GUIDEWIRE

WHITE PAPER

Taming the Uncertainty of Ransomware Risk

Introduction

It's easy to find ourselves doing much but achieving little when it comes to cyber perils. The particular cyber risk that is top of mind—ransomware—is no exception. In this white paper, Guidewire Cyence advances both a new mindset and a practical means to tame ransomware risk uncertainty with an innovative approach. We can be more empowered to understand and anticipate this peril by harnessing the signals of distress that correlate with individual ransomware incidents.

It's easy to find ourselves doing much but achieving little when it comes to cyber perils. The particular cyber risk that is top of mind—ransomware—is no exception. The field of play is familiar:

- **Dynamic and adaptive threats**
- **Interconnected risks and accumulated losses that defy temporal, organizational, and geographic boundaries**
- **Skewed and disconnected data**
- **Deficiencies in cross-enterprise metrics**
- **Skepticism about model methodology**
- **Information asymmetries between risk transfer supply and demand (poor data sharing)**
- **Complexity of cyber policies**

For data breach, at least, we've built up enough data to enable a degree of "uncertainty equilibrium" to price for incident severity and frequency. We've been able to do this primarily due to public incident reports that illuminated corporate victim information (including industry sector and revenue) and asset losses (sensitive record count), augmented by insights into threat and vulnerability data.

But ransomware incidents don't have similarly widespread legal requirements for public disclosure[1], leaving companies and insurers relatively unmoored from a shared understanding of the rates and costs of ransomware attacks and the volatility of trends that derive from collective measurements.

In this white paper, we advance both a new mindset and a practical means to tame ransomware risk uncertainty with an innovative approach. We can be more empowered to understand and anticipate this peril by:

- **Harnessing the signals of distress that correlate with individual ransomware incidents**
- **Leveraging the playbooks that steer cyber risk stakeholders[2]**

Reducing Conceptual Uncertainty: A Playbook Approach

In the evolving battle between offense and defense, ransomware is the latest chapter in the red team (offense) and blue team (defense) playbooks that are used in the tradecraft of cybersecurity and risk management. Lacking collective truth about ransomware's victimology, prevalence, payout rates, demand amounts, and other costs,[3] risk managers and underwriters face challenges in coverage, risk

[1] Entities covered by HIPAA that are infected with ransomware are presumed to have a reportable data breach unless it can be shown that there was a low probability that the protected health information (PHI) has been compromised.

[2] Stakeholders are defined here as threat actors, target organizations, security service and control vendors, and cyber insurance entities.

[3] Costs include lost business income, restoration and recovery of data and systems, forensics, and litigation.

selection, premium, and capital allocation. To be sure, there's a growing body of descriptive and predictive statistics from cyber risk and security products as well as services vendors—each with varying and limited scope—resulting in unstable trend analyses and risk-transfer paralysis.

In 2017, Petya/NotPetya and WannaCry ransomware made everyone sit up and notice, causing cross-industry, accumulative, and global losses that climbed into the hundreds of billions of dollars. The following year brought another round of notorious incidents, this time targeting public institutions. The cities of Baltimore and Atlanta rejected bounty payouts and incurred losses just under $20 million from recovery and mitigation alone. Meanwhile, the ransomware variants read like an assemblage of gamer tags or sci-fi characters: Ryuk, REvil, Matrix, BitPaymer, Cerber, Hermes, CryptoLocker, Robinhood, CryptoWall, Maze, Dharma, GandCrab, Emotet, iEncrypt, LeChiffre, LockCrypt, Megacode, LockerGoga, Nymain, PewCrypt, and SamSam.

A more in-depth look at recent history provides no shortage of descriptive analyses and predictive assessments of ransomware frequency and severity. For example, in 2019:

- **The average ransomware payment increased 1,150% [Coveware]**
- **The average down time caused by ransomware attacks increased by a factor of 2.6 [Coveware]**
- **The average cost of ransomware-caused downtime increased by more than 200% [Datto]**
- **The number of ransomware incidents increased 37% between Q1 and Q2 [Beazley Breach Insights]**
- **The number of ransomware claims increased by more than 2,000% from 2014 to 2019 and is anticipated to continue [Net Diligence 2020 Spotlight on Ransomware]**

Nevertheless, we find little convergence around the following questions:

- **What is the year-over-year increase in the number of ransomware attacks?**
- **Are attacks targeted or opportunistic?**
- **Is there a typical ransomware victimology? Are the attacks distinguishable by industry sector, company size, revenue, geographic or virtual footprint, or some other measurable feature? What proportion of the number of total victims does a particular feature account for?**
- **What is the average cost of a ransomware attack? How does the cost break down proportionally between business interruption (downtime/lost income), recovery, and restoration? What is the average downtime sustained by victims?**
- **What is the average bounty payout? How often is a bounty paid? Are demands increasing or decreasing? By what percentage and over what time period are they increasing or decreasing?**
- **Does paying the ransom result in an increased number of attacks for organizations in a given industry sector?**

- **What is the default rate on paid bounties (when payment is made but files and systems are not decrypted)?**
- **What percentage of ransomed companies pay the bounty to avoid having to disclose the specific incident, the harm to their risk reputation, and other costs?**
- **How much data and how many systems are recovered after a decryptor tool is provided?**
- **Are there predominant and consistent attack vectors? Do they vary according to victimology features or other observable characteristics?**
- **What are the most common families of ransomware? What attributes account for their popularity?**
- **What are the common controls that can be implemented to specifically prevent and mitigate ransomware attacks?**

Despite this volatility in statistical trends, companies and insurers can reduce uncertainty by invoking the playbooks that underpin attacker and defender interactions. These playbooks consider the strategic, tactical, and operational levels of the cat-and-mouse game that is cyber crime and cybersecurity. Although real-time specifics at the operational level continue to defy certainty, recognizing the tactical techniques and high-level strategies can go far in taming uncertainty.

Cyber crime strategy and tactics are foreseeable. Whether it is ransomware, data breach, or DDoS, the motive, means, and opportunity (MMO) script for cyber criminal developers and distributors has remained the same:

Motive (why): The attackers' objectives are to disrupt an organization and/or extract value for their own gain.

Means (how): The core blueprint for how attackers accomplish their objectives often comprises reconnaissance, target selection, evasion, and system/data incapacitation or theft.[4]

Opportunity (when, where, what): The resources, timing, and placement of attacks are a function of technology, process, and human vulnerabilities.

Although these fundamentals are unchanged, technology has enabled their evolution and points to where trendlines are heading. Because attackers are rational economic, ideological, or geopolitical actors, they embrace technology to optimize the execution of their mission—just like their legitimate business counterparts and targets. Hallmarks of this evolution are automation, cryptocurrency, and a service-oriented business model.

Ransomware optimizes **motive** by focusing on higher, more likely returns at lower risk: Why would attackers resort to extracting value from selling access to resources, credit cards, or

[4] In security industry terminology, these are referred to as the TTPs (tactics, techniques, and practices).

personal data in a volatile and saturated underground market? Or why would attackers settle for the exposure and limitations of committing identity fraud and money laundering when they can lock down a system or merely threaten to expose data, and then collect a quick and certain payout in bitcoins.

Ransomware optimizes the **means** by automating the steps involved in reconnaissance and attack, thus enabling more-efficient ratios of cost-of-effort to reward.

Ransomware optimizes **opportunity** by adopting a specialization business process that turns ransomware attacks into a modular service-oriented ecosystem. Similar to the familiar XaaS (where X is platform, infrastructure, software, or data, among others), ransomware as a service (RaaS) involves a supply chain of developers, aggregators, operators, and affiliates. They perform different roles with associated rewards in executing a ransomware attack: from setting up malware portal storefronts[5] to selling plug-and-play malware kits, finding and deploying them on victim systems, and finally liquidating the bounty demand in cryptocurrency.

PowerShell-based malware (a variant of which was described by some as "novel" at the time of this white paper's publication) illustrates this automation-driven evolution, while at the same time affirming the red team playbook for ransomware. First, the attack is triggered when either a duped user clicks infected email or a poorly authenticated (no MFA) server allows a password to be guessed with brute force. Reconnaissance is automated by leveraging SMB (Server Message Block), a network protocol used by Windows-based computers that enables systems in the same network to share files. Distribution is automated using a PowerShell command that instructs all other hosts on the network to fetch a malware-infected file from a remote server and execute the instructions. Evasion is automated by dynamically loading the payload modules to evade static detection tools. Detection is further reduced by obfuscating the recon and distribution in commonly used systems and administrator processes, as well as by killing other processes that leave attack artifacts. The play finishes with a script that generates an encryption key that locks down file types of the attacker's choosing and then offloads the key to the attacker's server and an HTML ransom note to the victim.

Why does shifting to a playbook mindset matter? The right frame of reference helps us manage the unknowns. It defines the vantage point through which we interpret risks. By framing ransomware risk according to the strategy and tactics of the red team, defenders and risk professionals can avoid blindspots that constrain the solutions for managing this type of risk. By putting a box around ransomware—reducing uncertainty at the strategic and tactical levels—at-risk companies and risk professionals can improve situational awareness, risk identification, and risk management.

[5] Storefronts include bundle discounts, support service, and customer reviews.

Methodology

- Using publicly reported ransomware incidents from 2010 through March 2020, Cyence analyzed the relationship between these events and approximately 40 cyber risk factors, as well as a company's risk rating, expressed on a scale from 100 (lowest risk) to 400 (highest risk).
- Cyence conducted the analyses both for "all companies" with incidents (730) and "companies with greater than $20 million in revenue" (531). While this count is certainly not the totality of ransomware incidents over the approximately 10-year time frame, the subset analyzed were those with reported incidents for which there was risk factor data. Cyence continues to collect and curate valuable data to best reflect the actual number of ransomware incidents.
- Cyence minimized a potential collection bias effect by partitioning the correlation analysis into two revenue bins. This is because it can be challenging to pick up signal for some risk factors for very small companies. Also, there are millions of small- and medium-sized businesses but very few reported incidents, so reported incidents skew toward larger companies, which usually have higher risk scores.
- Cyence backtested the ransomware risk factor correlation by applying the association analysis to historical incident data from 2017 to see how accurately the method would have predicted actual results. We chose 2017 (the year with most reported ransomware incidents collected) as the incident observation year, thereby assuming the perspective of someone in December 2016 seeking to leverage risk factors to know the likelihood of a firm experiencing a ransomware attack in the subsequent 12 months.

Reducing Empirical Uncertainty: Risk Factors as Risk Differentiators

In addition to playbook framing, in practice, ransomware uncertainty can be tamed by leveraging signals of distress that actually correlate to this peril. Cyence has developed meaningful risk insights to help identify which companies may be at higher risk of a successful ransomware incident. Rather than relying on instinct or educated guesses, insurers and organizations can now turn to threat and exposure signals to select and quantify ransomware risk at both firm and portfolio levels, as well as deploy controls that prevent and mitigate ransomware incidents. While correlation of these risk factors and risk ratings with ransomware more directly assists with frequency estimates, risk selection, and risk prevention, it can have comparative ranking and trend implications for risk pricing and capital allocation.

Using threat and exposure signals along with real incident data, Cyence conducted correlation analyses to discover the discrimination power of risk factors and risk ratings with ransomware incidents.

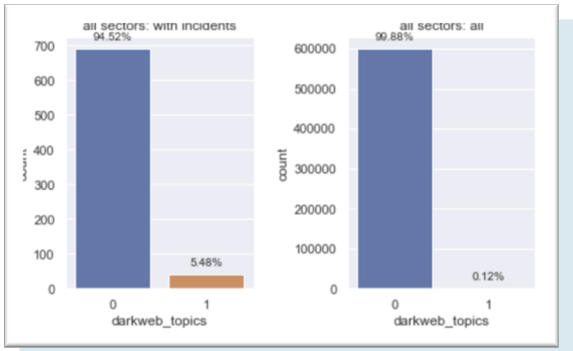
The results show the discrimination power of risk factors for ransomware. Specifically, the likelihood of a company having a ransomware incident increases by the following multiples if an organization exhibits the following risk signals:

45x	22x	22x	20x	7x	4x	3–4x
Targeted Darkweb Chatter	Leaked User Accounts	High Risk Rating	Compromised User Passwords	DNS Leakage	Spam Activity	Email Misconfig

The results below indicate the predictive value of each specific risk factor for all industry sectors. The top four sectors with historical ransomware incidents are education and research, healthcare, public administration, and financial services.

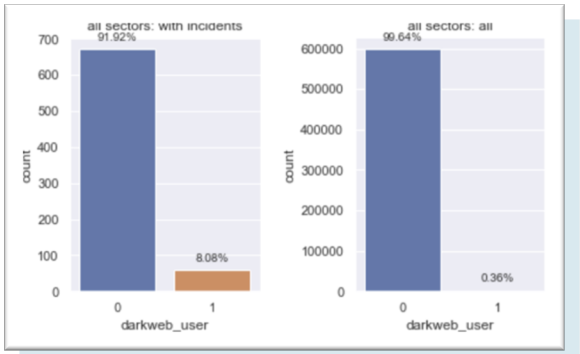
Targeted Darkweb Chatter (45x Discrimination Power)

- **This risk factor comprises specific and focused discussions about the company within the hacker community. Hackers use underground forums for a variety of activities: sharing ideas, tips, and tricks; planning and executing attacks; and even boasting about the latest successful hack. Increased hacker forum activity generally signals that a company has recently been breached or is being targeted for an attack.**



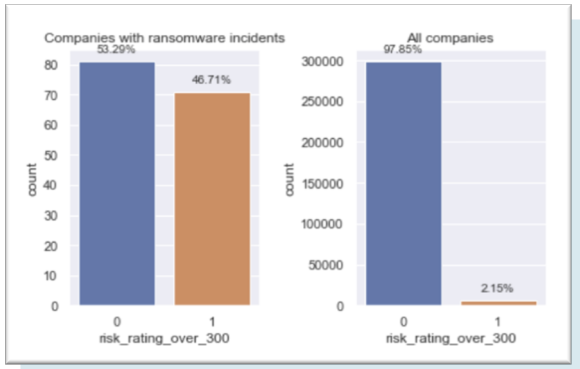
Leaked User Accounts (22x Discrimination Power)

- **Leaked user accounts can include email addresses used for registration, subscription, or account login.**
- **User accounts and online identities shared or sold on the dark web could be used to target individuals in spear phishing campaigns aimed at obtaining sensitive information.**
- **Ransomware is distributed by actors buying access to a company's secure network. Leaked user accounts on the dark web enable bad actors to access machines to deploy ransomware.**



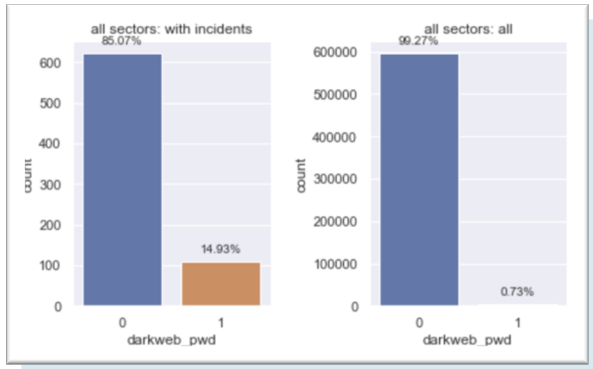
High Risk Rating (22x Discrimination Power)

- **While the current Cyence risk rating measures the probability that an individual company will experience a data breach in the next 12 months, it also has predictive advantage for identifying which companies are at higher risk of a ransomware incident. If a company has a risk rating of greater than 300, the company is 22 times more likely to have a ransomware incident.**
- **This correlation between risk rating and ransomware is logical—the risk signals are indicative of an attacker's capability to expose a company's weaknesses and compromise its defenses. Whether they get in and exfiltrate data or lock down systems, the entry point is just a means to that end, so we should see some similarity in those risk signals.**



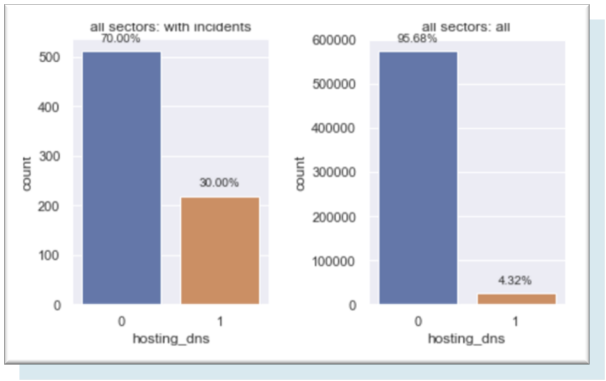
Compromised User Passwords (20x Discrimination Power)

- **Compromised login credentials consist of username and password pairs, which may be used to hack into private accounts. Combinations of employee usernames and passwords may be used by malicious actors to gain access to corporate accounts, especially given the prevalence of password reuse.**



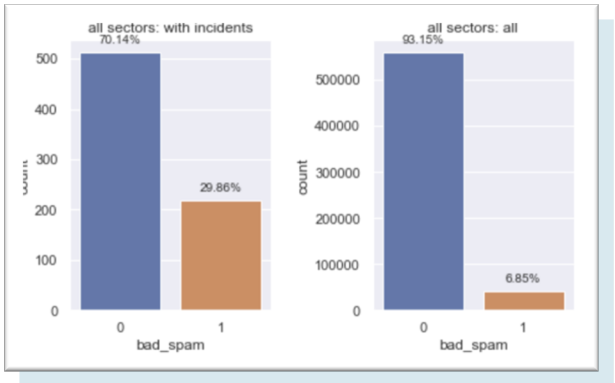
DNS Leakage (7x Discrimination Power)

- **This risk factor indicates a network misconfiguration or malicious signaling from internal network configurations that is published openly on the internet. It could allow attackers to monitor traffic requests and behavior.**



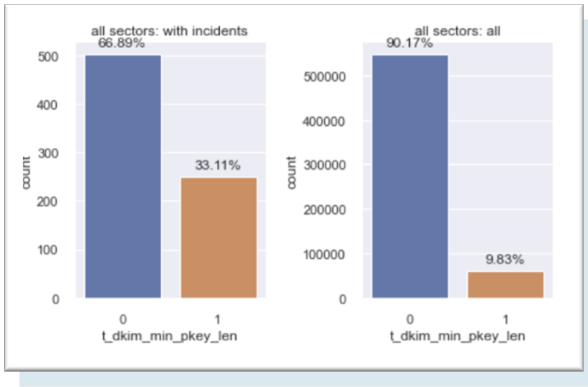
Spam Activity (4x Discrimination Power)

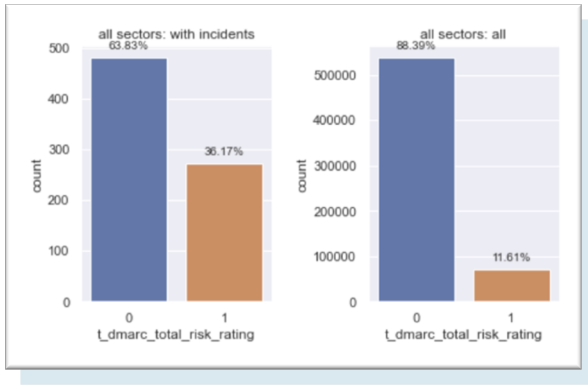
- **Spam involves the propagation of unsolicited junk email distributed to a large number of recipients. Mail servers being used to distribute spam may indicate system misconfiguration or compromised user credentials, which puts the company at risk.**



Email Misconfiguration (3–4x Discrimination Power)

- **DKIM (DomainKeys Identified Mail) is a leading standard for email authentication, which provides assurances that content has not changed from the sender's mail server to the recipient. Technically, this is achieved by an implementation of the standard public/private key signing process.**
- **DKIM duration and key length are features that are relevant, as they minimize exposure of keys to compromise by adversaries.**
- **DMARC (Domain-based Message Authentication, Reporting, and Conformance) is an email authentication, policy, and reporting protocol that allows a sender to indicate that their messages are protected by SPF and/or DKIM. It signals email authenticity, which helps combat spam and phishing.**





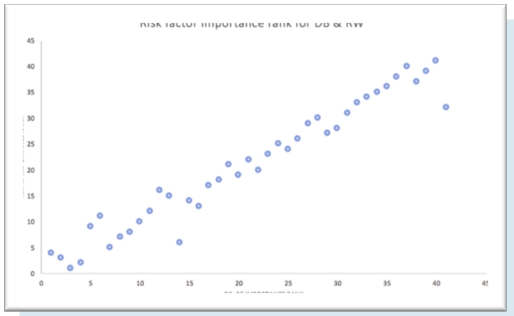
Relating Ransomware Risk Factors and Data Breach

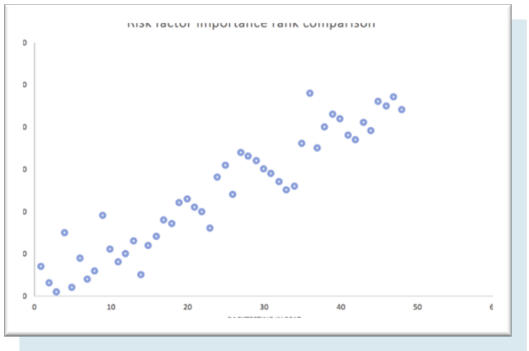
To test the theory that predictive risk factors for ransomware can also forecast data breach, Cyence used a similar methodology to detect the relationship between ransomware risk factors and data breach incidents. The results show the discrimination power of these risk factors for predicting whether a company will have a data breach incident.

As with ransomware association analysis, we calculated how the observation of specific risk factors can be used to help identify companies with a higher risk of data breach. This calculation can be used as a metric of risk factor importance in data breach incident prediction.

Note that this initial analysis focused on the single-variable discrimination power of risk factors. Further analysis aims to determine the predictive power of multi-variable analysis. The metric was calculated both against data breach incidents and individual ransomware incidents, which enabled us to compare how similarly/differently a risk factor impacts likelihood of different perils. Analysis revealed a monotone trend, meaning that when a risk factor is stronger in predicting a ransomware incident, it is also likely to predict a data breach incident more effectively.

- **Ranking the risk factors by descending importance (where 1= most important for both perils) and then plotting the data breach risk factors against ransomware, we discovered that risk factor ranking is highly linear between the two perils. This helps validate that ransomware and data breach share certain risk signals, and therefore we can leverage the risk factor value proposition for identification and selection of both types of incidents.**
- **As with ransomware, we backtested against data breach incidents in 2017 to compare the importance of risk factors between backtesting and previous analysis using all incidents. Again, we discovered a strong linear relationship, thus showing consistency between backtesting and holistic analysis for correlation between certain risk factors and data breach incidents.**





Insurance Implications

When reining in ransomware risk uncertainty, our perspective determines how we perceive risk. Since the ancient Egyptians, humans have leveraged visual aids to improve our ability to detect and use the world around us. When it comes to measuring and using signals of cyber risk, effective insurance demands better optics along the entire continuum: from cross-sector systemic risk, to interconnected portfolio risk accumulation, and down to granular enterprise-level exposures. While the past is not necessarily indicative of the future with regard to cyber risk, the history of threat playbooks, perennial vulnerabilities,[6] and security control deficiencies shows us that history does often repeat itself. Cyence empowers both a telescopic and microscopic capability to help anticipate and forecast ransomware and other cyber perils.

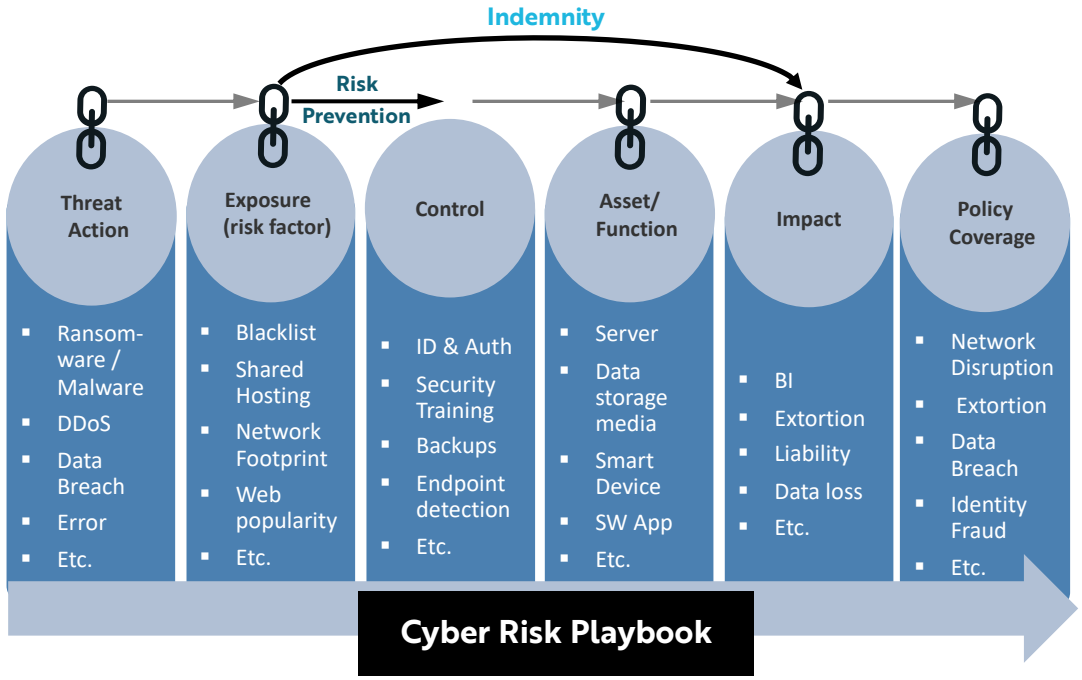
Specifically, discriminatory risk factors can help insurers to:

- **Identify if a company is at a higher risk than its peers of sustaining a successful ransomware incident**
- **Engage in meaningful conversations with current and potential policyholders about proactive risk controls and security management**
- **Comparatively rank firms in a portfolio based on categorical risk factors**
- **Draw on qualitative heuristics to zoom in on quantitatively derived questions related to the cost of ransomware attacks, as well as premium and sublimit strategies**
- **Inform trend analyses of cyber threats and exposures to help calibrate qualitative model output**
- **Lower loss ratios based on discriminatory risk factors, by tailoring policies and engaging in proactive risk management**

Closing the Gap Between Security Risk Management and Insurance

"What's driving a particular risk?" and "What are the maximum and expected loss exposures?" are top questions facing both companies and insurers. There is no simplistic single model for increasing the certainty of the answers. What's needed, rather, is a combination approach composed of data and model variables that offer insight into the cyber risk playbook: malicious *threats* that exploit *vulnerabilities* in systems and devices because of deficiencies in *controls*, which negatively impact valuable *assets and/or functions*, and result in *losses* that are transferable via insurance *policies*. The better we are at collecting and mapping data and variables according to this cyber risk playbook, the more we lower the inference risk and close the gap between risk inputs and negative outcomes.

[6] For an example, see https://www.us-cert.gov/ncas/alerts/aa20-133a.



Ransomware risk and anticipated losses are not due to unknown threat actions and indefinable defenses. Empirical incident response shows us that the following security controls are effective in preventing or mitigating this peril:

- **An IT business continuity and disaster recovery plan that includes multiple backups of important data on different media onsite and offsite, which is secured using industry- standard encryption**
- **Multi-factor authentication (MFA)**
- **Dynamic endpoint malware detection**
- **Network segmentation**
- **Employee phishing security training**
- **Vulnerability patching**

From a security controls perspective, there's no mystery regarding what can be done to lower ransomware risk. Yet from a risk-management perspective, the realities of resource constraints, information asymmetries, and risk triage conspire to keep these questions in the forefront.

By collecting and measuring many of these data and model variables according to the cyber risk playbook, Cyence reduces uncertainty in ransomware risk identification, quantification, selection, and pricing. Also, by mapping the discriminatory power of these risk signals to outcomes and controls, Cyence offers actionable insights and value propositions for both indemnity and risk prevention, respectively. Amid the growing body of risk signals, being able to triage the most impactful is key. Cyber risk factors and scores that lack relational associations to controls and impacts invite uncertainty and they prevent meaningful risk benchmarking. They leave one wondering, "So what?"

In addition, there is much room for contributory data and information sharing among insurers, policyholders, and other cyber risk stakeholders. Cyber risk models that are informed by combined risk signals along with incident claims and on-the-ground losses have appreciably stronger predictive power compared to models composed of only one part of the playbook.

Reducing the level of inference between components in the risk playbook will yield more-reliable cyber risk prediction, risk-management capability, and certainty. It achieves this by:

- **Reducing information asymmetries (i.e., how correlated are cyber perils, proving loss to an insurer, what cybersecurity risk controls are deployed) between the companies seeking to transfer risk and those providing risk transfer (insurers)**
- **Synchronizing the typically siloed intra-firm IT and risk-management functions**

Specifically, the level of cyber risk uncertainty will improve substantially if companies provide security and event measurements, loss validation, and more-detailed incident reports to cyber insurers, either directly or via impartial risk analytics intermediaries, like Cyence.

The auto insurance industry has only recently instrumented vehicles to collect digital telemetry, yet it has already incentivized the sharing of that data in exchange for premium reductions and tailored coverages. Companies have been producing security telematics from their systems for much longer, yet contributory risk data has been poorly incentivized in cyber insurance. Advances in secure multiparty computation and other disclosure-control technology[7] can allow unprecedented insights from sensitive, cross-organizational data while assuring the confidentiality of its sources. Also, cyber risk technology, like Cyence, that can observe, contextualize, and financially model telescopic and microscopic data is an essential component to optimize the transfer of dynamic cyber risk via continuous underwriting capability. Imagine having more-complete knowledge of claims, losses, and

[7] Other example technologies include differential privacy, homomorphic encryption, and federated model learning.

near-misses across insurers and policy lines without sacrificing competitive intelligence. Overlaying these capabilities atop a growing aggregated cyber risk leads to certainty between risk inputs and harmful outcomes—a mapping that both sides of the risk-transfer market have lost sight of.

Future-Proofing Ransomware Risk Uncertainty

Technology will continue to drive cyber crime innovation. The current trend toward automating ransomware execution tasks will allow attackers to focus on strategy tasks, resulting in higher-quality and higher-volume targeting.

Technology is raising the cyber peril cat-and-mouse game to the power of data science. This will manifest as red team playbooks enabled by machine learning and AI. Threats will be accelerated by machine-learning models that exploit vulnerabilities and hasten intelligent evasion, system infection and hijacking, and data acquisition faster than non-automated defenses can patch or react. Intelligent targeting, for example, will leverage training data about how individual employees communicate and respond to various phishing messages to create machine-learning models that successfully impersonate legitimate messages—to dodge spam filters and click-bait victims into visiting infected websites or sending sensitive data to criminals.

The implications for risk control and transfer are nontrivial, yet the uncertainty is manageable. Automated and intelligent red team playbooks can be countered with defense playbooks that execute at machine speed with machine-learning and AI prescience. The foundational knowledge to develop, train, and calibrate these advanced security models depends on scalable observation, synthesis, and orchestration of these risk signals according to the cyber risk playbook. Platforms that can manage, coordinate, and model this telescopic and microscopic data and knowledge are prerequisite for enabling such a foundation. In addition to aiding automated defenses, such platforms can further optimize cyber risk management and underwriting by increasing data quality and reducing acquisition costs associated with risk and control selection. Just as *science* is the discipline of how we change what we know, *Cyence* enables the observation and experimentation of how we change what we know about the structure and function of evolving cyber risk.

Guidewire is the platform P&C insurers trust to engage, innovate, and grow efficiently. We combine digital, core, analytics, and AI to deliver our platform as a cloud service. More than 380 insurers, from new ventures to the largest and most complex in the world, run on Guidewire. For more information, contact us at info@guidewire.com.

© 2020 Guidewire Software, Inc. For more information about Guidewire's trademarks, visit http://guidewire.com/legal-notices. Document Published: 2018-08-02