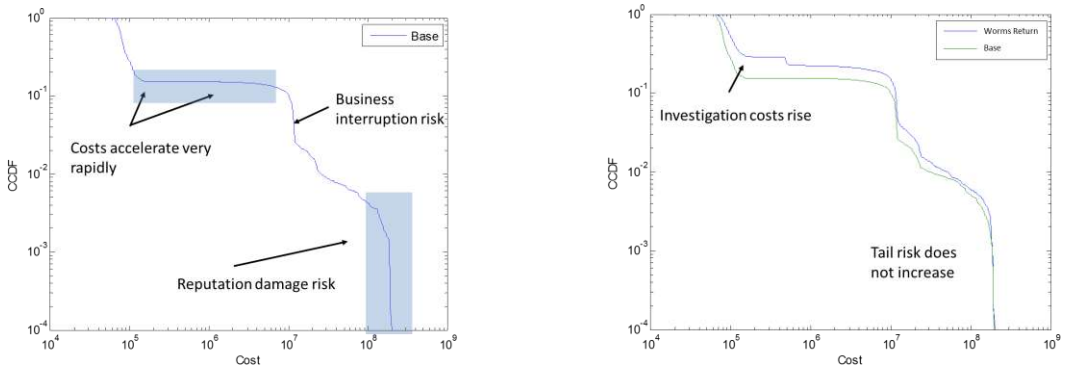


Figure 48: Risk curves for malicious emails (a), and analysis of worm incidents (b).

5.5.4 Malicious Email Safeguard Modeling

The first line of defense against malicious emails involves blocking them before they get to the user. Many enterprise technologies exist to filter spam or malicious attachments. Other safeguard technologies are not specific to malicious emails, but are general security features that reduce the probability of a successful compromise from several vectors. For example, sinkholes identify compromised machines that are attempting to communicate with known malicious IPs. These communication requests are blocked, which triggers an alert to the security team. These IP blacklists reduce the probability of malicious email compromises, web attacks, and other malware incidents.

One of the most effective safeguards against stealing credentials is two-factor authentication (TFA). Instead of requiring a simple password to access a system (which can easily be observed, intercepted, or even elicited via a phishing email), TFA requires another form of verification from the user. Many TFA solutions exist on the market to suit a variety of needs.

***Tokens*:** One of the most common forms of TFA is the use of physical security tokens. These devices are carried by the user and display a unique string of characters that change rapidly (usually every 30–60 seconds). A user authenticates using their password and the token string.

***Mobile text/Duo App*:** Several TFA mechanisms interact with a user's mobile device. For example, a text message with a unique string can be sent to the user's mobile device, or an application can be installed that requires a user to push a button on their mobile device to authenticate. While some security professionals have expressed concerns regarding an adversary's ability to compromise a mobile device, mobile TFA is a convenient and cheap solution that can defeat a wide range of adversaries.

***PIV Smart Cards*:** Another technology that is widely being considered is a personal identity verification (PIV) smart card, which often takes the form of an ID and a chip that cryptographically authenticates. While this technology is another effective form of TFA, it can require costly infrastructure (namely card readers installed at all endpoints). For example, most tablets do not have PIV smart card readers built in, requiring an additional piece of hardware.

Two-factor authentication primarily impacts the severity of incidents. Incidents continue to occur at the same frequency, but the impact will be reduced. For example, if a user exposes their credentials through a phishing email, the security operations center still needs to document the

incident, reset the user's password, and verify any malicious failed login attempts. In this case, TFA limits a potentially severe incident to a very small investigation.

Given the wide range of attackers and attack types for malicious emails, organizations could choose to develop very detailed models to assess the effectiveness of TFA. For the purposes of this case study, several simplifying assumptions are made to make the model clearer. First, Space Corp experts assess that the cost of implementing TFA is $500,000 per year. In reality, there will be a larger upfront fixed cost to develop the infrastructure and an additional annual cost for software licenses and support. However, $500,000 per year is a good estimate for Space Corp.

Malicious email incidents involve a range of attacker types, meaning that two-factor authentication may have a different effectiveness rate against different types of adversaries. Space Corp models the effectiveness of TFA by first analyzing the proportion of malicious email incidents attributed to three attacker types: spammers, criminals, and persistent adversaries. Next, the effectiveness of TFA is assessed against each of these adversaries. For example, 35% of malicious email incidents are attributed to spammers and 90% of spammer attacks are defeated using TFA. Each of the incidents that are defeated are then limited in impact. For example, Space Corp assumes that the 90% of malicious email incidents due to spammers will be resolved in one hour. The total cost due to malicious email incidents is then calculated via the standard process. Figure 49 shows the assessed TFA rate reduction for each attacker type, and the new risk curve when TFA is implemented.

Attacker	Proportion of current incidents	TFA rate reduction
Spammers	35%	90%
Criminals	45%	20%
Persistent Attackers	10%	20%

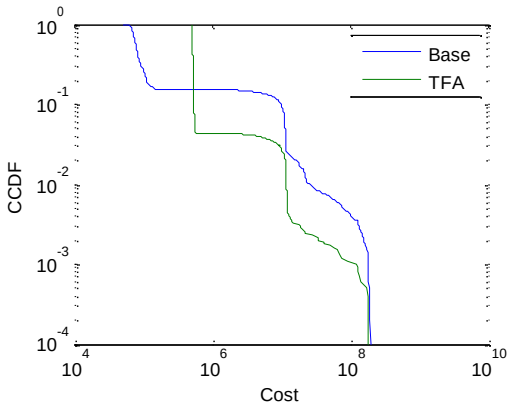


Figure 49: Risk curve for malicious email with different effectiveness rates of TFA.

Several important insights can be derived from these model runs. First, it is clear that while the small increase in malicious email attacks is important to consider, it is not a major driver of the risk. In fact, sensitivity analysis shows that the growth rate is hardly noticeable in the short term, meaning that the results are very robust to assumptions about the increase in malicious email attacks. Analyzing the cost-effectiveness of TFA, several other insights are apparent. TFA requires a significant investment and immediately alleviates some investigation costs, but not at a sufficient

level to make the technology immediately cost-effective. However, TFA reduces the variability in investigation costs a large amount while also significantly reducing losses by between $1 million and $10 million. The risk reduction is also considerably larger than DLP's effect on data spillage incidents because malicious email incidents tend to involve larger monetary losses than data spillage incidents. Based on the estimate of TFA costing $500,000 per year and a careful analysis of the risk curves, Space Corp finds that TFA is a cost-effective technology. The yearly cost is somewhat limited by the reduced investigation times and the tail risk is improved as well.

Although the analysis shows that TFA is cost-effective in this case, it is important to emphasize that the details matter. TFA can be implemented smoothly or poorly, and the user experience can cause the entire investment to be bad. For example, it is critical that the TFA codes are implemented with a strong random number generator. Google authentication codes have been reported to never start with a zero, making it easier (but still very difficult) for an attacker to guess the rest of the code (Dmitrienko, Liebchen, Rossow, & Sadeghi, 2014). Other bugs can exist in the implementation. For example, if a code is not used, it may remain valid for an extended period of time instead of expiring, leaving the system open to additional attacks.

Usability concerns are also important to consider (De Cristofaro, Du, Freudiger, & Norcie, 2013). TFA mechanisms that rely on a mobile device without any alternative authentication backup have the potential to be incredibly disruptive when an employee loses their device or has a dead battery. The speed of authentication also needs to be preserved, since a 30-second delay to receive and enter the code can quickly add up to many frustrated users. Finally, the security of the overall system needs to be based on the weakest link. The POODLE exploit involved forcing a connection to default to a weak encryption protocol that could be broken easily (Möller, Duong, & Kotowicz).[64] Similarly, if a backup authentication mechanism involves calling a helpdesk and answering easy-to-guess security questions, then the entire TFA authentication system becomes weak and vulnerable to many attackers.

Email Filtering and User Training

A separate strategy for eliminating malicious emails involves filtering, either at the machine level when an email arrives so that it never reaches a user, or by educating users to better identify malicious emails. The ability to filter malicious emails pre-user has resulted in one of the most interesting tradeoffs between security and usability. Malicious attachments can be almost

[64] The POODLE exploit was publicly announced in October 2014. The attack allowed web traffic to be decrypted under certain situations.

completely eliminated using sandboxes, which are virtual machines[65] that open all of the attachments of inbound email in a controlled environment. If the attachment is malicious, the virtual machine is infected, but the intrusion is contained. The email can be filtered and the user is not impacted. Some adversaries have created special attacks that attempt to bypass this security control by identifying whether the malware is on a real machine or a virtual machine (for example, it might look for a pattern in mouse movements indicating that a true human is on the device), and only detonate when it is on a user's machine. This one-upmanship dynamic has created a rich field of attacker-defender games that has been explored by other researchers.

Sandboxing is a useful method for preventing malicious attachments, but detecting malicious links is much more difficult due to unintended consequences. A security device that explores every link in an email will inadvertently cause many actions that may not be intended by the recipient. For example, links may automatically unsubscribe to a mailing list, confirm an invoice, or cause any number of other results. Therefore, security is often in the hands of the end user, who is trusted to make intelligent inferences about the validity of different links.

However, significant evidence exists that shows that users should not be trusted, because mistakes are inevitable. Certain phishing emails are relatively easy to recognize: they often contain bad grammar, urgent language, and links to websites that are clearly not authentic.[66] Despite these obvious signals, however, research, industry reports, and internal data consistently show that phishing success rates are roughly 20%. One vendor who specializes in training employees about the danger of phishing emails found that quarterly training sessions result in a 19% phishing success rate, bimonthly training drops success rates to just 12%, and monthly training decreases the success rate to 4% (ThreatSim, 2013). However, for large organizations, a 4% success rate virtually guarantees an adversary at least one compromise.[67]

Empirical evidence may also exist at an organization that directly illustrates the difficulty of preventing users from being compromised via malicious emails. Figure 50 shows two malicious email exercises at a large organization. Malicious emails were designed based on previous attack campaigns, and were sent to a sample of employees across the organization. Data was collected on

---

[65] Virtual machines are instances of an operating system that run on top of another machine. For example, a single laptop could run a dozen virtual machines that each look like a distinct laptop on the network. In security, virtual machines are useful due to their hierarchical controls. If a virtual machine is hacked, it can be easily deleted and replaced.

[66] Some attackers use specially crafted web addresses, such as "www.bankofthe**vv**est.com" instead of "www.bankofthe**w**est.com".

[67] For example, for an organization of 10,000 people and assuming every user clicks on a phishing email with p=0.96, the probability that at least one person is compromised is 1- 0.96^10,000, or .999, with another 174 nines. This silly example ignores real-world safeguards, such as shutting down email servers in the case of a large attack, but nevertheless illustrates the difficulty in preventing attacker success.

the proportion of employees that opened the bogus email (which in certain limited cases could have led to a successful compromise), and the proportion of employees that clicked on an attachment or entered their credentials.

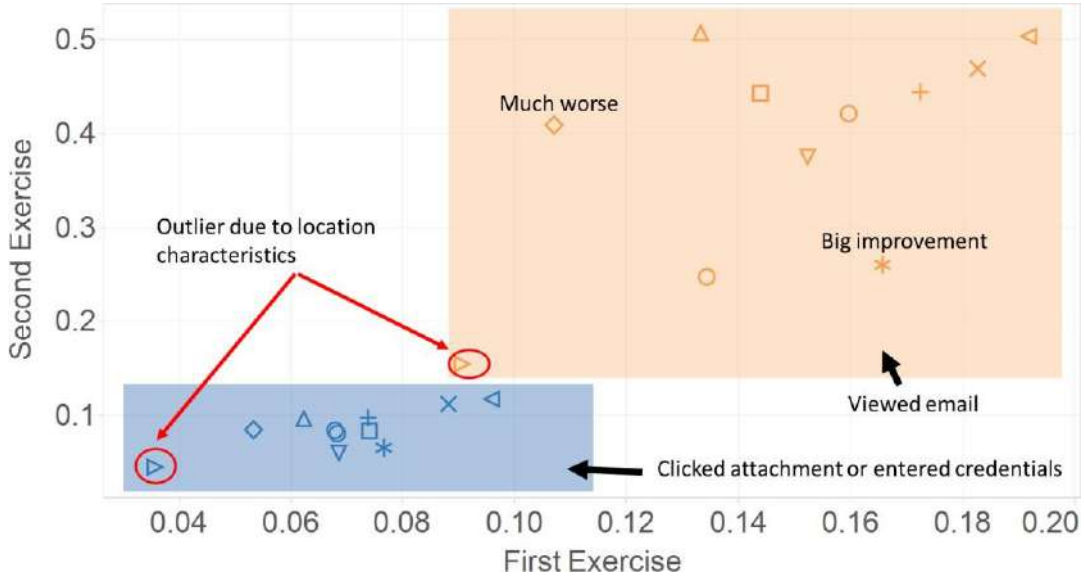


Figure 50: Data from two malicious email campaigns used for training purposes. Note that the compromise rates are between 4% and 15% for users who clicked on the attachment or entered their credentials.

As discussed above, the effectiveness of email filtering is largely dependent on the core technology and what type of emails are being sent into the organization. For example, targeted emails are much more difficult to filter compared to spam emails that are sent to a large number of recipients.

The majority of phishing success rates are between 30% and 5%. User training could reduce this rate, although the effect of training is likely to wear off over time. Organizations need to balance the frequency, effectiveness, and cost of user training. Based on conversations with security professionals, Space Corp could implement a yearly employee training program at a cost (curriculum development plus lost productivity of trained employees) of $100,000. The training is forecast to reduce the phishing success rate from the current 20% to a rate of 15%. Figure 51 shows the new risk curve under these assumptions.

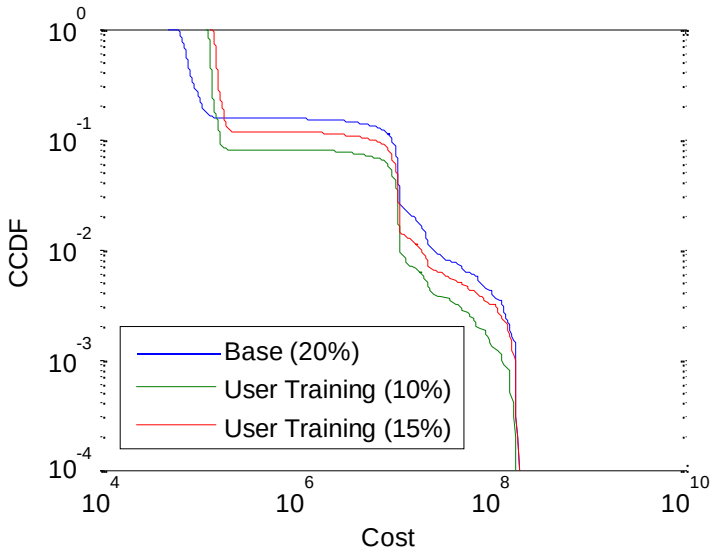


Figure 51: Cost-effectiveness of user training.

User training has the same tradeoffs observed in other safeguards, consisting of a higher rate of losses during most years but a smaller risk of large impact incidents. The exact effectiveness of user training is highly uncertain,

moderate risk reduction.

Malicious email attacks a

a certain proportion of the popul

given time and will open an ema

anxiously anticipated. Attackers

emails. At one organization, adv

visited another corporation and se

even contained an explanation fo

LinkedIn, a social networking sit

identify the administrative assista

of targeted attacks may be too co

exploit this critical business func

good security investments for Spa

a costly attack vector. New pre

however, to further secure the org

5.6 Websites

Websites are public facing and are the front door for cyber attacks. Organizations use websites to satisfy a number of objectives such as advertising or educational outreach, or to satisfy some core business component such as online banking or remote email access for employees (webmail). Websites can be difficult to fully secure. Developers will inevitably make errors leading to misconfigurations that leave websites vulnerable. Patching is difficult to implement with high fidelity as well, given the steady stream of vulnerabilities that occur in many applications. Overall, website security can be a major challenge.

Several types of attacks can be launched against a website. Criminals often deface websites, which can result in minor reputation damage but is more analogous to graffiti on the side of a building (an eyesore without any actual structural damage). Organizations that use websites to interact with their customers can be vulnerable to specially crafted attacks designed to steal information from databases. For example, standard query language (SQL) is a method for interacting with data. Specially designed queries can sometimes inject code (a process known as SQL injection) to extract data from a system.

A number of other attack techniques exist against websites. Cross-site scripting (XSS) injects malicious code into a web application. Brute force attacks can gain access to exposed login pages by trying all user and password names exhaustively. Directory traversal attacks are used when an adversary jumps to an unauthorized portion of a web server by guessing the file hierarchy.

Website attacks generally result in one of three outcomes. Lone hackers and hacktivists generally try to deface a website, making detection very easy. Other lone hackers or criminals might attempt to extract information from the server or database either to demonstrate their skill or to search for data that can be sold. Finally, a subset of attackers will attempt to compromise a website to gain a foothold on a network to establish persistence, extract intellectual property, or cause damage to a network. Once access to a web server is gained, the attacker can install backdoors, upload privilege escalation tools, and move laterally through the network.

Websites are attractive targets because of the low resource investment required to attack them, because they are always exposed, and because many websites have a steady stream of vulnerabilities. Attackers can immediately attempt many exploits against a website without having to invest time into profiling an organization or learning background information that would be required for a targeted phishing attack. Many popular web application development platforms, such as Adobe's ColdFusion or Apache HTTP web servers have numerous bugs that are found regularly. If a development team is not closely monitoring vulnerability alerts and patching quickly, outdated software can quickly lead to major vulnerabilities. Some of the highest profile, most wide-scale

vulnerabilities have impacted web servers. In 2014, researchers disclosed the Heartbleed vulnerability, which allowed unauthorized individuals to obtain information from a web server that could be used to steal credentials and other information.

Web attacks are not just limited to web servers, but can target other network equipment as well. For example, a common method for transferring files involves FTP (File Transfer Protocol) servers. FTP servers send files between a client and a server. However, most of these servers send information unencrypted, leaving the data vulnerable to interception by an adversary. Further, many FTP servers are running outdated software that has known vulnerabilities, meaning that attackers can use a compromised FTP server to gain a foothold on the network.

5.6.1 Website Attack Frequency and Impact

Data analysis is a powerful tool, but understanding the context in which the data was recorded is also critical. The rate of website incidents at Space Corp was typically low from 2009 to 2013, except for an isolated period of time where the organization experienced incidents from a sustained attack campaign. However, in 2013 incidents began to rise, and in mid-2014 incident recording exploded. Monthly totals were over four times the historical maximum. While these trends seem to signal a huge increase in the number of attacks, the changes were actually driven by management issues. Prior to 2013, attempted intrusions were typically not recorded in the data. New incident reporting guidelines were issued in that year, however, leading to a dramatic increase in the number of recorded incidents, mostly driven by attempted intrusions.

Figure 52a shows that over 95% of website incidents take less than two hours to investigate. 46b shows the investigation time by year, and the number of incidents per month. Note that the tail section of investigation time is roughly the same each year, but offset due to due different amounts of small scale incidents. The two years with the greatest offset (2013 and 2014) also correspond to the large increase in incidents per month.

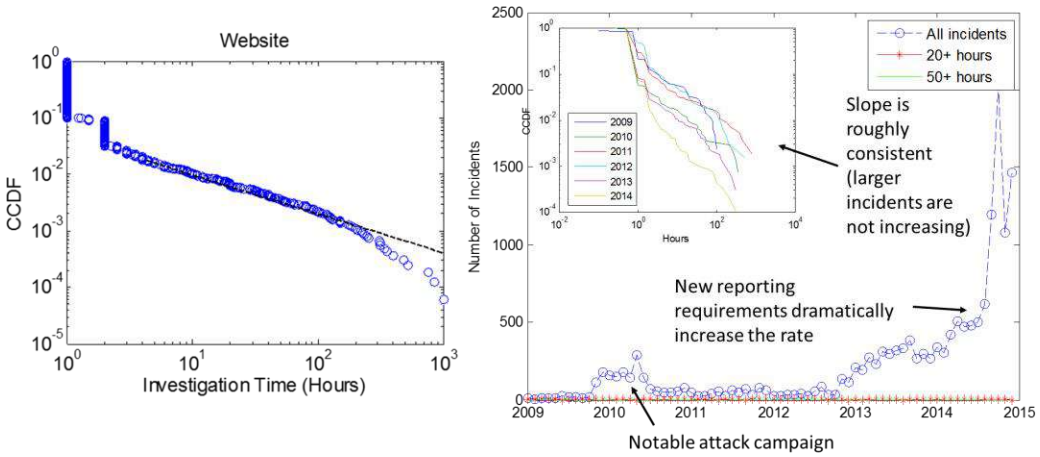


Figure 52: The CCDF for website incident severity (a), and the rate of incidents over time (b).

Analyzing the data in another way, it is clear that the overall change in the number of incidents is driven by small incidents. When small incidents are filtered out, the rate of larger incidents is relatively constant over time. Figure 53 shows that large incidents did not become significantly more frequent during this historical time period.

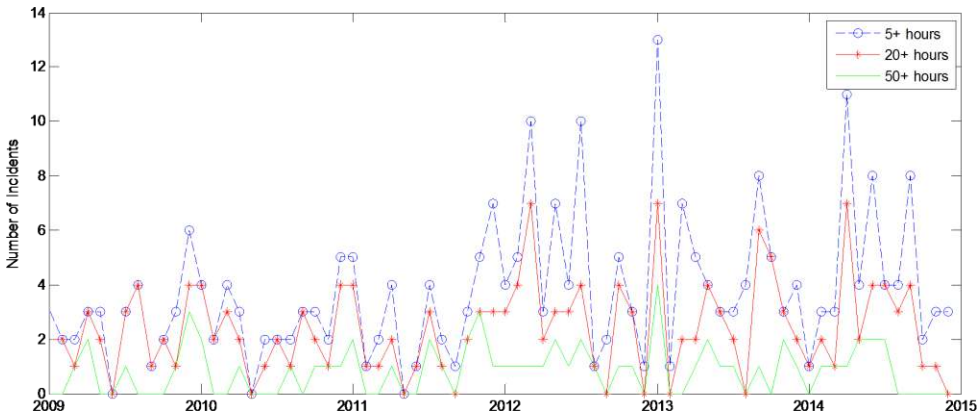


Figure 53: The rate of large website incidents over time.

The rate and severity of incidents over time presents a unique view into cyber security at an organization. One organization had the perception that website attacks increased when media attention was high, and would often staff the Security Operations Center with more personnel in anticipation of extra attacks. However, the data made it clear that attacks did not correlate with media attention, allowing the organization to scale back the surge personnel.

5.6.2 Website Attack Modeling

Due to the change in reporting guidelines, the rate of initiating events, including SQL injections, XSS attacks, brute force attacks, or others, is not well known. Instead, the modeling of the impacts is emphasized because of the higher-quality data.

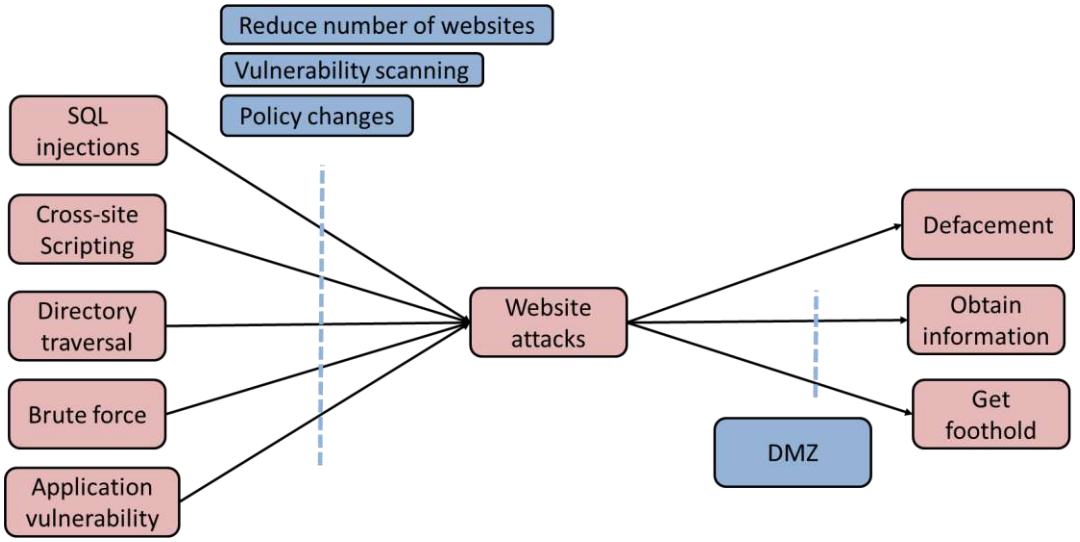


Figure 54: A general model for the website attacks and defenses.

Space Corp is considering several website safeguards to improve security. Other safeguards are available as well, but Space Corp has identified three investments that are of particular interest.

Website Footprint Reduction: Some organizations have a large number of websites, many of which are legacy pages that are no longer maintained. For example, academic institutions are home to many research groups that are constantly creating websites. When a faculty member leaves the university or a student group disbands, the website persists with no owner. These untended, unpatched websites are an excellent entry point for attackers. An effective management strategy could require that all websites be audited periodically so that legacy websites are removed in a timely manner.

Vulnerability Scanning: In conjunction with removing legacy websites, a vulnerability scanning process can significantly reduce the number of website exposures. Nearly two months after Heartbleed was disclosed, roughly 1.5% of the most-visited 800,000 sites were still vulnerable to the exploit (Leyden, 2014). Patching systems quickly is a critical component of website security.

***DMZ*:** One of the most effective website controls involves setting up a DMZ (demilitarized zone) for websites. In some organizations users can set up web servers on the core network. If that server

is compromised, an attacker essentially has access to the core network and can move laterally without significant difficulty. A DMZ aggregates web servers on a specially segmented portion of the network. These websites are typically easier to manage and protect due to their central location. Additionally, any attack that compromises a web server will be unable to move easily to the core network, since additional security controls will be in place. DMZs do not prevent website defacement, however, and have a limited ability to prevent information dumps. However, they can be very effective at limiting lateral movement.

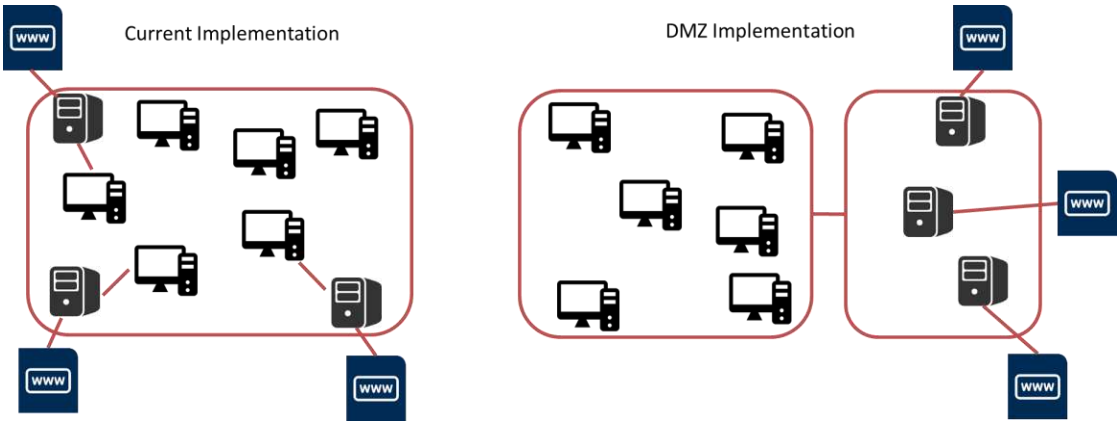


Figure 55: A schematic of a DMZ implementation. Note that the web servers are isolated from the core network, so that an attacker who compromises a web server is contained to a certain part of the network.

5.6.3 Website Impacts

Website outcomes can be broadly classified into different scenarios such as website defacements, information dumps, or network intrusion attempts. Each scenario has a typical set of impacts associated with it. As discussed earlier, the impact of website defacements is similar to graffiti on a building, or throwing a brick through the window of a store. Some investigation cost is required to identify how the attacker compromised the system and to ensure that the attack did not place any persistent code on the server or gain access to other parts of the network. Some business interruption costs are possible as well. For example, a Texan hacked into a NASA system and prevented 3,000 users from accessing science data for several days. NASA estimated the total losses at over $66,000 (Martin, 2013). Reputation damage can also result from website defacements, but the level of damage is largely dependent on the media cycle. For example, in 2015, the Syrian Electronic Army hacked www.army.mil, the US Army's official web page, and briefly posted a critical message on

the website (Weise, 2015). This incident received considerable news attention, whereas little attention was paid to at least twelve other defacements of US army webpages from 2010 onwards.[68]

Organizations should determine the most useful modeling frame to analyze different attack types. The dataset used for Space Corp does not contain sufficient data to distinguish between different website attack types. The suspected adversary is also inconsistently recorded. In certain cases, intrusions are attributed to certain groups or individuals, but the majority of incidents are not. Therefore, Space Corp chooses to model all website attack impacts probabilistically.

Many of the impacts from website attacks can be modeled using the same distribution used for data spillage and malicious email incidents (although the investigation threshold and probability of additional costs may be modified). Figure 56 defines the monetary impacts for website attacks. Investigation time is sampled from historical event where the maximum investigation time is 1,000 hours.[69] The loss of personally identifiable information (PII) is modeled using a 400-hour threshold that triggers a 20% chance of an additional cost drawn from the PII loss distribution used in other parts of the model (a uniform distribution between $60,000 and $5 million). Intellectual property loss occurs with a 400-hour threshold that triggers a 20% chance of additional costs drawn from the power law distribution used to model email incidents. Business interruption is also modeled borrowing a similar process to malicious email business interruption impacts: a 400-hour trigger and an additional cost drawn from a beta distribution ($100,000 to $10 million).

Direct costs are modeled uniquely and come from historical data combined with expert assessments. This scenario also represents the plausible worst case scenario where an adversary is able to gain a foothold on the Space Corp network, move laterally through the system, establish control over critical infrastructure (satellite control equipment in this case), and cause actions that lead to the loss of a spacecraft. Experts assess that this scenario is relatively unlikely (3% chance per large incident), but would result in substantial costs to the organization. Finally, Space Corp assesses that the reputation damage associated with website incidents can be modeled using the same reputation damage process used for data spillage and malicious email incidents (although with a 400-hour trigger).

[68] See zone-h.org for archived defacements.
[69] Two larger incidents were removed because the recorded investigation time included system downtime, which is classified as business interruption in the model.

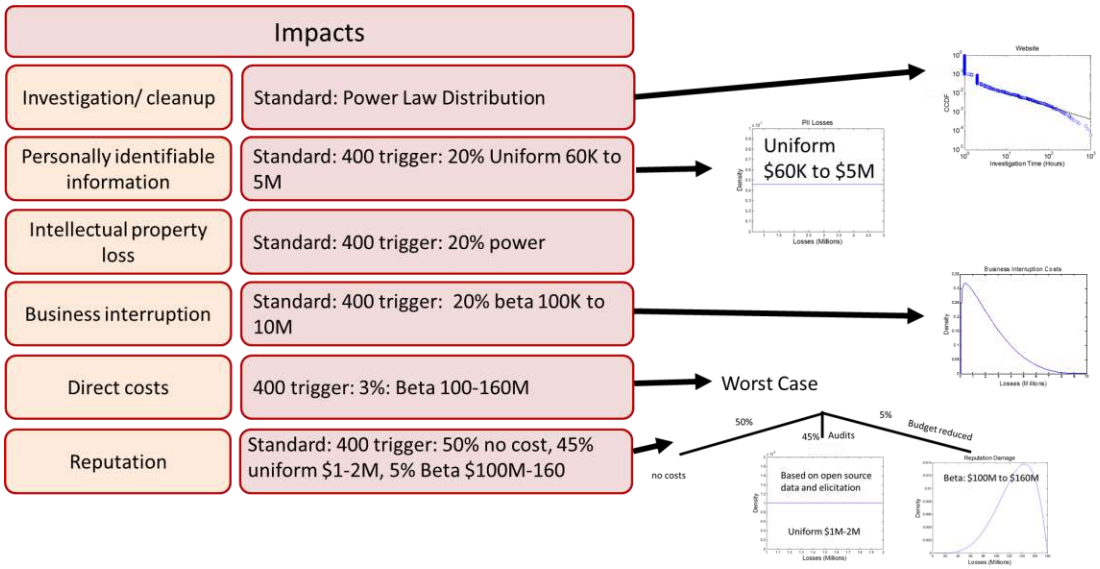


Figure 56: Website impacts.

Using all of the information above, a Monte Carlo simulation is run to assess the impact of website incidents. It is immediately clear that website incidents are the costliest attack vector for Space Corp. Attacks occur frequently and at all scales. The median cost per year is over $2 million, and there is a substantial likelihood of impacts ranging between $5 million and $50 million.

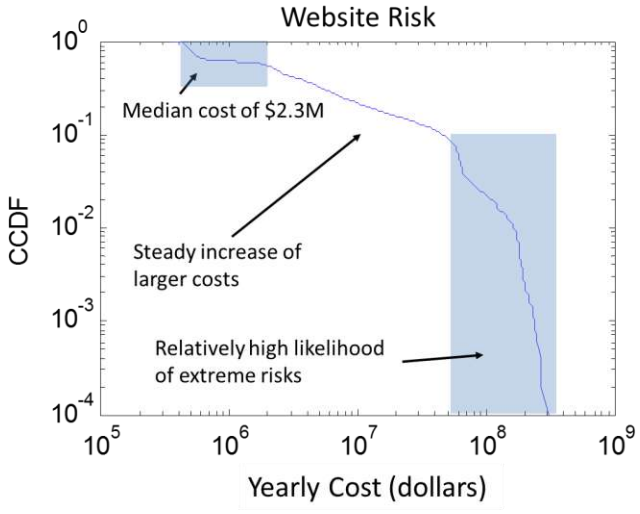


Figure 57: Website risk curve.

The high risk of websites for Space Corp does not come as a huge surprise. The historical data clearly show a consistent pattern of costly large-impact incidents originating via websites. Sensitivity analysis on several of the model parameters can be run to confirm these results. For example, figure 58 shows a sensitivity analysis on the investigation threshold. The risk curve corresponding to an investigation time trigger of 500, 300, and 200 hours is observed. While the threshold shifts the risk curve up and down, sometimes changing the losses by millions of dollars,

the overall risk consistently exceeds the losses due to other incident types by a wide margin. In fact, it is difficult to construct a plausible scenario where website incidents are not the major risk driver for Space Corp.

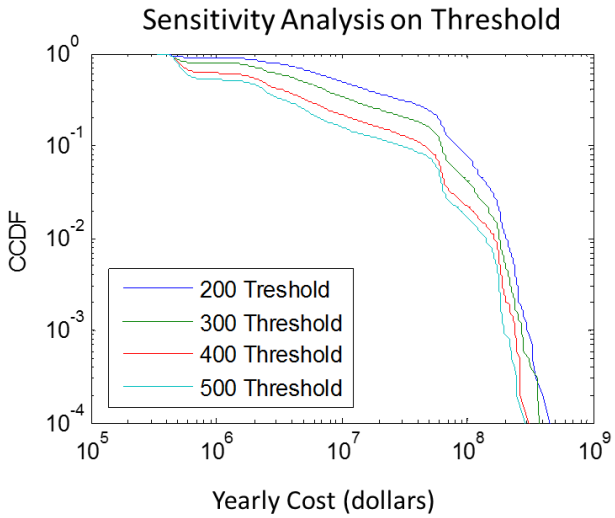


Figure 58: Sensitivity analysis for trigger threshold.

Sensitivity analysis allows an analyst to confirm that the results of the model are robust. This also demonstrates the value of quantitative risk analysis in the absence of perfect data. Significant uncertainty surrounds all of the assessments that populate the model. Despite this, Space Corp can be confident about the rank ordering of the different risk vectors.

5.6.4 Website Safeguards

As described earlier, Space Corp is considering several different website safeguards, including reducing the number of websites, implementing a more advanced vulnerability scanning and patching program, and implementing a DMZ. These safeguards impact the risk model by reducing either the rate of incidents (website removal and vulnerability management) or the impact (DMZ).

Experts at Space Corp assess that website reduction could be a high leverage activity, given that a large number of outdated, unmanaged websites remain on the network. Roughly 50% of these websites could be eliminated or aggregated to a centrally managed web server, leading to a 40% reduction in the rate of website attacks. The estimated cost of this website deprecation is $400,000.

Implementing a DMZ is a costlier prospect, given that a major restructuring of the network would be required. In this case, Space Corp is considering a limited-scope program that would relocate public-facing websites to a segmented network so that compromised machines would be prevented from gaining access to the core network. This restructuring is non-trivial and risks may still remain from websites that were not identified and transferred to the segmented network, and from misconfigurations that advanced adversaries might exploit to move past the segmented

network. However, experts assess that a significant risk reduction would still be accomplished with the new network structure. The cost of this program is estimated at $2 million, but would reduce the probability of personally identifiable information loss, intellectual property loss, business interruption costs, direct costs, and reputation damage by a factor of two.

Figure 59 shows the risk curves with different website safeguard strategies. Website reduction is clearly a good investment, and one of the most cost-effective safeguards that Space Corp could implement. The savings in figure 59 are also underestimated, since removing old websites is required only once. In future years, the savings are realized without the $400,000 investment, meaning that the savings rapidly pay for themselves. Implementing the DMZ requires a substantial upfront investment, but reduces tail risk more effectively than any other safeguard considered thus far. The DMZ may not prevent a million-dollar impact in its first year, but within five years it is very likely to have prevented a very large incident.

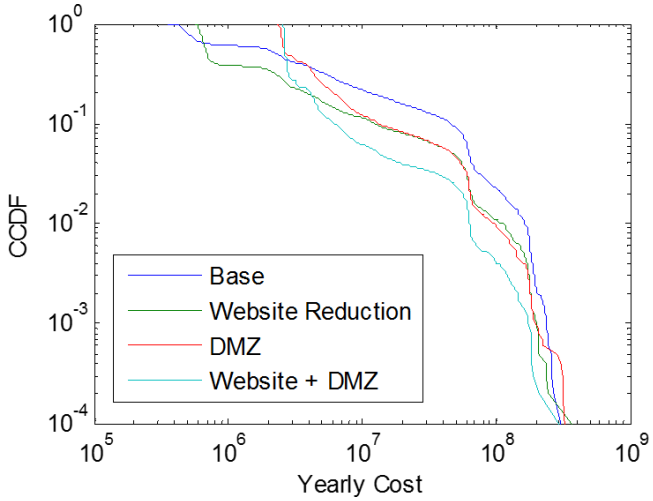


Figure 59: Website safeguard risk curves.

Data-driven methods clearly show that website attacks have been the costliest vector at Space Corp during the considered time period. Other sophisticated attacks often get much more attention (for example, malicious insiders, supply chain attacks, or targeted phishing emails), but the evidence shows that the mundane website vulnerabilities require more attention. Website attacks occur both at a relatively high frequency and at all severity levels. The power law exponent for website incident investigation time is very small, meaning that the distribution is extremely heavy-tailed. In fact, the exponent is so small that the incident investigation time has no mathematically defined mean or variance.

Despite these issues, several safeguards have proven success against website attacks. Reducing the number of websites at Space Corp is one of the most cost-effective investments that

could be implemented. Implementing network segmentation is also a good investment, despite the large cost and complicated restructuring that would be required. Implementing email safeguards would improve Space Corp's security by a large margin.

5.7 Lost and Stolen Devices

Laptops in organizations (industries, government, academic and medical institutions) are stolen or misplaced routinely.[70] Lost devices can come at a high cost to organizations for many reasons: devices must be replaced, employee productivity is disrupted, and reputation damage or credit monitoring costs may be required in the case of lost devices without encryption that contain personally identifiable information. If a determined adversary obtains access to a lost laptop, it is trivial to bypass the operating system authentication and access all the information on the device. As a result, almost all US states have laws requiring the notification of people whose personally identifiable information is on a lost laptop, and broader federal laws are currently being drafted.[71]

Data breaches resulting from stolen or misplaced laptops are frequently in the news, but because the information about the actual loss is seldom shared, the probability distribution of the costs of lost laptops is unknown. Further, this distribution has a heavy tail, meaning that many small losses occur frequently, but very large losses also occur (although more rarely). Therefore, the average is an insufficient measure of the losses associated with lost devices. Probabilistic methods provide a more precise picture of losses, and can be used to compare the value of different safeguards, including full disk encryption, asset recovery, theft awareness training, and data backup programs. The probabilistic methods require data and information from different sources (e.g., statistics and expert opinions) that are generally accessible in-house.

Lost laptops have resulted in large data breaches at many organizations. In 2010, a laptop stolen from a government contractor contained the social security numbers, names, and addresses of over 207,000 individuals (Krebs, 2010). In 2013, 74,000 records were lost by Coca-Cola on an unencrypted laptop. Although the laptops were later recovered, a year of credit monitoring services was offered to each of the affected employees (Esterl, 2014). Advocate Health Care had four unencrypted laptops stolen in 2013, resulting in over 4 million records being breached (McCann,

[70] Lost devices include stolen and misplaced devices. For example, stolen laptops are most frequently taken out of cars or residences through forced entry. Misplaced laptops occur when an employee leaves a laptop in a plane, cab, or other location and is unable to recover the laptop. Misplaced devices may be stolen if they are left in insecure places. These thefts are opportunistic and unlikely to be the result of a targeted attack. The distinction between targeted and non-targeted losses becomes important when persistent adversaries are concerned.

[71] Some sectors have existing data breach notification requirements, such as the financial sector (Gramm-Leach-Bliley Act) and healthcare (Health Insurance Portability and Accountability Act HIPAA). For information on State laws, see National Conference of State Legislatures, "Security Breach Notification Laws," available at http://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx.

2013). Dozens of other data breaches resulting from laptop theft are documented in the Privacy Rights Clearinghouse database.[72]

5.7.1 Laptop Theft

Laptop theft is recognized as a concern by many organizations, but a limited amount of work has been dedicated to the study of the rate of laptop theft or the impact of safeguards. Dimkov, Pieters, and Hartel studied the effectiveness of security cameras and access control to limit laptop theft at universities, using data from reported stolen laptops and penetration tests (Dimkov, Pieters, & Hartel, 2010). Kitteringham provided a review of industry statistics of lost laptops, discussed impacts and safeguards, and called for additional research (Kitteringham, 2008).

Many surveys and cyber security reports present summary statistics about laptop theft, but the value of these data is limited by vague definitions and insufficient metrics. Another source of information is publicly available databases, such as Privacy Rights Clearinghouse and the VERIS community database. Unfortunately, open-source data contain massive survey bias, since incidents are publicized only if a disclosure is mandated by law (usually involving personally identifiable information) or if it is sufficiently large.

Misplaced laptops are likely to continue to be an issue for organizations, given human error. Laptop theft is also likely to continue. Laptops are relatively light and usually physically unsecured, making them easy to steal. Further, a robust market exists for stolen laptops: after criminals wipe a hard drive the laptops are sold through a series of distributors (usually for about $50–$150). The combination of these facts makes them an attractive target.

5.7.2 Rate Assessment of Lost Devices

Most organizations have data that can be used to determine the rate of lost devices, but sometimes these data need to be cleaned. Figure 60 shows the cumulative number of lost devices over time, shown previously in Chapter 4. Note that the rate of lost devices was relatively constant until roughly 2012, when the rate suddenly accelerated. This transition corresponds to a change in reporting guidelines. Before 2012, only lost laptops were recorded. After 2012, mobile devices, security tokens, and other devices were recorded as well. Anomalies may also occasionally occur. For example, in 2014, a large number of devices were recorded by an organization as missing in a short time span. An investigation into these incidents revealed that an internal audit had found

[72] www.privacyrights.org

several dozen devices that had gone missing over several years but catalogued them as missing all together, leading to a large spike in the number of recorded incidents.

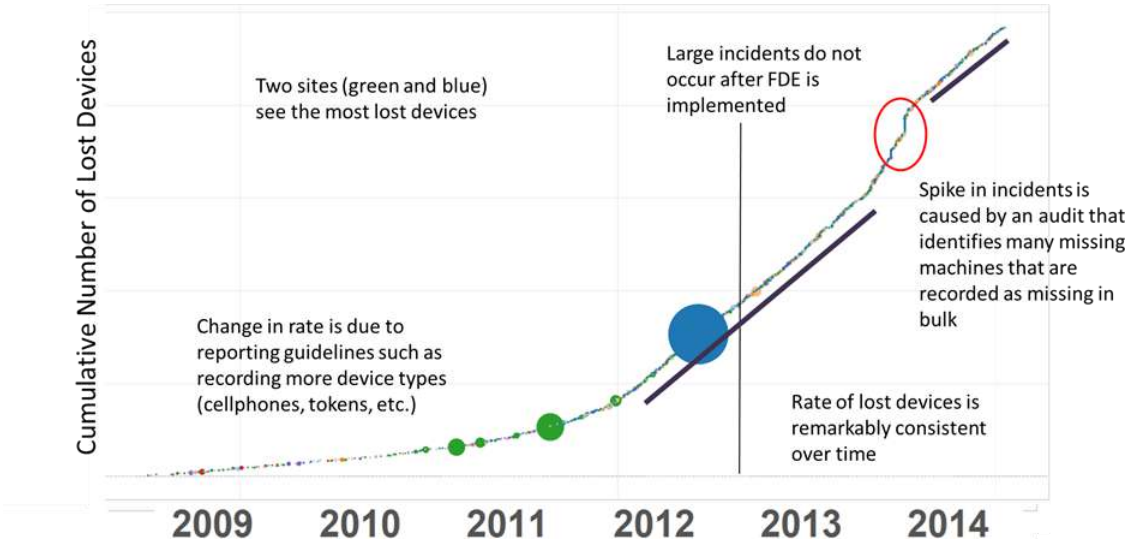


Figure 60: The cumulative number of lost devices over time.

After cleaning the data, the rate of lost devices is found to be remarkably constant over time. Other types of crime sometimes exhibit seasonal trends, but this is not the case here.[73] An analysis of lost devices at several independent organizations might show that the rate of lost devices is proportional to the number of employees at that organization and can therefore be modeled quite well.

Assessing the rate of lost devices is useful for several reasons. The stochastic nature of lost devices can be used to inform the risk analysis model. Additionally, quantifying the frequency of lost devices could help security professionals identify indications that the organization is under attack by certain groups. Without the data analysis, security operators are left to their instincts to determine whether a string of lost devices is related. Using historical data, analysts can identify clusters of thefts that are statistically unlikely, suggesting that an adversary is targeting laptops.[74]

5.7.3 Impact Assessment for Lost Devices

The impact of lost devices is relatively straightforward and easy to model. Losses may involve investigation time, lost personally identifiable information, lost intellectual property, business interruption costs (in the form of lost productivity for the impacted employee), and direct costs (in

[73] For example, violent crime goes down in the winter months because it is too cold for criminals to be outside.

[74] Another pattern that has been theorized, but not yet identified in a real-world dataset, is the tendency of employees to "lose" their phones when a new model comes out, so that they have an excuse to upgrade. With statistics, an organization can determine whether this is actually occurring and implement policy changes as necessary.

the form of replacing the device). The direct costs can be quantified given that historical information exists on how many devices of each type were lost over the past year (see figure 61).

Probability	Device	Average Cost
0.34	Cellphone	$400
0.32	Token	$100
0.20	Laptop	$1000
0.07	Other	$300
0.05	Desktop	$1000
0.02	Tablet	$700

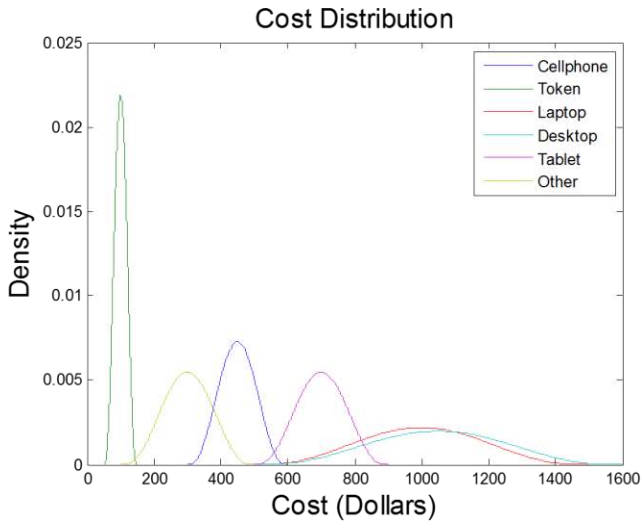


Figure 61: Costs due to lost devices.

Other costs are modeled using historical data (e.g., investigation costs) and the standard cost models used in other attack scenarios (i.e., PII, IP, and reputation damage). Business interruption is modeled under the assumption that employees always lose productivity while the lost device is replaced, which typically takes between three and eight hours (uniformly) with a time cost of $100 per hour.

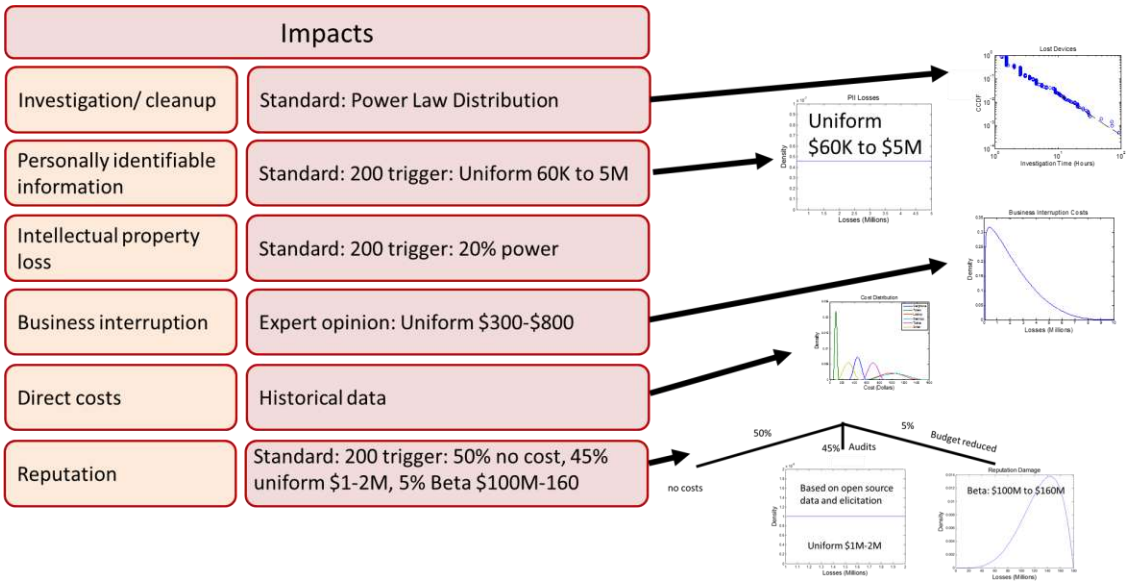


Figure 62: Impact model for lost devices.

A Monte Carlo simulation is used to generate the risk curve shown in figure 63. Overall, lost devices are found to have a cost of roughly $1 million per year, given the current environment.

Most of this cost is due to employee downtime (roughly 50%). About $360,000 comes from the direct cost of replacing hardware. The rest comes from incident investigation.

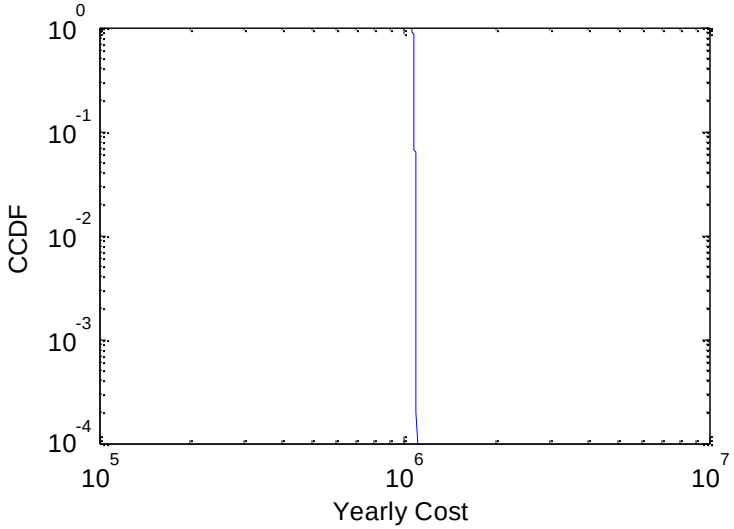


Figure 63: Risk curve for lost devices. Note that very little variation exists in the yearly costs.

The yearly cost has very little variation. This is a direct result of the fact that full disk encryption (FDE) is already implemented on laptops across Space Corp. FDE effectively prevents any large-impact incidents from occurring, meaning that reputation damage, IP loss, and PII loss rarely occur. Note that these results are also based on the fact that Space Corp has a comprehensive laptop backup program enabled. Therefore, lost devices rarely have information on them that is not stored somewhere else, meaning lost data can easily be recovered. For organizations where this may not be true, a different cost model would be used.

5.7.4 Lost Device Safeguards

Space Corp is considering three safeguards for reducing the risk associated with lost devices, namely asset recovery software, theft awareness training, and a laptop loaner program. Further, Space Corp has already implemented FDE and wants to determine whether it was a cost-effective investment.

Full Disk Encryption

Full disk encryption (FDE) protects the data on a computer by requiring a password before a computer boots up. File encryption secures files individually, while FDE encrypts the entire hard drive, including the operating system files. Without FDE, the operating system password can easily be bypassed. FDE makes it considerably harder for an adversary to obtain information off a lost

laptop.[75] In fact, most data breach laws do not require notification in cases where the stored information is encrypted.

FDE is an obvious benefit to many organizations, but the cost-effectiveness of the technology implementation still needs to be considered. A number of FDE solutions are available, including BitLocker (Windows native), FileVault 2 (Macintosh OSX native), and PGP. The costs of these technologies range from free (default on machines) to $200 per license per year. In this analysis, FDE externalities such as the added time cost of typing a password before booting, the costs associated with the software implementation, and computer crashes due to software errors are not modeled because Space Corp believes that these costs are negligible. Otherwise, these additional costs can easily be incorporated into the analysis.

FDE has two primary effects on the costs of lost laptops.[76] First, it eliminates reputation damage, PII losses, and IP losses since data cannot be easily extracted from a lost device. Second, investigation times go down considerably because laptops with FDE do not need to be searched for sensitive information. Space Corp implemented full disk encryption roughly halfway through the six years of data. The investigation time of lost devices can be compared before and after the implementation to calculate the effect of FDE. Figure 64 shows the change in investigation time before and after FDE was implemented. FDE corresponded with the elimination of large-impact incidents.

[75] While some attacks, such as cold-boots and off-line brute force attacks, are possible, they are sufficiently difficult that most criminals will not attempt them. Note that in the case where motivated nation states are a threat, FDE attacks should be modeled.

[76] A simplifying assumption is that FDE is implemented completely and perfectly across the organization. In practice, organizations have missed encrypting a handful of laptops, leading to data breaches after an FDE policy is enacted.

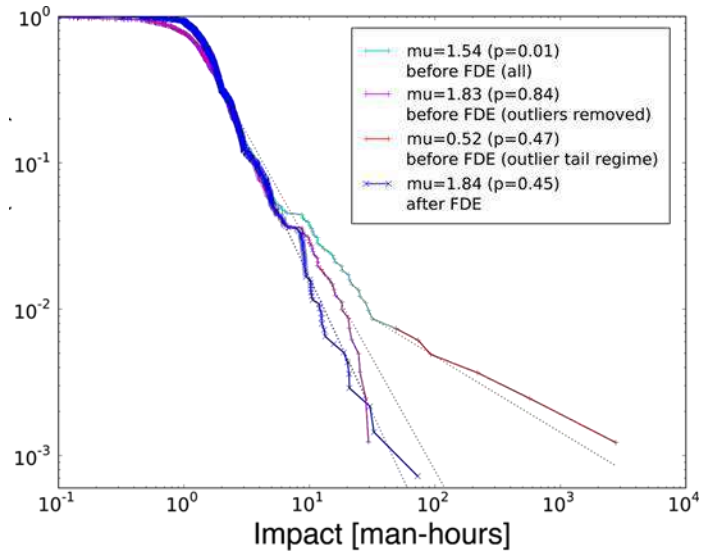


Figure 64: The CCDF of investigation time for lost devices. Before FDE was implemented, incidents followed a main tail distribution (power law) and an outlier tail that involved devices with PII data. After FDE was implemented, the outlier tail disappeared. Taken from (Kuypers, Maillart, & Paté-Cornell, 2016).

The data show that the large-impact incidents disappeared after FDE was implemented. Using this information, the risk curves before and after FDE can be compared (figure 65). FDE costs Space Corp $2 million per year to implement across the organization. Therefore, removing an FDE requirement would save $2 million per year. However, high-impact incidents would return along with reputation damage, credit monitoring costs from PII disclosures, and costs due to IP theft. Figure 65 shows that while "no FDE" is less expensive per year than FDE roughly 50% of the time, unencrypted laptops can come with very large costs.

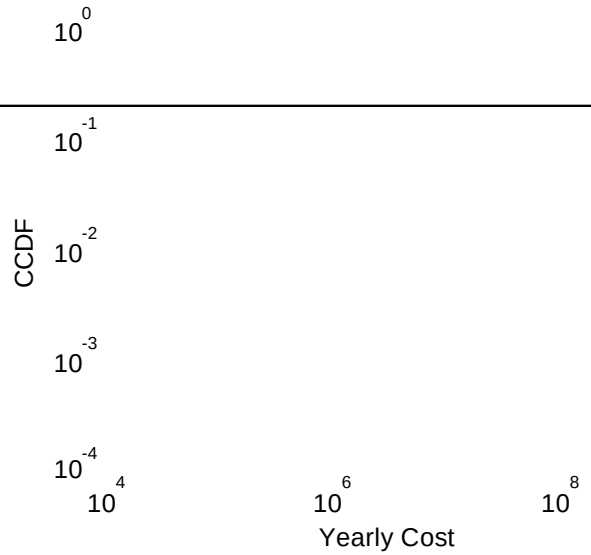


Figure 65: The effect of FDE on yearly costs.

Several other security safeguards can be considered as well. Theft awareness training attempts to reduce the rate of lost laptops by educating employees about simple strategies to reduce theft. For example, laptops should not be visible when left in a car, or be left outside the home overnight. Experts assess that a training program can be implemented for a fixed cost of $15,000 for training materials, plus 0.25 hours for each employee to take the training. Based on limited rollouts of training programs in the past, Space Corp believes that the training would reduce the number of lost devices by 20%.

Laptop loaner programs are also considered. Over 50% of the yearly costs due to lost devices are due to employee downtime. Space Corp could cut the downtime in half by creating an office that loans out laptops to employees until the laptop backups can be completed. The cost of stocking these additional laptops and providing personnel is estimated to be $100,000 per year.

Figure 66 compares the cost-effectiveness of these different laptop safeguards. Theft awareness training is not cost-effective. Although the time cost per employee is small, the large number of employees means that theft awareness is very expensive to implement, and the reduction in lost devices does not recover this cost. The loaner program is cost-effective for similar reasons. Reducing employee downtime makes up for the cost of the program because a small reduction in downtime compounds into large savings. Loaner programs may also have extra benefits in the form of aiding employees who simply forget their laptop at home.

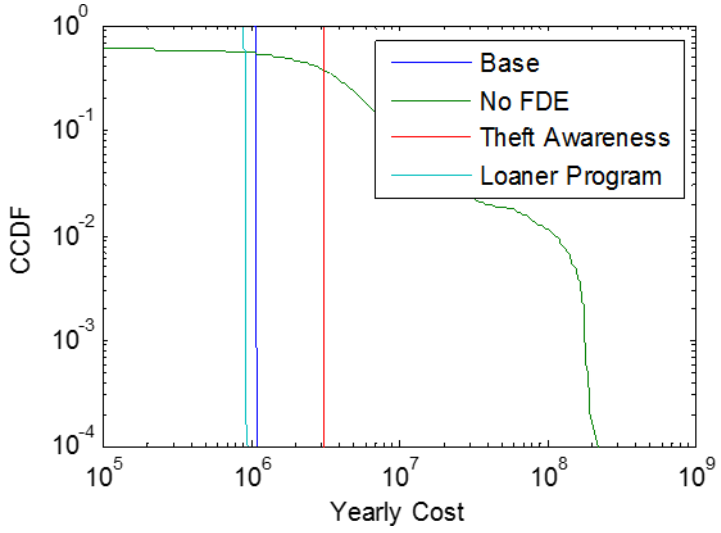


Figure 66: Risk curve for several lost device safeguards.

A final safeguard that is considered is asset recovery software, a technology that aids in the recovery of a lost laptop. Typically, software will automatically connect to the Internet and send information about its physical location to a user or police agency. Several asset recovery solutions exist, including LoJack and exo5, and cost around $100 for a three-year subscription.

The cost-effectiveness of asset recovery technology essentially represents a tradeoff between the value of the hardware and the cost of the software. Space Corp assumes that 60% of lost laptops with asset recovery software are eventually recovered; the rest are likely destroyed, harvested for parts, or wiped before connecting to an Internet source. The investigation time and the other cost vectors remain the same.

Quick calculations show that asset recovery software is a poor investment. The cost of the software is 20,000 employees * $33 per year = $660,000. The expected value of the lost equipment is only $362,000 (of which only 60% is recovered). It is immediately clear that the software does not cover the cost of the lost devices. Asset recovery could, however, be cost-effective for certain high-value laptops. For example, certain custom machines with computational power have a value of about $10,000. But asset recovery software is a bad deal for Space Corp in general.

5.8 Other Incident Types

Some cyber incidents have not occurred in the database with enough frequency to merit a specific attack model. For example, supply chain attacks, malicious insiders, and denial of service attacks all pose risks to an organization. However, the past data contain very few (if any) examples of these attack types, requiring a scenario-based modeling approach.

When conducting a risk analysis, it is important that perception matches reality. Many organizations are concerned with insider attacks, given the leaks that have impacted the Department of Defense. Supply chain attacks have also been a significant concern in organizations. In 2010, senators raised concerns about Huawei, a Chinese telecommunications company. The CEO of Huawei is a former member of the People's Liberation Army and there were worries that using Huawei equipment in the telecommunications infrastructure of the US could introduce secret backdoors. These concerns have prevented Huawei from successfully entering the US market.

While insiders and supply chain attacks are frequently in the news, these attack vectors are not always relevant to organizations. Organizations could easily bankrupt themselves trying to prevent rare incidents that would cost the organization about as much as a severe website incident. At Space Corp, no work on national security- or military-related technologies is conducted, meaning that it is easier to quantify the value of security with an upper bound. While other attacks are possible, the vectors modeled here are the most relevant. Focusing on the supply chain and malicious insider risk is analogous to reinforcing the walls of a bank vault while leaving the front door unlocked.

Other attack types (e.g., DDoS) could be large risks to certain organizations, in which case they should be modeled. For example, the financial sector has historically seen many DDoS attacks that are very costly because they prevent customers from using online services. Governments and defense contractors might need to worry about supply chain vulnerabilities, and certain organizations may need to be concerned about malicious insiders. Each organization should perform a careful analysis to determine the scope of attacks that will be considered.

5.9 Model Results

Combining the results of the previous case studies results in figure 67. The yearly losses due to each attack vector are shown, along with the total cost due to all vectors.

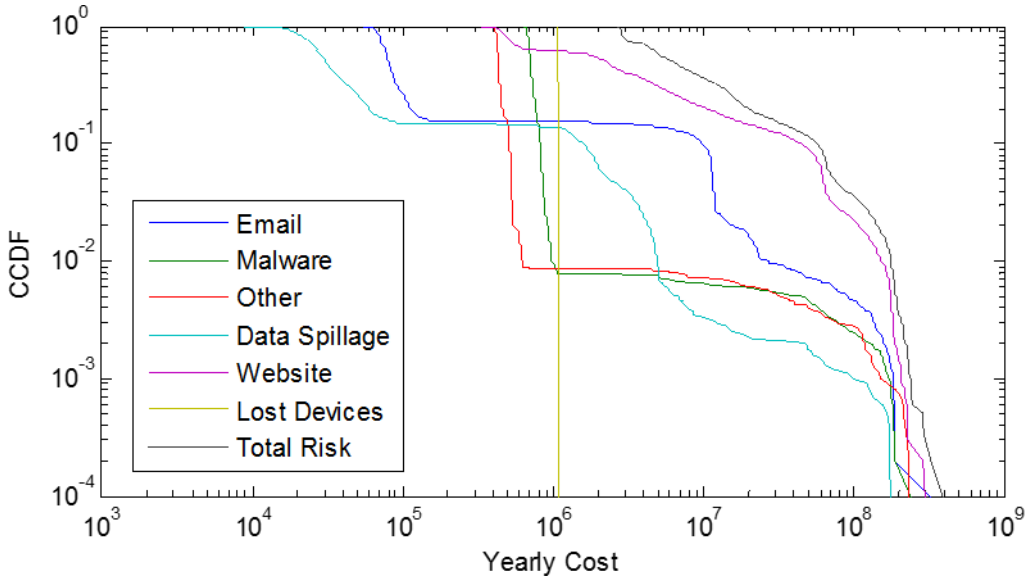


Figure 67: Cyber risk at Space Corp.

The risk curves shown in figure 67 enable a much higher quality discussion about cyber risk at an organization. Data spillage is found to be low risk, costing Space Corp less than $10,000 most years. Malicious email attacks require more investigation time, and carry much more risk in the tail due to consequences associated with the impact an adversary causes after compromising a system. Lost devices have an extremely low variance in cost due to full disk encryption, which was a good investment for Space Corp. Website attacks represent the largest cost to Space Corp due to investigation time and tail risk. Website incidents occur frequently and at all scales, resulting in very heavy-tailed losses.

The risk curves can also be used to rank cyber security investments. Website safeguards should be a first priority. The quantification of losses due to the different attack vectors provides a rigorous justification for changing the security budget (in many cases, increasing it). In many boardrooms, CISOs ask for more money to improve security, but cannot answer questions about how much more secure the organization will be given the additional expenses. Qualitative assessments are unconvincing, since evidence for why a risk goes from red to yellow is often lacking. Quantitative risk analysis (and empirical analyses of cyber incidents) demonstrate where

investments should be made and why. Decision makers are forced to acknowledge cyber risk, but are enabled to make intelligent tradeoffs with other forms of classic business risk.

Risk quantification also is useful for evaluating the need for cyber insurance. Interestingly, most organizations see cyber insurance as overpriced, while underwriters view it as underpriced. The stochastic nature of cyber impacts (i.e., the existence of heavy tails) is a strong indication that cyber insurance could be a sustainable market. In many cases, it would be too expensive for organizations to maintain the manpower and resources required to deal with a heavy-tailed event. Since they are rare, insurers can spread risk and provide an influx of capital to respond to and recover from large cyber incidents, in the form of third-party investigations, breach notification expenses, or legal litigation.

The results in figure 67 appear very simple, and a ranking of different attack vectors is clear. It is easy to forget that the seemingly obvious conclusions (i.e., data spillage is low risk, website risk is highest) are in fact unintuitive before the analysis has been run. The final results are also fundamentally different than the results obtained by taking the average severity and cost of different incident vectors. The variance in the level of costs is a key component of the overall risk. Costs from website attacks may be less than costs from lost devices in some years, but can exceed lost device costs by two orders of magnitude in other years. Analysts could easily make incorrect conclusions about the rank ordering of risks if a small sample of yearly costs is used, instead of modeling the entire system.

6 Conclusion

Quantitative cyber risk is in its infancy. Data are difficult for outside researchers to obtain, and organizations often lack the most basic situational awareness of their systems. The first step for many organizations in quantifying risk frequently becomes to simply determine what they have that attackers value, be it computing resources, designs, credit cards, or health records. The second step is to gather statistical data on cyber incidents that can be used to determine ground truth. The third step involves conducting a complete probabilistic risk analysis that includes both data analysis and scenario analysis for large incidents and emerging trends. Organizations that do not rely on data-driven methods are likely to struggle when making investment decisions.

Data provides an anchor that can improve decision making. For example, one organization had been dealing with lost laptops. After a laptop went missing, an investigator would interview the employee about information that was on the device. The employee would often state that no sensitive information was on the laptop. However, after the data were examined, user reporting was found to be extremely unreliable. In about 50% of the cases, sensitive information was found on the machine (using backups or the recovered device) despite the user's assessment. This is just one example of a case where accurate information can improve an organization's decision making, in this case to phase out reliance on user reporting. Here, recording both the reported type of information and the confirmed type of information on a device led to more useful data.

Some readers may be skeptical that sufficient data exist at their organization to conduct a similarly detailed risk analysis as outlined in this dissertation. Yet again, many organizations have more information than they think they do, and starting to quantify risk, no matter how small-scale, will begin to drive organizational change. Other readers may encounter difficulty in obtaining cost estimates through expert elicitation and willingness-to-pay methods. Expert elicitation can in fact be very challenging, requiring the right method and analyst, and, most importantly, an expert who is willing to quantify their own uncertainty. Giving probabilistic assessments may be a foreign exercise that requires practice.

6.1 Limitations

Risk models have inherent limitations stemming from the quality of the input data, the structure of the model, and the fundamental uncertainty of the system. While the monetary impact of many cyber incidents is known, surprises still occur. For example, one cyber insurer had an expensive claim filed when an airport's baggage sorting system was disabled by malware. The airport hired temporary workers to hand sort luggage until the system was repaired, leading to large costs that were not anticipated beforehand. Cyber will continue to impact systems in unique and surprising

ways for the foreseeable future, leaving the possibility that the risk models leave out important aspects of risks.

There are also limitations in using statistics based on past historical incidents to inform the risk model, given that some hacks may not have been detected, or may simply not yet have occurred. Certain discrete events can suddenly change the risk landscape. For example, Stuxnet introduced a plausible scenario where physical destruction could occur via a cyber attack. In these cases, a scenario-based model, expert opinion, and careful analysis (for example, of near-misses) are needed to fill the gap.

Questions remain about the applicability of the trends identified in the organization studied here to other organizations. Circumstantial evidence suggests that some observations will be consistent across many organizations. Laptop theft is likely to take place at a rate proportional to the number of devices in an organization, meaning that the lost device incidents at one organization can be applied to another. Other cyber incidents may exhibit differences across organizations of different sizes, industries, and maturity levels. For example, large financial institutions will likely face a different number of website defacement attempts compared to educational institutions. As more cyber incident data becomes available, researchers will be able to test these hypotheses.

Longitudinal studies are also needed to examine the effect that other variables, like industry or organizational culture, have on cyber incidents. In the data analyzed in this dissertation, it is remarkable that the rate and severity of cyber incidents is generally stable over time, after removing incident recording changes. Over the six years, the organization with 60,000 incidents experienced changes in its workforce and management. Significant changes also occurred in the network structure, and a large number of cyber security safeguard technologies were implemented. Despite all of this, the only action that was reliably observed was the impact of full disk encryption on reducing the severity of lost device incidents. More research is needed to determine whether the lack of change in other attack severities is due to the evolution of adversaries, organizational culture preventing effective security safeguard implementation, or some other cause.

There is also a large question about the role of observational bias in recording cyber security incidents. A common criticism is that larger cyber incidents may be occurring but not identified, recorded, or reported, leading to an inaccurate assessment of cyber risk. While this is indeed possible, the repeated identification of successful attacks from advanced and persistent adversaries indicates that the organization studied here has good detection capabilities. Still, it is possible that some severe incidents are not identified. Other sources show that a large proportion of cyber incidents are discovered by third parties (Verizon, 2014). An important point is that incidents that are discovered by third parties are still recorded in incident management systems, but

there may be a delay of months or years from when an attack occurs to its identification. The danger of this delay is lessened given the length of the data sample used here (six years), since incidents that are identified years later are still in the sample (although the last year might be missing some severe incidents).

As with any real-world dataset, the quality of the data could be improved. Typos are identified in the hours of investigation, duplicate incidents exist, and some incidents are only partially completed (missing an incident description or investigation time). Overall, however, these incomplete incidents make up a small proportion of the total number of incidents. In the case of the Space Corp database, data cleaning was used to identify errors and the security operations center was closely consulted to ensure that the data were accurate. While concern existed that differences in security analysts' investigative methods would increase the variability of the data, clear trends emerged as shown in this study.

6.2 Future work

Cyber security is one of the richest areas for additional research. Getting the right order of magnitude in cyber risk is often groundbreaking because it represents a first step in a more sophisticated assessment. The data used in this dissertation constitute one of the examples where detailed cyber security incidents have informed and validated a cyber security risk model.

This dissertation presents a limited number of attack types at a specific organization designed to evaluate the effectiveness of several cyber security safeguards. Cyber security products are abundant and many more sub-models need to be developed. For example, the risk reduction due to antivirus software, strong password requirements, or whitelisting is largely unknown.[77] Some technologies will require more advanced privilege escalation and lateral movement models. For example, file integrity monitoring (FIM) calculates a unique hash of critical files and sends them to a central repository that cannot be deleted. If a file is changed, a new hash is reported to the central logging system and an alert is generated, making it more difficult for an adversary to remain undetected. Assessing the value of FIM requires more information about an organization's internal structure, and the most probable sequence of actions that an adversary would take. Further, FIM can detect attacks that occur via websites, email, web browsing, and other vectors, requiring an organization to develop a comprehensive, multi-feature security model.

[77] Whitelisting involves specifically designating which programs can talk to which IP addresses, while blacklisting forbids named programs or IPs. For example, blacklisting address X.X.X.X means that devices can communicate with any IP address except X.X.X.X, while whitelisting means that devices can only communicate with Y.Y.Y.Y.

The model presented here is also targeted at an aerospace organization. While the general method is widely applicable, the risk model will vary depending on the organization type. Retail organizations may need to maintain a positive public image more than manufacturers do, for example. Different sectors will have unique tradeoffs that need to be modeled as well. Educational institutions, for instance, need to operate in a heterogeneous environment that is largely open: IP addresses from all areas of the globe routinely connect to university webpages, given the diversity of the student population. Small consulting firms, on the other hand, may have the ability to monitor web traffic much more closely. If the business does not operate in Asia, then any VPN connection attempts from Asia could be automatically blocked or scrutinized.

The model could also be extended to project-level risk management solutions. CISOs are often faced with the decision to approve waivers for special project requirements. For example, a high-performance computing group may require a waiver that disables timed screen lockouts, since diagnostics need to be run overnight and a lockout would disrupt the test. In this case, the decision maker needs a way to evaluate the additional risk versus the project's requirement.

Another area that is ripe for improvement is the monetary estimation of cyber impacts. Currently, most organizations record cyber impacts using a qualitative scale (e.g., low, medium, high). An improvement (present in the dataset analyzed here) records a quantitative proxy for monetary cost (i.e., hours of investigation). Ideally, organizations should record the estimated monetary cost for each incident, incorporating investigation time, reputation damage, business interruption costs, and others. Separately recording each of these cost vectors results in the highest quality data for assessing cyber risk.

More work is needed for risk quantification, not just at the entity level but at the industry and global levels as well. Cyber insurance is a well-established market, but the techniques for assessing cyber risk and pricing policies are still insufficient. Underwriters are likely to develop a significant portion of risk quantification, given their need to accurately price polices. Cyber insurers also need to assess aggregate risk for the industry. Hurricanes and earthquakes have the potential to impact large numbers of insureds, leading to very large losses. Reinsurance (insurance for insurers) exists to limit aggregate risk, but accurate assessments of heavy-tailed events are needed. In cyber, a regional Internet outage, data center failure, or wide-spread vulnerabilities have the potential to result in a new form of aggregate risk, requiring the development of new tools and techniques.

6.3 Summary

Many organizations struggle with the quantification of cyber security risk. Qualitative frameworks are common, but provide limited decision support and do not permit the ability to quantitatively assess the effectiveness of cyber security safeguards. In this dissertation, a method for quantifying cyber security risk is shown using 60,000 cyber security incidents recorded at a large US-based organization over six years. Extreme events (i.e. the theft of intellectual property) that have not occurred yet are included in the analysis to generate complete risk curves.

The risk model presented in this dissertation builds on previous quantitative risk models by adding high-quality data from a large organization and incorporating uncertainty about the rate, severity, and monetary impacts of cyber incidents directly into the analysis. Cyber security incidents have been shown to have heavy tails, meaning that using the mean can be highly misleading. Further, while many sources of data now exist to estimate difficult-to-quantify impacts (e.g., reputation damage), the monetary costs can still have large ranges. Encoding this uncertainty into the model results in the ability to compare frequent low-cost incidents with rare high-cost incidents.

The quantitative treatment of risk allows decision makers to prioritize cyber investments, justify cyber security budgets, and improve situational awareness by identifying the highest risk areas of a network. Powerful quantitative tools like probabilistic analysis, sensitivity of the results to the chosen parameters, and the value of information tests can indicate where additional money should be spent.

Risk quantification can improve risk communication. Many board members will be surprised to learn how much the steady stream of nuisance cyber incidents (e.g., spammers, website defacements, lost devices, data spillage) costs the organization. While other types of business risk have a long history of quantification, cyber risk has mostly been assessed on an unrelated scale that prevents its comparison to other areas. Putting all outcomes in monetary terms allows a better comparison of risk across an organization.

Cyber security is often a reactive process. Organizations routinely fail to prioritize countermeasures to cyber attacks, resulting in exposures that can eventually lead to a massive cyber incident. Bad outcomes elevate cyber security concerns to the board level, where members perhaps make excessive investments into security. In time, after memory of the cyber incident fades, interest stagnates and cyber is ignored until the next incident, leading to a repetitive cycle. Quantitative cyber risk addresses this sub-optimal process. Because the threats may constantly vary, the risk has to be regularly updated and safeguard investments re-evaluated.

Overall, considerable potential exists for cyber risk management. The data are getting better, an increasing number of experts are studying the problem, and interest in the field is growing. Risk quantification and risk management are a critical piece of the enterprise management puzzle. In the future, cyber risk may be as well understood as car insurance or homeowner coverage, but may still vary as new attackers and types of attacks emerge.

References

Abelson, H., Anderson, R., Bellovin, S. M., Benaloh, J., Blaze, M., Diffie, W., ... & Rivest, R. L. (2015). Keys under doormats: mandating insecurity by requiring government access to all data and communications. Journal of Cybersecurity, tyv009.

Alberts, C. J., & Dorofee, A. (2002). Managing information security risks: the OCTAVE approach. Addison-Wesley Longman Publishing Co., Inc.

Anderson, R., Barton, C., Böhme, R., Clayton, R., Van Eeten, M. J., Levi, M., ... & Savage, S. (2013). Measuring the cost of cybercrime. In The economics of information security and privacy (pp. 265-300). Springer Berlin Heidelberg.

Andrijcic, E., & Horowitz, B. (2006). A Macro-Economic Framework for Evaluation of Cyber Security Risks Related to Protection of Intellectual Property. Risk Analysis, 26(4), 907–923.

Biener, C., Eling, M., & Wirfs, J. H. (2015). Insurability of Cyber Risk: An Empirical Analysis†. The Geneva Papers on Risk and Insurance–Issues and Practice, 40(1), 131–158.

Blum, D. (2012). Probabilistic Models for Warning of National Security Crises. Stanford dissertation. 2012.

Bojanc, R., & Jerman-Blažič, B. (2008). An economic modelling approach to information security risk management. International Journal of Information Management, 28(5), 413–422.

Bojanc, R., & Jerman-Blažič, B. (2013). A quantitative model for information-security risk management. Engineering Management Journal, 25(2), 25–37.

Buckshaw, D. L., Parnell, G. S., Unkenholz, W. L., Parks, D. L., Wallner, J. M., & Saydjari, O. S. (2005). Mission oriented risk and design analysis of critical information systems. Military Operations Research, 10(2), 19–38.

Campbell, K., Gordon, L. A., Loeb, M. P., & Zhou, L. (2003). The economic cost of publicly announced information security breaches: empirical evidence from the stock market. Journal of Computer Security, 11(3), 431–448.

Cardwell, D. (2014, September 1). Solar Company Seeks Stiff U.S. Tariffs to Deter Chinese Spying. Retrieved from http://www.nytimes.com/2014/09/02/business/trade-duties-urged-as-new-deterrent-against-cybertheft.html

Cavusoglu, H., Mishra, B., & Raghunathan, S. (2004). A model for evaluating IT security investments. Communications of the ACM, 47(7), 87–92.

Cavusoglu, H., Mishra, B., & Raghunathan, S. (2004). The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers. International Journal of Electronic Commerce, 9(1), 70–104.

Cavusoglu, H., Raghunathan, S., & Yue, W. T. (2008). Decision-theoretic and game-theoretic approaches to IT security investment. Journal of Management Information Systems, 25(2), 281–304.

Chapman, I. M., Leblanc, S. P., & Partington, A. (2011, April). Taxonomy of cyber attacks and simulation of their effects. In Proceedings of the 2011 Military Modeling & Simulation Symposium (pp. 73–80). Society for Computer Simulation International.

Condon, E., He, A., & Cukier, M. (2008, November). Analysis of computer security incident data using time series models. In Software Reliability Engineering, 2008. ISSRE 2008. 19th International Symposium on (pp. 77–86). IEEE.

Cox, T. (2008). What's wrong with risk matrices? Risk analysis, 28(2), 497–512.

Cyber Incident Handling Program. (2014, December 18). Chairman of the Joint Chiefs of Staff Manual. Retrieved from http://www.dtic.mil/cjcs_directives/cdata/unlimit/m651001.pdf

Daniels, M. (2014). Optimization of spacecraft architectures for earth-orbit satellite projects. Stanford dissertation.

De Cristofaro, E., Du, H., Freudiger, J., & Norcie, G. (2013). A comparative usability study of two-factor authentication. arXiv preprint arXiv:1309.5344.

DHS. (2015). Enhancing Resilience through cyber incident data sharing and analysis. DHS whitepaper.

Dimkov, T., Pieters, W., & Hartel, P. (2010, October). Laptop theft: a case study on the effectiveness of security mechanisms in open organizations. In Proceedings of the 17th ACM conference on Computer and communications security (pp. 666–668). ACM.

Dmitrienko, A., Liebchen, C., Rossow, C., & Sadeghi, A. R. (2014). Security analysis of mobile two-factor authentication schemes. Intel® Technology Journal, 18(4).

Edwards, B., Hofmeyr, S., & Forrest, S. (2015). Hype and heavy tails: A closer look at data breaches. WEIS.

Esterl, M. (2014, January 24). Coca-Cola: Stolen Laptops Had Personal Information of 74,000. Retrieved from http://www.wsj.com/articles/SB10001424052702304632204579341022959922200

Florêncio, D., & Herley, C. (2013). Sex, lies and cyber-crime surveys. In Economics of information security and privacy III (pp. 35–53). Springer New York.

Freund, J., & Jones, J. (2014). Measuring and managing information risk: a FAIR approach. Butterworth-Heinemann.

Garrick, B. J., Gekler, W. C., Goldfisher, L., Karcher, R. H., Shimizu, B., & Wilson, J. H. (1967). RELIABILITY ANALYSIS OF NUCLEAR POWER PLANT PROTECTIVE SYSTEMS (No. HN--190). Holmes and Narver, Inc., Los Angeles, Calif. Nuclear Div.

Glazer, E. (2015, August 3). J.P. Morgan to Accelerate Timeline for Cybersecurity Spending Boost. Retrieved from http://www.wsj.com/articles/j-p-morgan-to-accelerate-timeline-for-cybersecurity-spending-boost-1438641746

Goldstein, J., Chernobai, A., & Benaroch, M. (2011). An event study analysis of the economic impact of IT operational risk and its subcategories. Journal of the Association for Information Systems, 12(9), 606.

Gordon, L. A., & Loeb, M. P. (2002). The economics of information security investment. ACM Transactions on Information and System Security (TISSEC), 5(4), 438–457.

Gordon, L. A., Loeb, M. P., & Zhou, L. (2011). The impact of information security breaches: Has there been a downward shift in costs? Journal of Computer Security, 19(1), 33–56.

Greenberg, A. (2015, July 6). Hacking Team Breach Shows a Global Spying Firm Run Amok. Retrieved from http://www.wired.com/2015/07/hacking-team-breach-shows-global-spying-firm-run-amok/

Greisiger, M. (2013). Cyber liability & data breach insurance claims a study of actual claim payouts. Technical report, NetDiligence.

Guikema, S. (2002). Optimal resource allocation in an engineering design team with asymmetric information. Stanford dissertation.

Harrison, K., & White, G. (2011, January). A taxonomy of cyber events affecting communities. In System Sciences (HICSS), 2011 44th Hawaii International Conference on (pp. 1–9). IEEE.

Herrmann, A. (2013). The Quantitative Estimation of IT-Related Risk Probabilities. Risk Analysis, 33(8), 1510–1531.

Howard, R. A. (1968). The foundations of decision analysis. Systems Science and Cybernetics, IEEE Transactions on, 4(3), 211–219.

Howard, R. A., & Abbas, A. E. (2015). Foundations of decision analysis. Prentice Hall.

Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2011). Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Leading Issues in Information Warfare & Security Research, 1, 80.

Junio, T. J. (2013). How probable is cyber war? Bringing IR theory back in to the cyber conflict debate. Journal of Strategic Studies, 36(1), 125–133.

Kahneman, D. (2011). Thinking, fast and slow. Macmillan.

Kannan, K., Rees, J., & Sridhar, S. (2007). Market reactions to information security breach announcements: An empirical analysis. International Journal of Electronic Commerce, 12(1), 69–91.

Karam, E., & Planchet, F. (2015). Combining internal data with scenario analysis. Modern Economy, 6(05), 563.

Katsikas, S.K. (2009). Computer and information security handbook. Chapter 53. Morgan Kaufmann.

Killcrece, G., Kossakowski, K. P., Ruefle, R., & Zajicek, M. (2003). State of the practice of computer security incident response teams (CSIRTs) (No. CMU/SEI-2003-TR-001). CARNEGIE-MELLON UNIV PITTSBURGH PA SOFTWARE ENGINEERING INST.

Kitteringham, G. (2008). Lost laptops= lost data: Measuring costs, managing threats. In Crisp report, ASIS International Foundation.

Kjaerland, M. (2006). A taxonomy and comparison of computer security incidents from the commercial and government sectors. Computers & Security, 25(7), 522–538.

Korzak, E. (2014). Computer Network Attacks and International Law. Doctoral dissertation, University of London.

Krebs, B, (2010, May 13). Stolen Laptop Exposes Personal Data on 207,000 Army Reservists. Retrieved from http://krebsonsecurity.com/2010/05/stolen-laptop-exposes-personal-data-on-207000-army-reservists/

Krebs, B. (2014a, February 12). Email Attack on Vendor Set Up Breach at Target. Retrieved from http://krebsonsecurity.com/2014/02/email-attack-on-vendor-set-up-breach-at-target/

Krebs, B. (2015a, February 9). Anthem Breach May Have Started in April 2014. Retrieved from http://krebsonsecurity.com/2015/02/anthem-breach-may-have-started-in-april-2014/

Krebs, B. (2015b, August 14). Cyberheist Victim Trades Smokes for Cash. Retrieved from http://krebsonsecurity.com/2015/08/cyberheist-victim-trades-smokes-for-cash/

Krebs, B. (2015c, March 19). Are Credit Monitoring Services Worth It? Retrieved from http://krebsonsecurity.com/2014/03/are-credit-monitoring-services-worth-it/

Krebs, B. (2016a, January 30). Sources: Security Firm Norse Corp. Imploding. Retrieved from http://krebsonsecurity.com/2016/01/sources-security-firm-norse-corp-imploding/

Kucik, P. (2007). Probabilistic modeling of insurgency. Stanford dissertation.

Kuypers, M. A., & Paté-Cornell, M.E. (2016). Department of Energy Cyber Security Incidents. http://cisac.fsi.stanford.edu/publication/department-energy-cyber-security-incidents.

Kuypers, M. A., Maillart, T., and Paté-Cornell, M. E., 2016. An Empirical Analysis of Cyber Security Incidents at a Large Organization, in press.

Leyden, J. (2014, May 20). AVG on Heartbleed: It's dangerous to go alone. Take this (an AVG tool). Retrieved from http://www.theregister.co.uk/2014/05/20/heartbleed_still_prevalent/

Maass, P., & Rajagopalan, M. (2012). Does Cyber Crime Really Cost $1 Trillion? Pro Publica, 1st August,{Online Resource} Available at: http://www. propublica. org/article/does-cybercrime-really-cost-1-trillion [Accessed 26/11/12].

Maillart, T., & Sornette, D. (2010). Heavy-tailed distribution of cyber-risks. The European Physical Journal B, 75(3), 357–364.

Martin, P. (2012) NASA Cybersecurity: An Examination of the Agency's Information Security.

Martin, P. (2013) NASA's Information Technology Governance. Office of the Inspector General. Report NO. IG-13-015 (Assignment NO. A-12-018-00).

Mauw, S., & Oostdijk, M. (2005). Foundations of attack trees. In Information Security and Cryptology-ICISC 2005 (pp. 186–198). Springer Berlin Heidelberg.

McCann, E. (2013, September 6). Advocate Health slapped with lawsuit after massive data breach. Retrieved from http://www.healthcareitnews.com/news/AdvocateHealth-slapped-with-lawsuit-after-massive-data-breach

McLane, M., Gouveia, J., Citron, G. P., MacKay, J., & Rose, P. R. (2008). Responsible reporting of uncertain petroleum reserves. AAPG bulletin, 92(10), 1431–1452.

Meyers, C., Powers, S., & Faissol, D. (2009). Taxonomies of cyber adversaries and attacks: a survey of incidents and approaches. Lawrence Livermore National Laboratory (April 2009), 7.

Miura-Ko, R. A., & Bambos, N. (2007, June). SecureRank: A risk-based vulnerability management scheme for computing infrastructures. In Communications, 2007. ICC'07. IEEE International Conference on (pp. 1455–1460). IEEE.

Möller, B., Duong, T., & Kotowicz, K. (2014). This POODLE bites: exploiting the SSL 3.0 fallback. Google.

Moore, T., & Anderson, R. (2011). Economics and Internet Security: A Survey of Recent Analytical, Empirical, and Behavioral Research.

Mozur, P., & Wingfield, N. (2016, January 5). Microsoft Faces New Scrutiny in China. Retrieved from http://www.nytimes.com/2016/01/06/business/international/microsoft-china-antitrust-inquiry.html

Mukkamala, S., Janoski, G., & Sung, A. (2002). Intrusion detection using neural networks and support vector machines. In Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on (Vol. 2, pp. 1702–1707). IEEE.

NIST. (2012). Guide for Conducting Risk Assessments (NIST SP-800-30rev1). The National Institute of Standards and Technology (NIST), Gaithersburg.

Paté-Cornell, E., & Fischbeck, P. S. (1993). PRA as a management tool: organizational factors and risk-based priorities for the maintenance of the tiles of the space shuttle orbiter. Reliability Engineering & System Safety, 40(3), 239–257.

Paté-Cornell, E., & Guikema, S. (2002). Probabilistic modeling of terrorist threats: A systems analysis approach to setting priorities among countermeasures. Military Operations Research, 7(4), 5–23.

Paté-Cornell, E., & Guikema, S. (2002). Probabilistic modeling of terrorist threats: A systems analysis approach to setting priorities among countermeasures. Military Operations Research, 7(4), 5–23.

Paté-Cornell, M. E. (1993). Learning from the piper alpha accident: A postmortem analysis of technical and organizational factors. Risk Analysis, 13, 215–215.

Paté-Cornell, M. E. (1993). Risk analysis and risk management for offshore platforms: lessons from the Piper Alpha accident. Journal of Offshore Mechanics and Arctic Engineering, 115(3), 179–190.

Paté-Cornell, M. E., Lakats, L. M., Murphy, D. M., & Gaba, D. M. (1997). Anesthesia patient risk: a quantitative approach to organizational factors and risk management options. Risk Analysis, 17(4), 511–523.

Ponemon Institute. (2011). Cost of a Data Breach.

Ponemon Institute. (2014). 2014 Global Report on the Cost of Cyber Crime.

Ponemon, L. (2009). Business Risk of a Lost Laptop: A study of U.S. IT Practitioners.

Ponemon, L. (2009). The Cost of a Lost Laptop.

Reilly, S. (2015, September 11). Records: Energy Department struck by cyber attacks. Retrieved from http://www.usatoday.com/story/news/2015/09/09/cyber-attacks-doe-energy/71929786/

Reinhardt, J. C., Chen, X., Liu, W., Manchev, P., & Paté-Cornell, M. E. (2015). Asteroid Risk Assessment: A Probabilistic Approach. Risk Analysis.

Reuters. (2016, March 21). Fitch: Rapid Growth in Cyber Insurance Would Be Credit-Negative. Retrieved from http://www.reuters.com/article/idUSFit952109

Romanosky, S., Hoffman, D., & Acquisti, A. (2014). Empirical analysis of data breach litigation. Journal of Empirical Legal Studies, 11(1), 74–104.

Ruffle, S., Leverett, E., Coburn, A., Copic, J., et al. (2015). Business Blackout. Emerging Risk Report.

Ryan, J. J. C. H., & Jefferson, T. I. (2003, May). The use, misuse, and abuse of statistics in information security research. In Proceedings of the 2003 ASEM National Conference, St. Louis, MO.

Schneier, B. (1999). Attack trees. Dr. Dobb's journal, 24(12), 21–29.

Schwartz, M. J. (2011, May 30). Lockheed Martin Suffers Massive Cyberattack. Retrieved from http://www.darkreading.com/risk-management/lockheed-martin-suffers-massive-cyberattack/d/d-id/1098013?

Shostack, A. (2014). Threat modeling: Designing for security. John Wiley & Sons.

Simmons, C., Ellis, C., Shiva, S., Dasgupta, D., & Wu, Q. (2009). AVOIDIT: A cyber attack taxonomy.

Sinanaj, G., Muntermann, J., & Cziesla, T. (2015). How Data Breaches Ruin Firm Reputation on Social Media!–Insights from a Sentiment-based Event Study. In Wirtschaftsinformatik (pp. 902–916).

Sonnenreich, W., Albanese, J., & Stout, B. (2006). Return on security investment (ROSI)–a practical quantitative model. Journal of Research and practice in Information Technology, 38(1), 45–56.

Soo Hoo, K. J. (2000). How much is enough? A risk management approach to computer security. Stanford, Calif: Stanford University.

Stoneburner, G., Goguen, A. Y., & Feringa, A. (2002). Sp 800-30. risk management guide for information technology systems.

The White House. (2015, February 13). Fact sheet: White House Summit on Cybersecurity and Consumer Protection. Retrieved from https://www.whitehouse.gov/the-press-office/2015/02/13/fact-sheet-white-house-summit-cybersecurity-and-consumer-protection

Thomas, R. C., Antkiewicz, M., Florer, P., Widup, S., & Woodyard, M. (2013). How bad is it?–a branching activity model to estimate the impact of information security breaches. A Branching Activity Model to Estimate the Impact of Information Security Breaches (March 11, 2013).

ThreatSim. (2013). State of the Phish. Retrieved from https://info.wombatsecurity.com/state-of-the-phish

Tsai, C. F., Hsu, Y. F., Lin, C. Y., & Lin, W. Y. (2009). Intrusion detection by machine learning: A review. Expert Systems with Applications, 36(10), 11994–12000.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. science, 185(4157), 1124–1131.

US-CERT Federal Incident Notification Guidelines. Retrieved from https://www.us-cert.gov/incident-notification-guidelines

US–China Economic and Security Review Commission. (2011). Report to Congress. US Government Printing Office, Washington, November.

Verizon. (2014). Data Breach Investigations Report.

Weise, E. (2015, June 8). U.S. Army website hacked, Syrian group claims credit. Retrieved from http://www.usatoday.com/story/tech/2015/06/08/us-army-website-wwwarmymil-syrian-electronic-army-hack/28703173/