Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S.J.E. Taylor, eds.

A THEORY TO QUANTITATIVELY ESTIMATE AND BOUND SYSTEMIC CYBER RISK

Ranjan Pal[1], Konnie Duan[2], Sander Zeijlemaker[1], and Michael Siegel[1]

[1]MIT Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA
[2]Department of EECS, Massachusetts Institute of Technology, Cambridge, MA, USA

ABSTRACT

Business enterprises have grappled in the last one and half decade with unavoidable risks of (major) cyber incidents on critical infrastructure (e.g., power grid, cloud systems). The market to manage such risks using cyber insurance (CI) has been growing steadily (but not fast enough) as it is still skeptical of the extent of economic and societal impact of systemic cyber risk across networked supply chains in interdependent IT-driven enterprises. To demystify this skepticism, we study in this paper the role of (a) the statistical nature of multiple enterprise cyber risks contributing to aggregate supply chain risk and (b) the graph structure of the underlying enterprise supply chain network, in the statistical estimation and spread of aggregate cyber risk. More specifically, we provide statistical tail bounds on the aggregate cyber-risk that a cyber risk management firm such as a cyber insurer is exposed to in a supply chain.

1 INTRODUCTION

Digitally driven enterprise supply chains are becoming increasingly pervasive. Popular examples include the enterprise supply chains supported by critical infrastructure such as the power grid and cloud systems (see Figure 1). While the modern power grid built upon a cyber-physical system supports virtually all societal enterprise sectors, the cloud systems supporting SaaS, PaaS, and the IaaS service paradigms are the backbone of nearly every business today. Consequently a service disrupting cyber attack (or system configuration induced reliability faults) on such critical infrastructures will simultaneously cripple/disrupt the services offered by enterprises on a supply chain in a systemic fashion and increase business risk. The business risk arising due to such a cyber attack is often termed as systemic cyber risk.

To give examples of systemic cyber risk, on 16th November 2021, Google's cloud platform faced a global outage sourcing from a network configuration bug that lasted for two hours and cascadingly (systemically) bought down major services like Spotify and Facebook that themselves have millions of (business) clients who use their services. The *NotPetya* cyber attack of 2017 is an example of a malware-based supply chain attack that crippled the services of major organizations such as Maersk, FedEx, Mondelez, Reckett Benkeiser (all of them commonly using a tax-filing software affected by malware) for multiple weeks systemically affecting hundreds and thousands of their clients business services. Similar incidents, specific to the energy and utility service sectors, include the *Colonial Pipeline* cyber attack of 2021, and the Ukraine power grid cyber attack of 2015, where the services of multiple businesses reliant on energy and power were disrupted for hours/days. All of such business disruption incidents cost national economies hundreds of millions (if not billions) of dollars, and most of these costs till date are absorbed by affected enterprises on the digital supply chain.

1.1 Cyber Insurance Markets to Manage Risk

One could argue here that it is quite natural that enterprises around the globe should resort to cyber insurance (CI) to cover their business losses (be it first-party, or multi-party as usual in digital supply chain settings), similar to that in traditional sectors like health, property, life, and automobile.

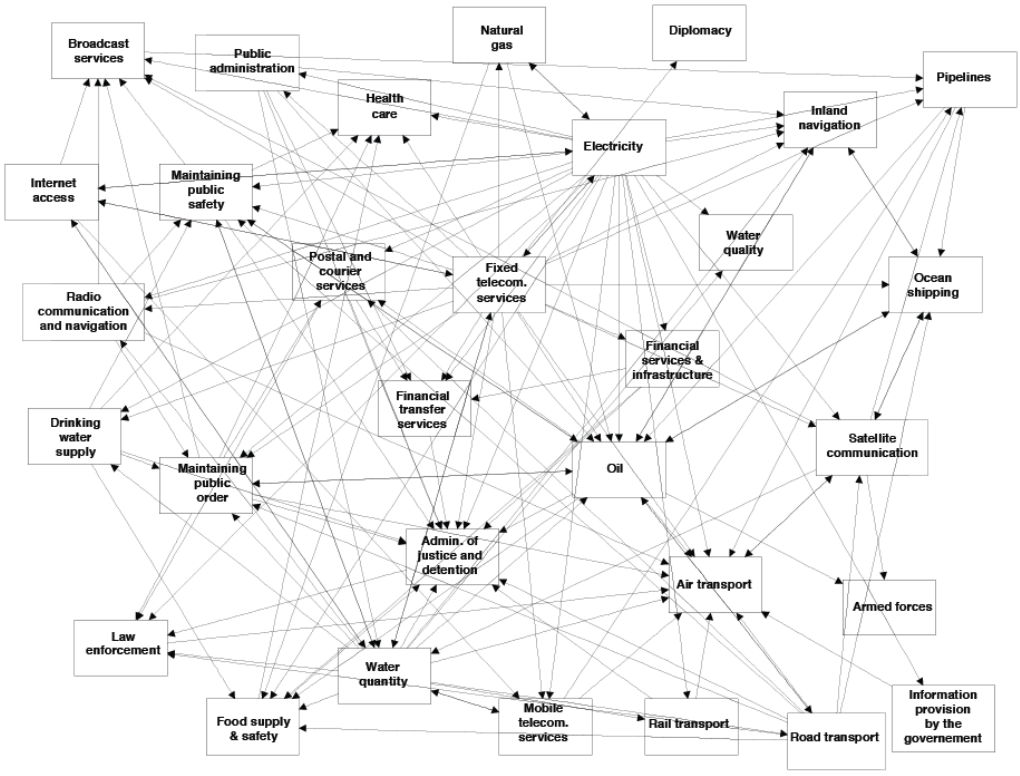

Figure 1: Showcasing the chart of complex service dependencies among networked enterprise sectors with critical cyber infrastructure. [*Source:* Netherlands Organization for Applied Scientific Research (TNO)].

1.1.1 Slow Growth in the CI Market

Though the concept of insurance for cyber-risk management was introduced in the late 90s, the market for such insurance had a sub-steady growth until early 2010s (Wolff 2022). The commercial cyber insurance (CI) market to manage enterprise cyber risk has seen a significant growth rate in the last five years in the USA (greater than 50% annually since 2020 according to *Conning*, an investment management firm for the insurance industry) and in many other countries around the globe. Most of it is due to the spike in ransomware claims (and business email compromise) since 2019 (based on a survey by insurance analytics firm *NetDiligence* in 2023) that has allowed cyber insurers to write higher premiums on a narrow exposure base with tightened exclusionary policies. The benefit of such policies apart from their obvious role in business loss coverage is that they have usually enforced that enterprises adhere to strictly recommended security controls (e.g., MFA, strong passwords) for attractive premiums and/or the necessary condition to get or renew insurance coverage (Wolff 2022). A recent survey by *Forrester* in 2023 provided statistics showing that such security controls necessitated on enterprises by (standalone) cyber insurance policies reduced the number of cyber incidents involving data breaches and also improved cyber resilience by reducing the mean time to incident detection, response, and recovery. Despite the promising statistics state above in favor of the growth of the CI industry, most cyber risk is absorbed by the insured client. In other words, there is big and growing annual supply-demand gap (of insurance premiums versus total cyber-crime costs) in the current cyber insurance industry running in atleast hundreds of billions of USD (see Figure 2). It is evident from Figure 2 that cyber insurers are skeptical to bear the bulk of global cyber-crime costs.

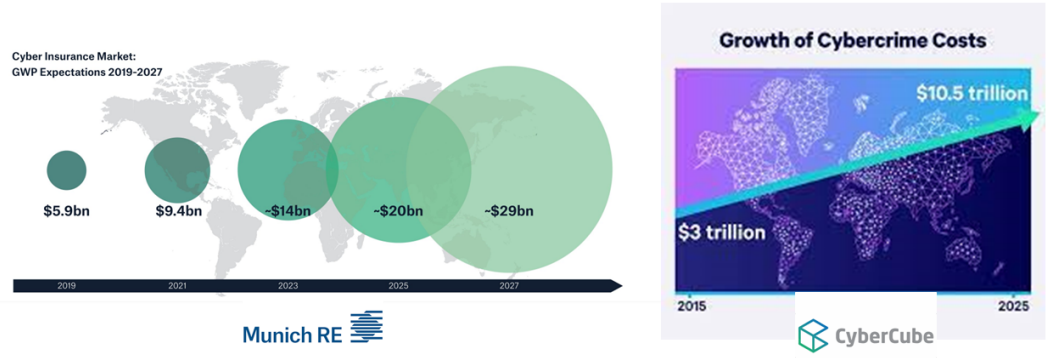

Figure 2: A large supply-demand gap exists in the global CI market. [*Source:* MunichRe and CyberCube]

1.1.2 The Core Challenge to CI Market Growth

The phenomenon of multi-sided *information asymmetry* (IA) is the core challenge to scaling CI market growth with respect to closing the wide supply-demand gap. The primary catalyzing factor behind the limited growth rate in CI markets is that the cyber risk terrain is extremely dynamic and expansive over time, and quite varying in space among enterprises. This factor makes it quite difficult for risk-averse cyber insurers to accurately assess business cyber-risk in both, non-systemic and systemic settings. This is characteristic of the multi-sided *information asymmetry* problem unique to cyber insurance environments (when compared to the single-sided information problem in traditional insurance) where both (and hence the term 'multi-sided'), the enterprise client and the insurance companies not have perfect knowledge of all the vulnerabilities in the hardware, software, and firmware stacks of computer systems (the vulnerability stacks combined forming the cyber terrain) in the enterprises, to be able to accurately assess business cyber risk (Pal et al. 2023; Pal et al. 2021; Pal et al. 2024). The vulnerability space in this cyber terrain will get much larger with the integration of AI in the computer systems space. In short, it often is too costly for both the (potentially) insured and the insurer to get and share the necessary cyber risk terrain information to price loss coverage contracts that are a 'win-win' for both the supply and demand sides of the CI business. Eventually, this leads to insurers either rejecting enterprise client demands to purchase cyber insurance, or designing contracts with high deductibles and premiums that attract relatively a low number of clients.

1.2 Information Asymmetry Challenge Amplifies in Systemic Settings

IT and the operational technology (e.g., IoT and CPS) driven enterprise sectors (mostly comprising of small and medium businesses) are becoming increasingly part of enterprise digital supply chains (see Figure 1), and in the near future will be the backbone of the daily needs of every society around the globe.

1.2.1 Multi-Sided IA in Systemic Settings

The net adverse economic impact from an adversary exploiting vulnerabilities (in multi-sided IA terms, the '(un)known-unknowns' of an enterprise) in a single enterprise on the digital supply chain is much amplified in this IT/OT-driven networked and interdependent enterprise setting that is quite susceptible to the relatively newer family of domino-style cascading/systemic cyber loss impact. This impact is an outcome of highly non-linear amplifier effects in any complex (supply chain) network. It is not only the IA impact, but the aforementioned multi-sided IA problem is also amplified in systemic supply chain settings due to networked dependency effects. In other words, the individual enterprise multi-sided IA problem is scaled on the order of the size of the underlying supply chain network. At the same time, risk information sharing among service interdependent supply chain enterprises has always been a challenge till date. Given that the CI market thus far has exhibited limited growth rates (with respect to closing the supply-demand

gap) in a fairly non-systemic risk environment, it is unlikely that this market trend would change (probably even worsen) in the evolving and potentially pervasive systemic cyber risk environments. We next state real-world examples where insurers due to lack of low historical incident data about systemic cyber attacks did not have enough cyber risk estimation model confidence to absorb a significant amount of economic and societal impact from such attacks.

1.2.2 Examples of IA Making CI Markets Conservative

Ransomware-as-a-service that started around 2014 hardly used to exceed multiple thousands of USD until 2019, when suddenly nation state actors demanded multi-million dollars on a single ransomware incident or from a systemic impact of a single cyber incident. *Such impact spikes are extremely difficult to predict using limited availability of historical data and cyber-posture information fueled by multi-sided IA*, and dampened cyber insurance market interest to cover business losses from such incidents post 2020. Add to this that systemic catastrophe loss events (including those larger than any we have seen till date) affecting multiple enterprises at once could very likely and unpredictably reverse the trend of CI markets growing on narrow exposure bases in the aftermath. As an example, the 2017 *NotPetya* supply chain cyber attack impact of nearly USD 10 billion initially pushed leading cyber insurers to opt out of providing demanded coverage amount. *The insurers simply never planned to cover such scales of systemic risk due to a multi-sided IA driven lack of historical data and a subsequent risk analysis based on stress testing, alongside and policy shortcomings.* As another very recent example, the global CrowdStrike IT outage of 2024 is likely cost the Fortune 500 companies, excluding Microsoft, at least USD 5.4 billion in direct financial losses, whereby cyber insurance will only cover 10% to 20% of the losses. *Such scales of under coverage is a direct outcome of (a) the multi-sided* information asymmetry *problem and (b) lack of a stress tested risk analysis, that together does not help generate sufficient risk statistics for insurers to improve coverage percentages.* According to *Parametric* and *CyberCube* - leading cyber insurance analytics firms, the cyber insurance market will likely face preliminary insured losses of between USD 400 million and USD 1.5 billion from the *CrowdStrike* incident - evidently a wide variation lacking statistical tightness of projections.

1.3 Research Motivation and Contributions

We state our research motivation and follow it up with our research contributions.

1.3.1 Research Motivation

The multi-sided information asymmetry problem is here to stay, and the opportunity cost for insurers to not tap into the multi-trillion dollar cyber risk coverage market is too high. With existing (a) cyber-posture estimating tools by firms such as *Bitsight*, and (b) third party cyber risk mitigation tools like *RiskRecon* by *Mastercard*, there is scope for getting more leverage from cyber risk data science in favor of systemic cyber risk management. In other words, a worst case stress tested analysis on systemic cyber risk statistics done apriori (using statistics estimated by industry standard cyber-posture tools like *Bitsight* and *RiskRecon*) can go a long way in mitigating the adverse effects of multi-sided IA and better systemic cyber risk management by the CI industry and reduce (re)insurer opportunity costs. Our research motivation in this paper is to quantitatively estimate and study tight (worst case) bounds on the systemic cyber risk faced by an enterprise on a supply chain and their impact on systemic cyber risk management by the CI industry.

1.3.2 Research Contributions

A particular cyber insurer will likely be exposed to an aggregate cyber risk stemming from first and multi-party claims in a digital supply chain. We provide a statistical theory to study the spread of the tail of an aggregate number of general enterprise cyber risks (reflecting both, light-tailed and heavy-tailed risk

statistics, and for both, i.i.d. and non i.i.d. settings) when the underlying network degree distribution of the enterprise supply chain network is both, light-tailed and heavy-tailed in nature (see Sections 2 to 4).

The *first* part (see Section 2) studies the spread and bound of the aggregate tail when individual and i.i.d. cyber risk distributions are light-tailed, but the number of such risks to be aggregated is sampled from a heavy-tailed distribution in the worst case. Alternatively, the sample of the number of cyber risks is from an any-tailed distribution (heavy or light). The *second* part (see Section 3) studies the spread and bound of the aggregate tail when individual i.i.d., cyber risk distributions are heavy-tailed, but the number of such risks to be aggregated is sampled from an any-tailed (heavy or light) distribution. The *third* part (see Section 4) studies the spread and bound of the aggregate tail when individual non i.i.d., cyber risk distributions are any-tailed (heavy or light), but the number of such risks to be aggregated is sampled from a light-tailed distribution. In practice, individual enterprise risk exposure distributions post any cyber attack could be dependent, i.e., non i.i.d. between enterprises. While this is true, it is also the case that individual enterprise IT infrastructures mutually differ. Hence, systemic enterprise impacts post a cyber attack are likely i.i.d. also. Our analyses in this paper accounts for both these settings. A summary of various (systemic) cyber risk settings for Sections 2-4 in showcased in Figure 4.

Our proposed statistical theory is based upon real-world enterprise cyber risk distribution types and supply chain network topology data obtained from *Bitsight*; Mastercard's *Cyber Quant* cyber risk quantification (CRQ) product; and Mastercard's *RiskRecon* multi-party risk evaluation product. Our proposed theory to study aggregate/sytemic cyber risk is the first (to the best of knowledge) to account for general cyber risk distributions individual enterprises in a supply chain is exposed to along with the underlying supply chain network structure impacting aggregate cyber risk.

2 RELATED WORK

In this section, we briefly review research related to residual cyber risk management markets.

Cyber Insurance to Improve Cybersecurity - It is only because of the inherent potential of cyber insurance to improve enterprise security governance that we have a market for third-party (systemic) risk transfer. This proven potential of cyber-insurance to improve cybersecurity has been mathematically shown in seminal papers (Lelarge and Bolot 2009; Shetty et al. 2010; Hofmann 2007; Pal and Golubchik 2010; Pal et al. 2014; Naghizadeh and Liu 2014; Pal et al. 2018; Pal et al. 2011; Pal et al. 2017; Yang and Lui 2014). In practice, cyber insurance markets have steadily seen an increase over the years (specifically, since the last decade and a half) with cyber insurance solutions demanding sufficient cybersecurity controls on part of enterprise clients to be contract-eligible, or receive significant coverage. This consequently supports these enterprises to effectively adapt, absorb, and respond to cyber incidents. The readers are referred to (Dambra et al. 2020; Marotta et al. 2017) for a review of the role of cyber insurance.

Methodologies to Manage Systemic Cyber-Risk - Systemic cyber risk management is one of the major applications feeding into CI markets. The authors in (Pal et al. 2023; Pal et al. 2021; Pal et al. 2024) have shown that optimally estimating and diversifying systemic cyber risk (an integral insurance operation to manage a portfolio of cyber risks) is NP-hard. However, the hardness of optimally estimating and diversifying (systemic) cyber-risk does not deter the existence of non-optimal but diversification sustainable portfolios of (systemic) cyber risk for cyber re-insurers. Recent theoretical efforts investigated the diversification sustainability problem for i.i.d. cyber risk portfolios. In a series of efforts (Pal et al. 2020; Pal et al. 2020; Pal et al. 2020; Pal et al. 2023; Pal et al. 2021), the authors have proved that diversifying a portfolio of *catastrophic* heavy-tailed cyber risks (each having infinite mean and potentially sourced from individual risks from multiple enterprises) that are identical and independently distributed (i.i.d.), i.e., not tail-dependent, *is not* an effective economically sustainable practice for reinsurers with respect to the industry-popular Value-at-Risk (VaR) tail risk measure. On the other hand, diversifying a portfolio of i.i.d. heavy-tailed cyber risks that are *not catastrophic* (risks with finite mean and sourced from individual risks from multiple enterprises) is economically sustainable for reinsurers. However, (systemic) cyber risks are often generated from non i.i.d. individual enterprise cyber risk sources. In (Pal et al.

2024; Pal et al. 2025), the authors derive the conditions for economic sustainability of systemic cyber risk portfolio diversification when such portfolios that cyber re-insurers are exposed to, consist of non i.i.d. risks with arbitrary tail nature. The authors in (Pal et al. 2023; Pal and Nag 2023) propose models on market efficiency of ILS-driven CI markets and their pricing, respectively - however, they do not propose a decision science as to when it is suitable for CI companies to (a) only invest in retaining cyber risk of their enterprise clients, (b) transfer their risk to reinsurance companies, and (c) transfer their risk using a combination of reinsurance and CAT bonds. This is the main decision problem for the CI industry.

Estimating and Bounding Systemic Cyber Risk - A fundamental principle from the field of management sciences is that if we cannot estimate something, we cannot manage that thing. This principle could not hold more true for (systemic/aggregate) cyber risk management. There have been works in the broad and historically rich field of risk analysis to estimate aggregate risk, that consequently envelop estimating systemic risk including cyber risk. Aggregate risk distributions are usually evaluated as a function of the distribution of the number of risks to be aggregated and the statistical nature of each such risk to be aggregated, and this evaluation approach directly extends to cyber risks. The authors in (Beard 2013; DasGupta and DasGupta 2008; Embrechts et al. 2013; Denisov et al. 2010; Stam 1973) assume (a) each of the risk distributions to be aggregated are i.i.d. and (b) both, the claims number distribution and individual risk distributions are light-tailed or at most sub-exponential. However, in reality and in the specific context of cyber risk, individual cyber risk distributions sourcing from enterprises are likely to be non i.i.d. in nature (due to correlated nature of cyber risk) alongside being statistically heavy-tailed (e.g., reflecting risk outcomes pertaining to the *NotPetya* cyber attack). In addition, the claim number distribution is likely to be heavy-tailed in certain situations of catastrophic systemic cyber risk (such as a power grid failure or a large public cloud provider outage). In relation to obtaining tight bounds of systemic cyber risk, the authors in (Wang and Wang 2007; Loukissas 2012; Ng et al. 2004; Lu 2012) design methodologies to estimate lower bounds of aggregated risk sourced from i.i.d. distributions, where each risk distribution to be aggregated is statistically light-tailed in nature. However, in the specific context of cyber risk, such distributions can be heavy-tailed as aforementioned alongside being non i.i.d. *In this paper, we extend the aforementioned works to fit the cyber context and alleviate existing modeling drawbacks related to the statistics of systemic cyber risk and the number of such risks (re)insurers are exposed to.* A review of (non-formal) methods to estimate cyber risk in non-systemic settings is surveyed in (Woods and Böhme 2021).

3 SCENARIO 1 - ANY-TAILED # OF IID LIGHT-TAILED RISKS

Such a cyber risk aggregation scenario is faced by a cyber reinsurer when, as examples, a non-catastrophic data breach cyber attack or DDoS service unavailability cyber attack affects multiple organizations in a systemic manner. Here, the tangible multi-party loss impact faced by individual enterprises under the coverage span of a cyber insurance (CI) firm is usually low/medium in size and hence a sample from a light-tailed risk distribution. Since individual enterprise IT infrastructures differ, individual enterprise loss claims are assumed to be i.i.d. post a systemic cyber attack incident. However, the number of enterprises affected by such incidents (i.e., # of claimants) is in the worst case likely very large and could be a sample from (the tail of) a heavy-tailed distribution. *In the context of an enterprise supply chain setting, this is the degree distribution of the supply chain network.*

Assume that N is the number of cyber risks a cyber insurer is exposed to post a cyber attack event, and each of these N risks are sourced from enterprises who are clients of the insurer. Each of these risks are reflective of multi-party claims an enterprise is exposed to, post a cyber attack incident, from its downstream supply chain of clients. The insurer transfers the aggregate of these N individual cyber risks to a re-insurer as the latter has significantly more capital to manage the tail of aggregate risk. Let R_i be the cyber risk distribution sourced from enterprise i. An illustrative figure of such an aggregate cyber risk transfer setting is shown in Figure 3. We have the following result estimating the tail of the aggregate $A_N = \sum_i^N R_i$, a cyber reinsurer is exposed to from a cyber insurer.

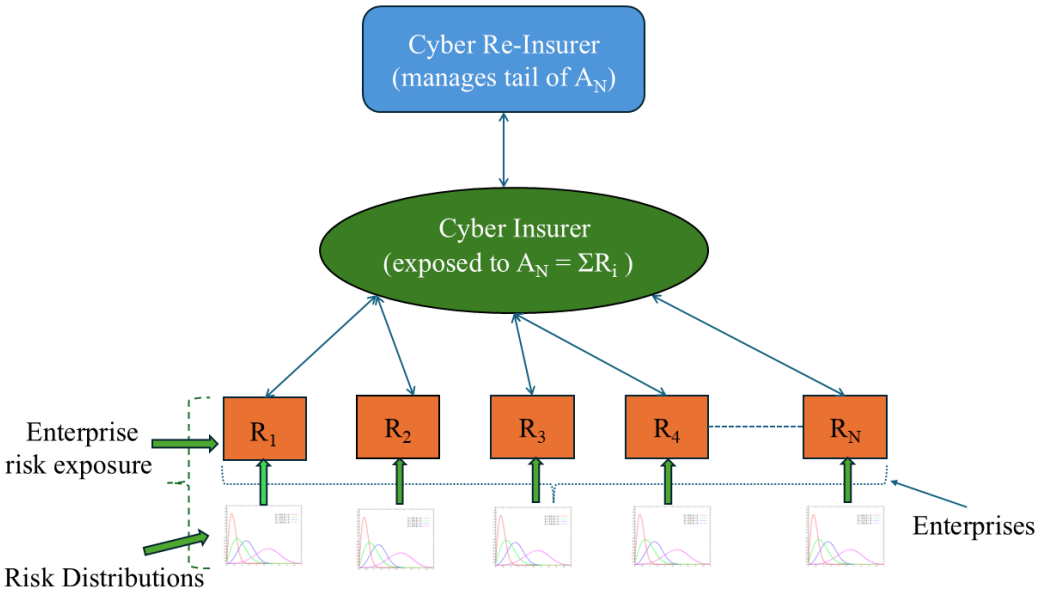

Figure 3: An logical structure of aggregate cyber risk transfer between enterprises and cyber (re)insurers.

Theorem 1 *Consider N sources (from being small to very large) of enterprise cyber risks* $\{R_i\}_{i=1}^N$ *that a cyber insurer is exposed to post a cyber attack event, whose aggregate is of interest to a cyber reinsurer (to whom the insurer often transfers significant portions of aggregate risk management liability). Let* $\{R_i\}_{i=1}^N$ *be statistically i.i.d. and each having a finite mean and variance (i.e.,* $\mathbb{E}[R_i^t] < \infty, \forall i$*, for some* $t > 1$*,* $\mathbb{E}[R_i] = \mu, \forall i$*), and let N be sampled from an any-tailed (worst case heavy-tailed) distribution with finite mean, whose tail is of consistent variation. Then,*

$$\mathbb{P}[A_N > r] \approx \mathbb{P}[N > \frac{r}{\mu}], \text{as } r \to \infty, \text{if } \mathbb{P}[R_i > r] = o(\mathbb{P}[N > r]), r \to \infty, \forall i.$$

Any distribution F for N is said to be of consistently varying tail (Cline 1994) if $\lim_{y\to 1} \lim_{x\to\infty} \sup \frac{\bar{F}(xy)}{\bar{F}(x)} = 1$*, where* $\bar{F} = 1 - F$*.*

Proof Sketch - Given the tail of N has consistent variation, we have the following relation:

$$\lim_{\varepsilon\to 0} \lim_{r\to\infty} \frac{\mathbb{P}[N > (1+\varepsilon)r]}{\mathbb{P}[N > (1-\varepsilon)r]} = 1.$$

Hence, the upper *Matuszewska* index (Bingham et al. 1989) $\alpha_N < \infty$, where $\alpha_N = \lim_{y\to\infty} \frac{-\log\{\bar{F}_*(y)\}}{\log(y)}$, and $\bar{F}_*(y) = \lim_{r\to\infty} \inf \frac{\bar{F}(ry)}{\bar{F}(r)}; y > 0$. If we pick a $\rho > \alpha_N$, then $r^{-\rho} = o(\mathbb{P}[N > r];\ r \to \infty$. Here, ρ_F for any distribution F is defined as: $\rho_F = \lim_{r\to\infty} \sup \frac{-\log\{\bar{F}(r)\}}{\log(r)}$. We also have the following relation from (Liu 2009)

$$\mathbb{P}[A_{\lfloor(1-\varepsilon(r)\rfloor} > r)] \leq \lfloor(1-\varepsilon)r\rfloor \mathbb{P}[R_i > v\varepsilon r] + C(\varepsilon r)^{-\rho}, \forall i.$$

Hence, $\mathbb{P}[A_{\lfloor(1-\varepsilon(r)\rfloor} > r)] = o(\mathbb{P}[N > r], r \to \infty$, and using *Chernoff* bounds (Hagerup and Rüb 1990), we get $\mathbb{P}[A_{\lceil(1+\varepsilon(r)\rceil} \leq r)] = o(\mathbb{P}[N > r], r \to \infty$. This proves the theorem.

Cyber Insurance Industry Implications - This result showcases that if cyber (re)insurers simply have (historical) knowledge about (a) the distribution of # of claimants post a systemic cyber attack incident, and (b) a rough average estimate of the mean loss impact value of individual light-tailed loss distribution post a cyber incident, they can estimate the tail of aggregate cyber risk, i.e., systemic cyber risk, sourcing from these individual claimants.

4 SCENARIO 2 - ANY-TAILED # OF IID HEAVY-TAILED RISKS

Such a cyber risk aggregation scenario is faced by a cyber reinsurer when, as examples, catastrophic (CAT) cyber attacks of the service unavailability type such as *NotPetya* or a public cloud outage for many hours/days affects multiple organizations in a systemic manner. Here, the tangible worst-case multi-party loss (from enterprise downstream supply chains) impact faced by individual enterprises under the coverage span of a cyber insurance (CI) firm is usually a sample from a heavy-tailed risk distribution and hence could be very large. Since individual enterprise IT infrastructures differ, individual enterprise loss claims are assumed to be i.i.d. post a systemic cyber attack incident. However, the number of enterprises affected by such incidents (i.e., # of claimants) could range from being medium in size to likely being very large and often a sample from any any-tailed distribution, i.e., a light-tailed distribution or a heavy-tailed distribution. *From an enterprise supply chain viewpoint, this is the degree distribution of the supply chain network.*

Assume that N is the number of cyber risks a cyber insurer is exposed to post a cyber attack event, and each of these N risks are sourced from enterprises who are clients of the insurer. The insurer transfers the aggregate of these (worst case heavy-tailed) N individual cyber risks to a reinsurer as the latter has significantly more capital to manage the tail of aggregate risk. Let R_i be the cyber risk distribution sourced from enterprise i. We have the following result estimating and bounding the tail of the aggregate $A_N = \sum_i^N R_i$, a cyber reinsurer is exposed to from a cyber insurer.

Theorem 2 *Consider N sources of enterprise cyber risks $\{R_i\}_{i=1}^N$ that a cyber insurer is exposed to post a cyber attack event, whose aggregate is of interest to a cyber reinsurer (to whom the insurer often transfers significant portions of aggregate risk management liability). Let $\{R_i\}_{i=1}^N$ be statistically i.i.d and each having a finite mean (i.e., $\mathbb{E}[R_i] = \mu, \forall i$) but possibly infinite variance, and let N be sampled from an any-tailed (either heavy or light-tailed) distribution with finite mean λ, whose tail is of consistent variation. Then,*

$$\mathbb{P}[A_N - \lambda\mu > r] \gtrapprox \lambda \bar{F}'(r + \lambda\mu)\ \forall \gamma \geq 0,$$

always holds uniformly for $r > \gamma\lambda$ and

$$\mathbb{P}[A_N > r] \gtrapprox \lambda \bar{F}' \forall \gamma \geq 0,$$

always holds uniformly for $r > \gamma\lambda$.

Here, (a) any distribution F for N is of consistently varying tail (Cline 1994) if $\lim_{y\to 1} \lim_{x\to\infty} \sup \frac{\bar{F}(xy)}{\bar{F}(x)} = 1$, where $\bar{F} = 1 - F$, and (b) $\bar{F}' = 1 - F'$; $\bar{F}'$ is the distribution of individual $\{R_i\}_{i=1}^N$.

Proof Sketch - Using the law of large numbers, the convergence in probability result is $\frac{A_N}{\lambda} = \frac{N}{\lambda}\frac{1}{N}\sum_{j=1}^N R_i \to_p \mu$. We then have

$$\mathbb{P}[A_N - \lambda\mu > r] \geq (1 - 2\delta)P'; P' = \mathbb{P}[A_N - \mu\lambda > r],$$

where $\delta \in (0, 0.5)$. This leads us to

$$\mathbb{P}[A_N - \mu\lambda > r] \geq (1 - 2\delta)(1 - \delta)\lambda\bar{F}'(r + \mu\lambda),\ \delta \in (0, 0.5).$$

Allowing δ to converge towards 0, we get $\frac{\mathbb{P}[A_N - \mu\lambda > r]}{\lambda\bar{F}'(r+\lambda\mu)} \gtrapprox 1$. This proves the theorem.

Cyber Insurance Industry Implications - This result implies that if cyber (re)insurers have knowledge about (a) mean of the # of claimants post a systemic cyber attack incident, and (b) a rough average estimate of the mean loss impact value of individual heavy-tailed loss distributions post a cyber incident, they can estimate and bound the tail of aggregate cyber risk, i.e., systemic risk, sourced from individual claimants.

5 SCENARIO 3 - LIGHT-TAILED # OF NON-IID ANY-TAILED RISKS

Such a cyber risk aggregation scenario is faced by a cyber reinsurer when, as examples, non-catastrophic and catastrophic cyber attacks of the service unavailability type such as a public cloud outage for many

IID/Non-IID	# of Claims	Type of Risk	Example Cyber Attack Scenario	Section
IID	LT	LT	small/medium data breach, malware affected IT firms	2.1
IID	LT	HT	ransomware-driven service outage of IT services	2.2
IID	HT	LT	cyber-driven IT failures in any given time interval	2.1
IID	HT	HT	massive data breach at private/public cloud provider	2.2
Non-IID	LT	LT	Data loss/theft from a cloud/email service provider	2.3
Non-IID	LT	HT	power grid outage for few days in a small town	2.3
Non-IID	HT	LT	DDoS attack on popular server, OS exploit shutdown	OP
Non-IID	HT	HT	public cloud outage and Internet outage for days	OP

LT: Light-Tailed Distribution (finite mean, finite variance)

HT: Heavy-Tailed Distribution (finite mean, infinite variance)

OP: Open Problem

The '# of Claims' is the distribution of the number of enterprises affected post a cyber attack

The '# of Claims' is a proxy of the degree distribution of an underlying enterprise supply chain network

Figure 4: A summary of systemic (aggregate) cyber risk settings (and related cyber attack scenarios).

hours/days, national power grid outages, or data breach impact affects multiple organizations in a systemic manner. *The core property of such aggregate risk, especially post catastrophic cyber attack, is that individual risks are correlated (i.e., non i.i.d.) with one another post a cyber attack event.* Here, the tangible multi-party loss impact faced by individual enterprises under the coverage span of a cyber insurance (CI) firm is usually a sample from a general light-tailed or a heavy-tailed risk distribution and hence could vary from being small to being very large. Similarly, the number of enterprises affected by such incidents (i.e., # of claimants) could range from being small/medium in size to likely being very large and often a sample from either from a light-tailed distribution or a heavy-tailed distribution. *In the context of an enterprise supply chain setting, this is the degree distribution of the supply chain network.* For the purpose of analytical tractability, we only work with the case when this number is sampled from a light-tailed distribution. The case of a sample from a heavy-tailed distribution is left for future work.

Assume that N is the number of cyber risks a cyber insurer is exposed to post a cyber attack event, and each of these N risks are sourced from enterprises who are clients of the insurer. The insurer transfers the aggregate of these (worst case heavy-tailed) N individual cyber risks to a re-insurer as the latter has significantly more capital to manage the tail of aggregate risk. Let R_i be the cyber risk distribution sourced from enterprise i. We have the following result estimating and bounding the tail of the aggregate $A_N = \sum_i^N R_i$, a cyber reinsurer is exposed to from a cyber insurer.

Theorem 3 *Consider N sources of enterprise cyber risks $\{R_i\}_{i=1}^N$ that a cyber insurer is exposed to post a cyber attack event, whose aggregate is of interest to a cyber reinsurer (to whom the insurer often transfers significant portions of aggregate risk management liability). Assume each R_i to have a distribution function F_i. Let $\{R_i\}_{i=1}^N$ be non i.i.d and each having a finite mean (i.e., $\mathbb{E}[R_i] = \mu, \forall i$) but possibly infinite variance. Let N come from a light-tailed distribution with finite mean $\lambda < \infty$, and a tail of consistent variation. Then,*

- ***Case A**: When $N = 2$, aggregate cyber risk A_N is bounded by*

$$F_{\min}(r) \lessapprox \mathbb{P}[A_N \leq r] \lessapprox F_{max}(r), \forall r \in \mathbb{R},$$

where

$$F_{min}(r) = \sup_{\vec{r}|r_1+r_2=r} \max\left\{F_1^-(r_1) + F_2^-(r-r_1) - 1, 0\right\}$$

and

$$F_{max}(r) = \inf_{\vec{r}|r_1+r_2=r} \min\left\{F_1(r_1) + F_2(r-r_1), 1\right\},$$

with $\{F_i^-\}_{i=1}^N$ being the left limit of distribution $\{F_i\}_{i=1}^N$.

- ***Case B**: When $N \geq 3$, aggregate cyber risk A_N is bounded by*

$$F_{\min}(r) \lessapprox \mathbb{P}[A_N \leq r] \lessapprox F_{max}(r), \forall r \in \mathbb{R},$$

where

$$F_{min}(r) = \sup_{\vec{r}|\sum_N r_i = r} \max\left\{\sum_{i=1}^{N} F_i^-(r_i) - (N-1), 0\right\}$$

and

$$F_{max}(r) = \inf_{\vec{r}|\sum_N r_i = r} \min\left\{\sum_{i=1}^{N} F_i(r_i), 1\right\},$$

with $\{F_i^-\}_{i=1}^N$ being the left limit of distribution $\{F_i\}_{i=1}^N$.

Here, (a) any distribution F for N is said to be of consistently varying tail (Cline 1994) if $\lim_{y\to 1} \lim_{x\to\infty} \sup \frac{\bar{F}(xy)}{\bar{F}(x)} = 1$, where $\bar{F} = 1 - F$, and (b) $\bar{F}' = 1 - F'$; $\bar{F}'$ is the distribution of individual $\{R_i\}_{i=1}^N$.

Proof Sketch - In the *Case A* setting, we have for arbitrary r and r_1 that

$$\mathbb{P}[A_N \leq r] \leq F_1(r_1) + F_2(r_2).$$

Hence, $\mathbb{P}[A_N \leq r] \leq F_{\max}(r)$. Now, $\mathbb{P}[R_1 < r_1] + \mathbb{P}[R_2 < r_2] - \mathbb{P}[R_1 < r_1, R_2 < r_2] \leq 1$. Therefore, the result:

$$\max\{F_1^{-1}(r_1) + F_2^-(r - r_1) - 1, 0\} \leq \mathbb{P}[R_1 < r_1, R_2 < r_2] \leq \mathbb{P}[A_N \leq r].$$

The analysis for the result on $F_{\min}$ follows similarly. In cases when $N \geq 3$, i.e., the *Case B* setting, we will use the principle of mathematical induction. $F_{\min}$ and $F_{\max}$ are non-increasing functions, and hence based on principles of mathematical analysis (Rudin et al. 1964), $F_{\min}(r)$ is left-continuous while $F_{\max}$ is right-continuous. Now choose an $\varepsilon > 0$, an u_0 such that $F_1(u_0 < \frac{\varepsilon}{2}$, and an r_0 such that $F_2(r_0 - u_0) < \frac{\varepsilon}{2}$. Hence, we have $F_{\max}(r) < \varepsilon$ for all $r \leq r_0$ when $r \to -\infty$. In situations when $r \to \infty$, choose an $\varepsilon > 0$, an u_1 such that $F_1(u_1 > 1 - \varepsilon$, and an r_1 such that $F_2(r_1 - u_1) > 1 - \varepsilon$. Now, given that

$$\inf_{u \leq u_1} \min\{F_1(u) + F_2(r_1 - u), 1\} \geq F_2(r_1 - u_1) > 1 - \varepsilon, \text{ and } \inf_{u > u_1} \min\{F_1(u) + F_2(r_1 - u), 1\} \geq F_1(u_1) > 1 - \varepsilon,$$

we have $F_{\max}(r) > 1 - \varepsilon$. The analysis for the result on $F_{\min}$ follows similarly. Applying the principle of mathematical induction, we can generalize the result for $N \geq 3$. This proves the theorem.

Cyber Insurance Industry Implications - This result implies that if cyber (re)insurers simply have knowledge about the individual enterprise cyber risk distributions, then *invariant of the degree of dependencies between these distributions*, a cyber (re-)insurer can estimate and bound the tail of aggregate cyber-risk only by an iterative computation procedure (as mentioned in (Frank et al. 1987)) that outputs the min/max of a set of a linear function performed on samples of aggregate cyber-risk. Alternatively, computing tail risk bounds of aggregate correlated cyber-risk is *easy*, i.e., not NP-hard.

6 SUMMARY

Systemic cyber risk management is steadily becoming a reality for the cyber insurance market. In this paper, we studied the role of (a) the statistical nature of multiple enterprise cyber risks contributing to aggregate/systemic supply chain risk and (b) the statistical nature of claim sizes in the underlying enterprise supply chain network, in the tail spread of aggregate cyber risk.

ACKNOWLEDGEMENT

This study has been supported by funding from Cybersecurity at MIT Sloan (CAMS).

REFERENCES

Beard, R. 2013. *Risk theory: the stochastic basis of insurance*, Volume 20. Springer Science & Business Media.

Bingham, N. H., C. M. Goldie, and J. L. Teugels. 1989. *Regular variation*. Number 27. Cambridge university press.

Cline, D. B. 1994. "Intermediate regular and Π variation". *Proceedings of the London Mathematical Society* 3(3):594–616.

Dambra, S., L. Bilge, and D. Balzarotti. 2020. "SoK: Cyber insurance–technical challenges and a system security roadmap". In *2020 IEEE Symposium on Security and Privacy (SP)*, 1367–1383. IEEE.

DasGupta, A. and A. DasGupta. 2008. "Saddlepoint Approximations". *Asymptotic Theory of Statistics and Probability*:203–224.

Denisov, D., S. Foss, and D. Korshunov. 2010. "Asymptotics of randomly stopped sums in the presence of heavy tails". *Bernoulli*:971–994.

Embrechts, P., C. Klüppelberg, and T. Mikosch. 2013. *Modelling extremal events: for insurance and finance*, Volume 33. Springer Science & Business Media.

Frank, M. J., R. B. Nelsen, and B. Schweizer. 1987. "Best-possible bounds for the distribution of a sum—a problem of Kolmogorov". *Probability theory and related fields* 74(2):199–211.

Hagerup, T. and C. Rüb. 1990. "A guided tour of Chernoff bounds". *Information processing letters* 33(6):305–308.

Hofmann, A. 2007. "Internalizing Externalities of Loss Prevention through Insurance Monopoly: An Analysis of Interdependent Risks". *The Geneva Risk and Insurance Review* 32:91–111.

Lelarge, M. and J. Bolot. 2009. "Economic Incentives to Increase Security in the Internet: The Case for Insurance". In *IEEE INFOCOM 2009*, 1494–1502. IEEE.

Liu, L. 2009. "Precise large deviations for dependent random variables with heavy tails". *Statistics & Probability Letters* 79(9):1290–1298.

Loukissas, F. 2012. "Precise large deviations for long-tailed distributions". *Journal of Theoretical Probability* 25:913–924.

Lu, D. 2012. "Lower bounds of large deviation for sums of long-tailed claims in a multi-risk model". *Statistics & Probability Letters* 82(7):1242–1250.

Marotta, A., F. Martinelli, S. Nanni, A. Orlando and A. Yautsiukhin. 2017. "Cyber-insurance survey". *Computer Science Review* 24:35–61.

Naghizadeh, P. and M. Liu. 2014, June 23-24. "Voluntary Participation in Cyber-Insurance Markets". In *Workshop on the Economics of Information Security (WEIS)*, 23–24. State College, PA, USA.

Ng, K. W., Q. Tang, J.-A. Yan, and H. Yang. 2004. "Precise large deviations for sums of random variables with consistently varying tails". *Journal of Applied Probability* 41(1):93–107.

Pal, R., K. Duan, and R. Sequeira. 2025. "A Theory to Estimate, Bound, and Manage Systemic Cyber Risk". In *To Appear in Proceedings of ACM SIGPADS*.

Pal, R., K. Duan, R. Sequeira, and M. Siegel. 2024. "Is Systemic Cyber Risk Management for Enterprises Sustainable?". In *2024 Winter Simulation Conference (WSC)*.

Pal, R. and L. Golubchik. 2010. "Analyzing Self-Defense Investments in Internet Security under Cyber-Insurance Coverage". In *IEEE 30th International Conference on Distributed Computing Systems*, 339–347. Genoa, Italy: IEEE.

Pal, R., L. Golubchik, and K. Psounis. 2011. "Aegis A Novel Cyber-Insurance Model". In *Decision and Game Theory for Security*, edited by J. S. Baras, J. Katz, and E. Altman, 131–150. Berlin, Heidelberg: Springer Berlin Heidelberg.

Pal, R., L. Golubchik, K. Psounis, and P. Hui. 2014, April 27-May 2. "Will Cyber-Insurance Improve Network Security? A Market Analysis". In *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, 235–243. Toronto, Canada.

Pal, R., L. Golubchik, K. Psounis, and P. Hui. 2017. "Security pricing as enabler of cyber-insurance a first look at differentiated pricing markets". *IEEE Transactions on Dependable and Secure Computing*.

Pal, R., L. Golubchik, K. Psounis, and P. Hui. 2018. "Improving Cyber-Security via Profitable Insurance Markets". *ACM SIGMETRICS Performance Evaluation Review* 45(4):7–15.

Pal, R., Z. Huang, S. Lototsky, X. Yin, M. Liu, J. Crowcroft *et al*. 2021. "Will Catastrophic Cyber-Risk Aggregation Thrive in the IoT Age? A Cautionary Economics Tale for (Re-) Insurers and Likes". *ACM Transactions on Management Information Systems (TMIS)* 12(2):1–36.

Pal, R., Z. Huang, X. Yin, M. Liu, S. Lototsky and J. Crowcroft. 2020. "Sustainable Catastrophic Cyber-Risk Management in IoT Societies". In *2020 Winter Simulation Conference (WSC)*, 3105–3116 https://doi.org/10.1109/WSC48552.2020.9384051.

Pal, R., Z. Huang, X. Yin, S. Lototsky, S. De, S. Tarkoma *et al*. 2020. "Aggregate Cyber-Risk Management in the IoT Age: Cautionary Statistics for (Re) Insurers and Likes". *IEEE Internet of Things Journal* 8(9):7360–7371.

Pal, R., P. Liu, T. Lu, and E. Hua. 2023. "How Hard is Cyber-Risk Management in IT/OT Systems? A Theory to Classify and Conquer Hardness of Insuring ICSs". *ACM Transactions on Cyber-Physical Systems (TCPS)* 6(4):1–31.

Pal, R., P. Liu, T. Lu, and X. Yin. 2021. "Cyber Re-Insurance Policy Writing is NP-Hard in IoT Societies". In *2021 Winter Simulation Conference (WSC)*, 1–12 https://doi.org/10.1109/WSC52266.2021.9715524.

Pal, R., T. Lu, P. Liu, and X. Yin. 2021. "Cyber (re-) insurance policy writing is NP-hard in IoT societies". In *2021 Winter Simulation Conference (WSC)*, 1–12. IEEE.

Pal, R., S. Madnick, and M. Siegel. 2023. "Trading in Catastrophe Bonds Can Boost Security Improving Cyber (Re-)Insurance Markets.". In *In Proceedings of Americas Conference on Information Systems (AMCIS)*.

Pal, R. and B. Nag. 2023. "A Mathematical Theory to Price Cyber-CAT Bonds to Boost IT/OT Security". In *2023 Winter Simulation Conference (WSC)*.

Pal, R., K. Psounis, J. Crowcroft, P. Hui, S. Tarkoma, A. Kumar *et al*. 2020. "When are Cyber Blackouts in Modern Service Networks Likely? A Network Oblivious Theory on Cyber (Re) Insurance Feasibility". *ACM Transactions on Management Information Systems (TMIS)* 11(2):1–38.

Pal, R., R. Sequeira, and S. Zeijlemaker. 2024. "How Hard is it to Estimate Systemic Enterprise Cyber-Risk?". In *2024 Winter Simulation Conference (WSC)*.

Rudin, W. *et al*. 1964. *Principles of mathematical analysis*, Volume 3. McGraw-hill New York.

Shetty, N., G. Schwartz, M. Felegyhazi, and J. Walrand. 2010. "Competitive Cyber-Insurance and Internet Security". In *Economics of Information Security and Privacy*, edited by T. Moore, D. Pym, and C. Ioannidis, 229–247. Boston, MA: Springer US.

Stam, A. 1973. "Regular variation of the tail of a subordinated probability distribution". *Advances in Applied Probability* 5(2):308–327.

Wang, S. and W. Wang. 2007. "Precise large deviations for sums of random variables with consistently varying tails in multi-risk models". *Journal of Applied Probability* 44(4):889–900.

Wolff, J. 2022. *Cyberinsurance policy: Rethinking risk in an Age of ransomware, computer fraud, data breaches, and cyberattacks*. MIT Press.

Woods, D. W. and R. Böhme. 2021. "SoK: Quantifying cyber risk". In *2021 IEEE Symposium on Security and Privacy (SP)*, 211–228. IEEE.

Yang, Z. and J. C. S. Lui. 2014. "Security Adoption and Influence of Cyber-Insurance Markets in Heterogenous Networks". *Performance Evaluation* 74.

AUTHOR BIOGRAPHIES

RANJAN PAL is a Research Scientist with the MIT Sloan School of Management, and an invited working group member of the World Economic Forum. His primary research interests lie in cyber risk and resilience management using interdisciplinary methods. He serves as an Associate Editor of the ACM Transactions on MIS. His email address is ranjanp@mit.edu.

KONNIE DUAN is a student in the Electrical Engineering and Computer Science (EECS) department at MIT. She is also a researcher with Cybersecurity at MIT Sloan (CAMS) at the MIT Sloan School of Management. Her primary research interest lies in cyber and financial risk management for business enterprise ecosystems. Her email address is konnie@mit.edu.

SANDER ZEIJLEMAKER is a Research Affiliate with the MIT Sloan School of Management. His primary research interest lies in developing cyber risk governance solutions based upon system dynamics. He is the President of the Security, Stability, and Resilience Special Interest Group of the System Dynamics Society. His email is szeijl@mit.edu.

MICHAEL SIEGEL is a Principal Research Scientist with the MIT Sloan School of Management. His primary research interest lies in cybersecurity management of information systems. He is the founding co-Director of the Cybersecurity at MIT Sloan (CAMS) center within the MIT Sloan School of Management. His email is msiegel@mit.edu.