SOCIETY OF ACTUARIES®

Modeling and Pricing Cybersecurity Risks in Fog Computing Based IoT Architectures

May 2021

Modeling and Pricing Cybersecurity Risks in Fog Computing Based IoT Architectures

AUTHOR Xiaoyu Zhang
Maochao Xu, PhD
Jianxi Su, PhD, FSA

SPONSOR General Insurance Research Committee

Caveat and Disclaimer

The opinions expressed and conclusions reached by the authors are their own and do not represent any official position or opinion of the Society of Actuaries or its members. The Society of Actuaries makes no representation or warranty to the accuracy of the information.

CONTENTS

Modeling and Pricing Cybersecurity Risks in Fog Computing Based IoT Architectures

Executive Summary

Research on cybersecurity risk modeling and pricing is becoming a spotlight in actuarial science. This paper pertains to the analysis of the cybersecurity risk inherent in the fog computing technology, which has been intensively deployed in assorted Internet of Things (IoT) applications. To this end, a structural model is established in order to describe the risk propagation mechanism in a fog network. We propose an interval approximation method to quantify the compromise frequencies for the network's elements, and under a smart home application, the compromise probabilities are computed explicitly. Applications of proposed models in the context of cyber insurance pricing are thoroughly explored.

Section 1: Introduction

Cybersecurity has been a ubiquitous matter in the present digital society, garnering extensive media coverage over the recent years. According to Cybersecurity Ventures[1] , financial damages due to cyber-related incidents are predicted to comprise six trillion US dollar globally in 2021. Such a magnitude of loss is comparable to about 7% of the world's GDP in 2019. Much effort has been made on the network infrastructures so as to enhance cybersecurity, vulnerabilities however cannot be fully eliminated in actual practice. To manage the residual cybersecurity risks, organizations including network service providers and users, are often advised to seek insurance solutions which secure a robust financial protection in case of cyber events (Böhme, 2005; Biener et al., 2015). In the wake of the market demand, an increasing number of insurance companies are driven to advance the cyber insurance products. Based on the latest figures published by the National Associate of Insurance Commissioners, the US alone has approximately 500 cyber insurance providers to date, with direct written premium amounted to two billion dollars (NAIC, 2019).

Quantitative study of cybersecurity risk is important not only for insurance company to properly price the cyber-related products, but also for customers to understand their needs for insurance coverage. In the context of cybersecurity risk modeling, two overarching strands of research stand out in the actuarial domain. The first strand of literature study cybersecurity risks from the macro-level perspective, aiming to understand the statistical properties of cyber-related losses and the associated economic implications. To name a few examples, Eling and Loperfido (2017) analyzed the loss distributions of different types of data breaches and found that the log-transformed loss distributions are right-skewed. Wheatley et al. (2016); Eling and Wirfs (2019) modeled the cyber loss frequency and severity by resorting to the toolkits originated from Extreme Value Theory. Eling and Jung (2018); Peng et al. (2018) applied copula methods to study the dependencies among different types of cyber losses. McShane and Nguyen (2020) conducted an empirical examination to study investor reactions to cyber events over time. Sun et al. (2020) developed a frequency-severity model by focusing on malicious hacking data breaches at the individual enterprise level, and further studied the application of the proposed model for ratemaking and pricing. Fang et al. (2021) studied the enterprise-level data breach risk, where a mixed D-vine dependence structure was employed to accommodate the complex dependence exhibited by the enterprise-level breach incident time series.

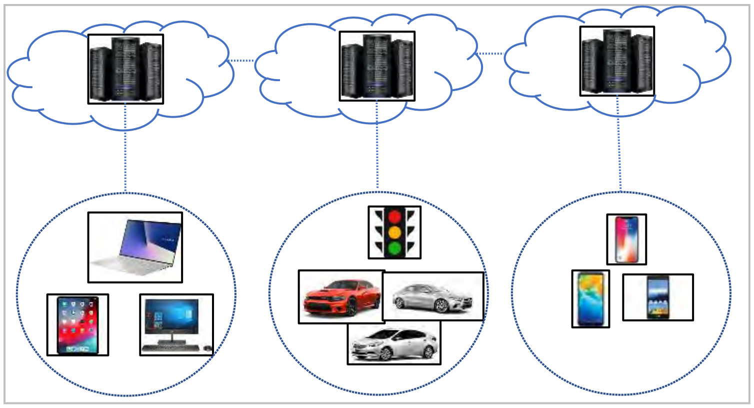
In the other strand of literature, cybersecurity risks are studied from the micro-level perspective, and focuses are placed on identifying the mechanisms that determine the extent of cybersecurity risks. To this end, structural models are frequently adopted to investigate the causal relationship between network characteristics and cyber losses. For instance, Fahrenwaldt et al. (2018); Xu et al. (2015); Xu and Hua (2019) deployed the susceptible-infectious-susceptible epidemic models over a deterministic network structure to study cyber losses. Jevtić and Lanchier (2020) proposed a class of dynamical percolation models for catering the cybersecurity risks within a tree-based network topology. Our paper falls into this strand of literature, which is more closely related to cyber insurance pricing.

It is fair to state that all the existing micro-level cybersecurity risk models in the actuarial domain only consider a relatively simple centralized network structure which may not be sufficient for the practical needs in this current era of Internet of Things (IoT). Although there is no universal definition on IoT (see, Lynn et al., 2020, for a variety of descriptions from either the technical or socio-technical perspectives), it may be heuristically understood as a network infrastructure of interconnected devices (i.e., things) for processing information from the physical and the virtual world. As the proliferation and consumerization of IoT

[1] Cybersecurity Ventures is the world's leading researcher and publisher for the global cyber economy, and it is widely accepted as a trusted source for cybersecurity facts, figures and statistics.

technology continue to evolve at a rampant pace, growing number of electronic devices are interconnected through multiple intricate networks, generating enormous data which need to be processed in real-time. In a traditional network system, the centralized cloud unit has a very high processing power and large memory storage so that low processing devices can run their respective computing in the cloud. Despite the broad utilization of cloud computing, due to the "long distance" between cloud-servers and end-users, a set of technical issues such as network congestion, high latency and cost, and scalability, etc., unfavorably arise. A new paradigm, namely fog computing, has been developed to circumvent the aforementioned technical limitations inherent in cloud computing (Puliafito et al., 2019). Specifically, fog computing is a decentralized computing infrastructure that extends the cloud service to the edge of the network, hence computational resources are closer to the position where the data is generated and used upon. Compared with the traditional cloud computing network, major advantages for adopting fog computing are the underlying superior user-experience and failure tolerance. Consequently, the fog computing has been widely deployed in a variety of domains which include smart home (Puliafito et al., 2019), health data management (Kraemer et al., 2017), intelligent transportation system (Darwish and Bakar, 2018), public services such as power grid, military defense and critical national infrastructure (Baccarelli et al., 2017). For the sake of illustration, a typical fog network is depicted in Figure 1, in which the multi-tenant (e.g., computers, laptops, smart devices, automated cars, traffic lights) and resource-sharing (e.g., the connected fog nodes) features are clear to recognize.

Figure 1
ILLUSTRATION OF A FOG NETWORK AND THE DIFFERENT TYPES OF END DEVICES



Although fog computing is emerging as a scalable, reliable and cost effective solution for big data analytics in IoT applications, its multi-tenant and resource-sharing architectures induce an unprecedented degree of cybersecurity risks to both the IoT service providers and users (Khan et al., 2017). The Ponemon Institute[2] estimated that the percentage of organizations who reported data breaches due to the unsecured IoT devices/applications has climbed from 15 percent in 2017 to 26 percent in 2019. In reality, the actual percentage may be even much higher since most organizations are not aware of the insecure IoT devices/applications threats in their work environment. The figures underscore the acute needs for IoT risk management improvement in which cyber insurance, as a nature tool for risk transfer and mitigation, should play a pivotal role.

[2] Source: https://www.ponemon.org/

Regardless of the increasing public concern about the security and privacy matters underlying IoT, quantitative methods for modeling the inherent cybersecurity risks seem to have gathered little attention thus far. Earlier studies related to fog networks are from the IT perspective, focusing on the construction and deployment of the technology in IoT. The only actuarially related work that we are aware of is Feng et al. (2018), where the risk management process of fog networks was formulated in a game theoretic framework. However, their work did not address the important issue of cybersecurity risk pricing. In this current paper, we follow a different route and aim to put forth a structural framework for modeling and pricing the cybersecurity risks in fog network from the micro-level perspective.

In this paper, we propose a novel class of propagation models to study the cybersecurity risks in fog computing cased IoT architectures, which are significantly different from the quantitative frameworks used in the existing cyber-related works within the actuarial literature (e.g., Fahrenwaldt et al., 2018; Jevtić and Lanchier, 2020; Xu et al., 2015; Xu and Hua, 2019). The rest of the paper is organized as follows. Beginning with a non-technical discussion about a few unique characteristics of cybersecurity risks in fog networks in Section 2, we propose a quantitative framework for modeling the cybersecurity risks in Section 3. To exemplify the proposed network models, in Section 4, we place the emphasis on the study of cybersecurity risks in a smart home system which is arguably one of the most popular IoT applications these days. Actuarial pricing of the cybersecurity risks in fog networks is considered in Section 5 with numerical illustrations. Sections 6 and 7 some further discussions and conclusions of the paper. In order to facilitate the reading, Appendix B contains a summary of the notation system used throughout the paper, and Appendix A contains the technical details which are in addition to our practical contributions.

Section 2: Characterizing the Cybersecurity Risks in Fog Networks

The cybersecurity risks associated with fog networks comprise a set of salient characteristics which must be addressed carefully. Firstly, fog networks feature a high level of heterogeneity and interdependency. To be specific, a fog network can consist of ample heterogeneous nodes which perform different functions such as controlling, networking, computing, and storing. These fog nodes can communicate with each other through wireless or wired transmission so that the computing resources can be shared. The aforementioned multi-tenant and resource-sharing natures of fog networks make the cybersecurity risk management very challenging. What is more, in the traditional centralized networks, patches and upgrades can be installed on the operating systems so as to limit the vulnerabilities existing in the network. However, the situation is quite different in fog networks due to the lightweight of operating systems and the relatively low computational capabilities of the IoT devices (Yu et al., 2015). The current security protocols of fog computing authenticate each edge device with the application before providing data or computation to perform. Hence, vulnerabilities hidden in the IoT devices form attractive entry points for attackers to penetrate into the network.

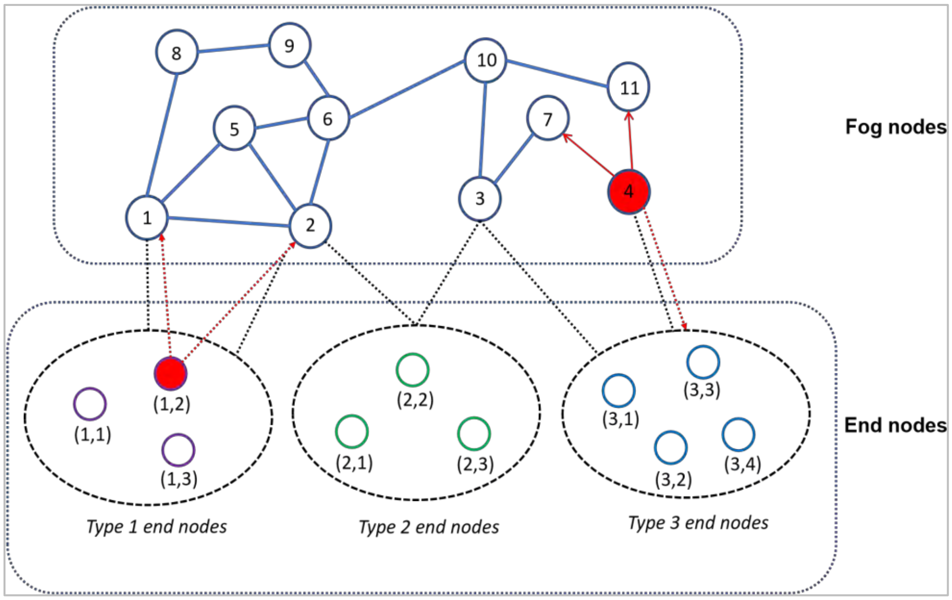
Secondly, fog networks are vulnerable to outside attacks. Typically, outside attacks are launched through unauthenticated devices or directly by external attackers. For example, an unauthenticated edge device can attack other devices or fog nodes, and an attacker can launch DDoS attacks directly to the fog nodes. Worse still, common vulnerabilities often exist in fog networks since similar computing nodes and end devices are operated under the same security configuration or software. These common vulnerabilities trigger the build-up of systemic cybersecurity risks. Namely, if a common vulnerability is identified and attacked by outsiders, then devastating damages may occur to the entire network. In cybersecurity risk pricing and management, it is critical to account for such high severity incidents.

Thirdly, fog networks are also vulnerable to inside attacks. The inside attacks are caused by the compromised authenticated devices which are inside the trusted network of applications. Once penetrated into the network, the attackers can advance toward the edge devices using the relations that exist among the vulnerabilities of different IoT devices. The compromised devices can attack fog nodes and other devices easily without being discovered as it has certain privileges in the network (Sohal et al., 2018).

For illustration, an abstract fog network is displayed in Figure 2, where there is one compromised fog node and one compromised end node, and the consequent cybersecurity risks may be propagated via the network structure. Specifically, fog node 4 is compromised by outside attacks, which can propagate the risk to its fog neighbors 7 and 11 via inside attacks. It can also propagate the risk to its end nodes, i.e., type 3 nodes. Similarly, the second type 1 end node is also compromised, and it can propagate the risk to its fog nodes 1 and 2. If fog node 2 is compromised, it can further compromise its neighbor fog nodes and end nodes, i.e., type 1 and type 2 end nodes. The formal modeling process is discussed in the following section.

Figure 2

ILLUSTRATION OF CYBERSECURITY RISKS FACED BY A FOG NETWORK IN THE IOT APPLICATION, WHERE THERE ARE 11 FOG NODES, AND 3 DISTINCT TYPES OF END NODES/DEVICES MARKED BY DIFFERENT COLORS. THE LABEL (d, i_d) INDICATES THE i_d-TH TYPE d END NODES, $d, i_d \in \mathbb{N}$. THE COMPROMISED NODES AND THE SURROUNDING PROPAGATION PATHS ARE INDICATED IN RED COLOR.



Section 3: Modeling the Cybersecurity Risks in Fog Networks

Modeling the occurrence of cyber-attack and the process of infection propagation plays an important role in assessing the cybersecurity risk within a fog network. To this end, we aim at putting forth a class of structural models for modeling the compromise frequency among the nodes of a fog network. In particular, the proposed infection models accommodate all the indispensable characteristics outlined in Section 2.

Let us begin with the notations for describing the inherent heterogeneous network components. For a fog network G, let $n^{\mathcal{F}}$ be the number of fog nodes, $n^{\mathcal{T}}$ be the number of end node types, and $n_d^{\mathcal{E}}$ be the number of end nodes that are of type $d \in \{1, \dots, n^{\mathcal{T}}\}$. Here and in the sequel, superscripts "$\mathcal{F}$", "$\mathcal{T}$" and "$\mathcal{E}$" indicate that an object of interest is related to the fog nodes, types of end node and end nodes, respectively. To illuminate, in the hypothetical fog network presented in Figure 2, we have

$$n^{\mathcal{F}} = 11, n^{\mathcal{T}} = 3, n_1^{\mathcal{E}} = n_2^{\mathcal{E}} = 3, n_3^{\mathcal{E}} = 4.$$

To quantify the frequency of cybersecurity risks in fog networks, special emphasis is placed on the set of compromise status RV's, $C_i^{\mathcal{F}} \in \{0,1\}$ and $C_{d,i_d}^{\mathcal{E}} \in \{0,1\}$, with value "$i$" and "$(d, i_d)$" indicate that the i-th fog node and the i_d-th type d end node is compromised, respectively, $i \in \{1, \dots, n^{\mathcal{F}}\}, d \in \{1, \dots, n^{\mathcal{T}}\}, i_d \in \{1, \dots, n_d^{\mathcal{E}}\}$. For brevity, we shorthand the compromise RV's by

$$\boldsymbol{C} = \left(\boldsymbol{C}^{\mathcal{F}}, \boldsymbol{C}_1^{\mathcal{E}}, \dots, \boldsymbol{C}_{n^{\mathcal{T}}}^{\mathcal{E}}\right), \quad \text{where } \boldsymbol{C}^{\mathcal{F}} = (C_1^{\mathcal{F}}, \dots, C_{n^{\mathcal{F}}}^{\mathcal{F}}), C_d^{\mathcal{E}} = (C_{d,1}^{\mathcal{E}}, \dots, C_{d,n_d^{\mathcal{E}}}^{\mathcal{E}}), d \in \{1, \dots, n^{\mathcal{T}}\}.$$

The study of $\boldsymbol{C}$ is further related to the frequency of outside and insider attacks which we are going to discuss next.

Denote by $O_i^{\mathcal{F}} \in \{0,1\}$ the outside attack status random variable (RV) of the i-th fog node, with $O_i^{\mathcal{F}} = 1$ means that the node is compromised due to an outside attack, and zero otherwise, $i = 1, \dots, n^{\mathcal{F}}$. Throughout, the i_d-th type d end node is labeled by (d, i_d), $d \in \{1, \dots, n^{\mathcal{T}}\}$ and $i_d \in \left\{1, \dots, n_d^{\mathcal{E}}\right\}$, for notational convenience. Then, $O_{d,i_d}^{\mathcal{E}} \in \{0,1\}$ denotes the outside attack status RV of the (d, i_d)-th end node. To account for the presence of systemic cybersercurity risk, we assume that there exist two types of vulnerabilities which may be exploited by outside attackers. Outside attacks through a common vulnerability imperil all the nodes that are of the same type, potentially causing systemic failures within a cohort of network components. In contrast, an idiosyncratic vulnerability may only exist in a particular node, through which outside attacks will infect the node solely. Let $V^{\mathcal{F}} \in \{0,1\}$ and $V_d^{\mathcal{T}} \in \{0,1\}$ indicate respectively whether a common vulnerability arises among the fog nodes and type d end nodes and is harnessed by an outsider to attack the network. Define

$$\mathbb{P}(V^{\mathcal{F}} = 1) =: v^{\mathcal{F}} \in [0,1], \quad \text{and} \quad \mathbb{P}\left(V_d^{\mathcal{T}} = 1\right) =: v_d^{\mathcal{T}} \in [0,1], \quad d = 1, \dots, n^{\mathcal{T}}.$$

Given that a common vulnerability is exploited by attackers, then with probabilities

$$\mathbb{P}(\{O_i^{\mathcal{F}} = 1, i = 1, \dots, n^{\mathcal{F}}\} | V^{\mathcal{F}} = 1) =: \pi^{\mathcal{F}*} \in [0,1]$$

and

$$\mathbb{P}\left(\left\{O_{d,i_d}^{\mathcal{E}} = 1, i = 1, \dots, n_d^{\mathcal{E}}\right\} \middle| V_d^{\mathcal{T}} = 1\right) =: \pi_d^{\mathcal{E}*} \in [0,1], \quad \text{for a fixed } d \in \{1, \dots, n^{\mathcal{T}}\},$$

all the fog nodes and all the type d end nodes will be compromised, respectively. In the probability notations above, the start sign "$*$" in the superscripts aims to emphasize that the compromise is caused by common vulnerability. Otherwise, individual devices may be attacked due to their own idiosyncratic vulnerabilities, and we have

$$\mathbb{P}\left(O_i^{\mathcal{F}} = 1 \middle| V^{\mathcal{F}} = 0\right) =: \pi_i^{\mathcal{F}} \in [0,1], \quad i = 1, \dots, n^{\mathcal{F}},$$

and

$$\mathbb{P}\left(O_{d,i_d}^{\mathcal{E}} = 1 \middle| V_d^{\mathcal{T}} = 0\right) =: \pi_{d,i_d}^{\mathcal{E}} \in [0,1], \quad i_d \in \left\{1, \dots, n_d^{\mathcal{E}}\right\}, d \in \{1, \dots, n^{\mathcal{T}}\}.$$

Denote by

$$\boldsymbol{O}^{\mathcal{F}} = (O_1^{\mathcal{F}}, \dots, O_{n^{\mathcal{F}}}^{\mathcal{F}}), \boldsymbol{O}_d^{\mathcal{E}} = \left(O_{d,1}^{\mathcal{E}}, \dots, O_{d,n_d^{\mathcal{E}}}^{\mathcal{E}}\right), \quad d = 1, \dots, n^{\mathcal{T}},$$

the sets of outside attack RV's. The following assumption is intuitive practically.

Assumption 3.1. *As outside attacks are launched randomly, we assume* $\boldsymbol{O}^{\mathcal{F}}, \boldsymbol{O}^{\mathcal{E}}_1, \dots, \boldsymbol{O}^{\mathcal{E}}_{n^{\mathcal{T}}}$ *to be mutually independent. Nevertheless, the outside attack RV's within the same node type are generally dependent because of the presence of common vulnerabilities. Specifically, given that* $V^{\mathcal{F}} = 1$ *(resp.,* $V^{\mathcal{T}}_d = 1, d \in \{1, \dots, n^{\mathcal{T}}\}$*, the coordinates of* $O^{\mathcal{F}}$ *(resp.,* $O^{\mathcal{E}}_d, d \in \{1, \dots, n^{\mathcal{T}}\}$ *are assumed to be identical almost surely. Practically, this assumption means that once a common vulnerability is exploited, all the nodes that are of the same type either get infected simultaneously if the attack succeeds, or all remain healthy if the attack fails. When* $V^{\mathcal{F}} = 0$ *(resp.,* $V^{\mathcal{T}}_d = 0, d \in \{1, \dots, n^{\mathcal{T}}\}$*, then the coordinates of* $O^{\mathcal{F}}$ *(resp.,* $O^{\mathcal{E}}_d, d \in \{1, \dots, n^{\mathcal{T}}\}$ *are assumed to be independent, since in this case, infections are caused by different attacks.*

We turn to consider the cybersecurity risks due to insider attacks which are launched by the existing compromised components through the connecting links. To this end, some additional notations for describing the possible paths of risk contagions are needed herein. For the i-th fog node, denote by $I^{\mathcal{F}}_{i\to j} \in \{0,1\}$ and $I^{\mathcal{F}}_{i\to(d,k_d)} \in \{0,1\}$ the activation status of the link to the j-th fog node and the (d, k_d)-th end node, respectively, with value "1" means active, "0" means inactive, and

$$\mathbb{P}\big(I^{\mathcal{F}}_{i\to j} = 1 \big| C^{\mathcal{F}}_i = 1\big) =: q^{\mathcal{F}}_{i\to j} \in [0,1], \quad \mathbb{P}(I^{\mathcal{F}}_{i\to(d,k_d)} = 1 | C^{\mathcal{F}}_i = 1) =: q^{\mathcal{F}}_{i\to(d,k_d)} \in [0,1],$$

for $d \in \{1, \dots, n^{\mathcal{T}}\}, i \neq j \in \{1, \dots, n^{\mathcal{F}}\}, k_d \in \{1, \dots, n^{\mathcal{E}}_d\}$. If a link is active, then a compromised node can lunch an insider attack to a healthy node via the link. Concerning the end nodes, note that in the security configuration of IoT applications, it is a common practice to limit the direct communications between end nodes so as to control the cybersecurity risk propagation. Thereby, we should only consider the direct communication from end nodes to fog nodes, and let $I^{\mathcal{E}}_{(d,i_d)\to j}$ represent the activation status of the link from the (d, i_d)-th end node to the j-th fog node, with

$$\mathbb{P}(I^{\mathcal{E}}_{(d,i_d)\to j} = 1 | C^{\mathcal{E}}_{d,i_d} = 1) =: q^{\mathcal{E}}_{(d,i_d)\to j} \in [0,1],$$

for $d \in \{1, \dots, n^{\mathcal{T}}\}, i_d \in \{1, \dots, n^{\mathcal{E}}_d\}, j \in \{1, \dots, n^{\mathcal{F}}\}$. For any two nodes between which there is no direct link, then the corresponding link status RV is equal to 0 with probability 1. To illustrate, consider the compromised fog node in the hypothetical network displayed in Figure 2, we have

$$q^{\mathcal{F}}_{4\to j} = \begin{cases} > 0, & j = 7, 11 \\ = 0, & \text{otherwise} \end{cases}, \quad q^{\mathcal{F}}_{4\to(d,i_d)} = \begin{cases} > 0, & (d, i_d) = (3,1), (3,2), (3,3), (3,4) \\ = 0, & \text{otherwise} \end{cases}.$$

For the compromised end node in the same network, it does not have any direct link to another end node but may have active links to fog nodes with transmission probabilities

$$q^{\mathcal{F}}_{(1.2)\to j} = \begin{cases} > 0, & j = 1, 2 \\ = 0, & \text{otherwise} \end{cases}.$$

Assumption 3.2. *Denote the set of all outside attack RV's by* $\boldsymbol{O}$ *and the set of all link status RV's associated with inside attacks by* $\boldsymbol{I}$*. Mainly, for mathematical elegance, we assume the coordinates of* $\boldsymbol{I}$ *to be mutually independent, meaning that a compromised node will attack its neighboring healthy nodes randomly and independently. Moreover, it is practical to assume that the outside attack RV's,* $\boldsymbol{O}$*, and inside attack RV's,* $\boldsymbol{I}$ *are independent.*

With the outside and inside attack notations in place, we now set out to establish a system of state equations for describing the compromise statuses of network nodes:

$$C_j^{\mathcal{F}} = 1 - \underbrace{\left(1 - O_j^{\mathcal{F}}\right)}_{①} \underbrace{\prod_{i=1, i\neq j}^{n^{\mathcal{F}}} \left(1 - C_i^{\mathcal{F}} I_{i\to j}^{\mathcal{F}}\right)}_{②} \underbrace{\prod_{d=1}^{n^{\mathcal{T}}} \prod_{i_d=1}^{n_d^{\mathcal{E}}} \left(1 - C_{d,i_d}^{\mathcal{E},[j]} I_{(d,i_d)\to j}^{\mathcal{E}}\right)}_{③}, \quad \text{for } j = 1, \dots, n^{\mathcal{F}}, \tag{1}$$

where

$$C_{d,i_d}^{\mathcal{E},[j]} = 1 - \underbrace{\left(1 - O_{d,i_d}^{\mathcal{E}}\right)}_{①} \underbrace{\prod_{i=1, i\neq j}^{n^{\mathcal{F}}} \left(1 - C_i^{\mathcal{F}} I_{i\to(d,i_d)}^{\mathcal{F}}\right)}_{②} \tag{2}$$

is state equation associated with the (d, i_d)-th end node while assuming that the j-th fog node is originally healthy (equivalently, excluding the j-th fog node from the state equations), and

$$C_{d,j_d}^{\mathcal{E}} = 1 - \underbrace{\left(1 - O_{d,j_d}^{\mathcal{E}}\right)}_{①} \underbrace{\prod_{i=1}^{n^{\mathcal{F}}} \left(1 - C_i^{\mathcal{F}} I_{i\to(d,i_d)}^{\mathcal{F}}\right)}_{②}, \quad \text{for } d = 1, \dots, n^{\mathcal{T}}, j_d = 1, \dots, n_d^{\mathcal{E}}. \tag{3}$$

Table 1 summaries the descriptions for the components in state equations (1) - (3). For a concise summary of the notation system introduced in this subsection, we refer the reader to Appendix B.

Table 1
DESCRIPTIONS OF THE COMPONENTS IN THE STATE EQUATIONS.

Number	Description
①	Compromise due to outside attacks.
②	Compromise due to inside attacks from infected fog nodes.
③	Compromise due to inside attacks from infected end nodes.

Remark that state equations (1) - (3) not only endogenize the stochastic compromise statuses of all nodes, but also capture the intricate risk contagions across the network. The set of compromise RV's are highly dependent. One origin of the dependence comes from the fact that the compromise RV specified in each state equation is interlinked with the compromise statuses of its neighboring nodes. Another origin is via the outside attack RV's involved in Equations (1) - (3), which are correlated among the same type of nodes because of the common vulnerabilities (also see, the discussion in Assumption 3.1). Consequently, it is considerably challenging to evaluate the joint compromise probabilities underlying $\boldsymbol{C}$. To the best of our knowledge, no explicit expression can be obtained for the distribution of $\boldsymbol{C}$. Numerical simulation must be adopted so as to tackle the problem, which may be computationally intensive. In some applications such as preliminary analysis, sensitivity testing, risk communication and so forth, an easy-to-implement approximation of the compromise probabilities is more likely to be appreciated by practitioners, which will be considered in the succeeding subsection.

3.1 INTERVAL APPROXIMATION OF COMPROMISE PROBABILITIES

We propose an interval method for approximating compromise probabilities:

$$\boldsymbol{p}^{\mathcal{F}} = (p_1^{\mathcal{F}}, \dots, p_{n^{\mathcal{F}}}^{\mathcal{F}})^\top, \quad \text{with } p_i^{\mathcal{F}} := \mathbb{P}(C_i^{\mathcal{F}} = 1), i = 1, \dots, n^{\mathcal{F}}, \tag{4}$$

and

$$\boldsymbol{p}_d^{\mathcal{E}} = (p_{d,1}^{\mathcal{E}}, \dots, p_{d,n_d^{\mathcal{E}}}^{\mathcal{E}})^\top, \quad \text{with } p_{d,i_d}^{\mathcal{E}} \coloneqq \mathbb{P}(C_{d,i_d}^{\mathcal{E}} = 1),\ d = 1, \dots, n^{\mathcal{T}},\ i_d = 1, \dots, n_d^{\mathcal{E}}. \tag{5}$$

Our main argument hinges on the notion of positive association in studying dependent RV's.

Definition 3.3. *A random vector* $\boldsymbol{X} = (X_1, X_2, \dots, X_n) \in \mathbb{R}^n, n \in \mathbb{N}$, *is said to be positively associated if*

$$\mathrm{Cov}(f(\boldsymbol{X}), g(\boldsymbol{X})) \geq 0$$

holds for all real-valued functions f, g *which are non-decreasing in each coordinate and such that the covariance exists.*

Recall that $\boldsymbol{C}, \boldsymbol{I}$ and $\boldsymbol{O}$ denote the sets of all RV's related to compromise statuses, inside attacks and outsider attacks, respectively. The next assertion shows that the aforementioned RV's are indeed positively associated. The succeeding lemma is of auxiliary importance.

Lemma 3.4 (Shaked,1982). *Assume that Borel measurable functions* $f_i: \mathbb{R}^n \to \mathbb{R}^m, i = 1, \dots, m$ *and* $m, n \in \mathbb{N}$, *are either all non-decreasing or all non-increasing component-wise. If* $\boldsymbol{X} \in \mathbb{R}^n$ *is positively associated, then* $(f_1(\boldsymbol{X}), \dots, f_m(\boldsymbol{X}))$ *is also positively associated.*

Proposition 3.5. *Under Assumptions 3.1 and 3.2, the compromise status, inside attack and outside attack RV's in a fog network, namely* $(\boldsymbol{C}, \boldsymbol{I}, \boldsymbol{O})$, *are positively associated.*

Proof. See, Appendix A. ■

Now, we are ready to spell out the interval approximation for compromise probability vectors $\boldsymbol{p}^{\mathcal{F}}$ and $\boldsymbol{p}_d^{\mathcal{E}}, d = 1, \dots, n^{\mathcal{T}}$. To facilitate the presentation, let us denote the outside attack probabilities by

$$\begin{aligned}\omega_j^{\mathcal{F}} &\coloneqq \mathbb{P}\big(O_j^{\mathcal{F}} = 1\big) = \mathbb{P}(V^{\mathcal{F}} = 0)\mathbb{P}\big(O_j^{\mathcal{F}} = 1 \big| V^{\mathcal{F}} = 0\big) + \mathbb{P}(V^{\mathcal{F}} = 1)\mathbb{P}\big(O_j^{\mathcal{F}} = 1 \big| V^{\mathcal{F}} = 1\big) \\ &= \pi_j^{\mathcal{F}} + v^{\mathcal{F}}\big(\pi^{\mathcal{F}*} - \pi_j^{\mathcal{F}}\big), \quad j = 1, \dots, n^{\mathcal{F}},\end{aligned}$$

and similarly

$$\omega_{d,j_d}^{\mathcal{E}} \coloneqq \mathbb{P}(O_{d,j_d}^{\mathcal{E}} = 1) = \pi_{d,j_d}^{\mathcal{E}} + v_d^{\mathcal{T}}(\pi_d^{\mathcal{E}*} - \pi_{d,j_d}^{\mathcal{E}}), d = 1, \dots, n^{\mathcal{T}}, j_d = 1, \dots, n_d^{\mathcal{E}}.$$

Let $\boldsymbol{l}^{\mathcal{F}} = (l_1^{\mathcal{F}}, \dots, l_{n^{\mathcal{F}}}^{\mathcal{F}})^\top$ with elements

$$l_j^{\mathcal{F}} = \max\left(\underbrace{\omega_j^{\mathcal{F}}}_{\textcircled{1}}, \underbrace{\bigvee_{i=1, i\neq j}^{n^{\mathcal{F}}} \beta_i q_{i\to j}^{\mathcal{F}}}_{\textcircled{2}}, \underbrace{\bigvee_{d=1}^{n^{\mathcal{T}}} \bigvee_{i_d=1}^{n_d^{\mathcal{E}}} \max\left(\omega_{d,i_d}^{\mathcal{E}}, \bigvee_{i=1, i\neq j}^{n^{\mathcal{F}}} \beta_i q_{i\to(d,i_d)}^{\mathcal{F}}\right) q_{(d,i_d)\to j}^{\mathcal{E}}}_{\textcircled{3}} \right), \tag{6}$$

where

$$\beta_j = \max\left(\omega_j^{\mathcal{F}}, \bigvee_{i=1, i\neq j}^{n^{\mathcal{F}}} \omega_i^{\mathcal{F}} q_{i\to j}^{\mathcal{F}}, \bigvee_{d=1}^{n^{\mathcal{T}}} \bigvee_{i_d=1}^{n_d^{\mathcal{E}}} \max\left(\omega_{d,i_d}^{\mathcal{E}}, \bigvee_{i=1, i\neq j}^{n^{\mathcal{F}}} \omega_i^{\mathcal{F}} q_{i\to(d,i_d)}^{\mathcal{F}}\right) q_{(d,i_d)\to j}^{\mathcal{E}} \right)$$

for $j = 1, \dots, n^{\mathcal{F}}$. Moreover, let

$$\boldsymbol{u}^{\mathcal{F}} = (u_1^{\mathcal{F}}, \dots, u_{n^{\mathcal{F}}}^{\mathcal{F}})^\top = (\mathbf{1} - \boldsymbol{A})^{-1}(1 - \boldsymbol{\gamma}), \tag{7}$$

in which $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_{n^{\mathcal{F}}})^\top$ with

$$\gamma_j := (1 - \omega_j^{\mathcal{F}}) \prod_{d=1}^{n^{\mathcal{T}}} \prod_{i_d=1}^{n_d^{\mathcal{E}}} \left[1 - q_{(d,i_d)\to j}^{\mathcal{E}} + q_{(d,i_d)\to j}^{\mathcal{E}} (1 - \omega_{d,i_d}^{\mathcal{E}}) \prod_{i=1, i\neq j}^{n^{\mathcal{F}}} \left(1 - q_{i\to(d,i_d)}^{\mathcal{F}}\right) \right],$$

and $\boldsymbol{A}$ is an $n^{\mathcal{F}}$ by $n^{\mathcal{F}}$ zero diagonal matrix having off-diagonal elements $a_{ij} = \gamma_i q_{j\to i}^{\mathcal{F}}$ for $i \neq j \in \{1, \dots, n^{\mathcal{F}}\}$. Lastly, for $d = 1, \dots, n^{\mathcal{T}}, j_d = 1, \dots, n_d^{\mathcal{E}}$, define $\boldsymbol{l}_d^{\mathcal{E}} = (l_{d,1}^{\mathcal{E}}, \dots, l_{d,n_d^{\mathcal{E}}}^{\mathcal{E}})^\top$with elements

$$l_{d,j_d}^{\mathcal{E}} = \max\left(\underbrace{\omega_{d,j_d}^{\mathcal{E}}}_{①}, \underbrace{\bigvee_{i=1}^{n^{\mathcal{F}}} l_i^{\mathcal{F}} q_{i\to(d,j_d)}^{\mathcal{F}}}_{②} \right), \tag{8}$$

and $\boldsymbol{u}_d^{\mathcal{E}} = (u_{d,1}^{\mathcal{E}}, \dots, u_{d,n_d^{\mathcal{E}}}^{\mathcal{E}})^\top$with elements

$$u_{d,j_d}^{\mathcal{E}} = 1 - \left(1 - \omega_{d,j_d}^{\mathcal{E}}\right) \prod_{i=1}^{n^{\mathcal{F}}} \left(1 - u_i^{\mathcal{F}} q_{i\to(d,j_d)}^{\mathcal{F}}\right), \tag{9}$$

where $l_j^{\mathcal{F}}$ and $u_j^{\mathcal{F}}, j = 1, \dots, n^{\mathcal{F}}$, are specified in Equations (6) and (7), respectively.

Theorem 3.6. *Consider a fog network described as per Section 3, and suppose that Assumptions 3.1 and 3.2 hold. The corresponding compromise probability vectors satisfy the following inequalities*

$$\boldsymbol{l}^{\mathcal{F}} \leq \boldsymbol{p}^{\mathcal{F}} \leq \boldsymbol{u}^{\mathcal{F}}$$

and

$$\boldsymbol{l}_d^{\mathcal{E}} \leq \boldsymbol{p}_d^{\mathcal{E}} \leq \boldsymbol{u}_d^{\mathcal{E}}, \quad \text{for } d = 1, \dots, n^{\mathcal{T}}.$$

Proof. See, Appendix A. ■

Here are some remarks about Theorem 3.6. Firstly, the expressions in Equations (6) to (9) only contain simple algebraic operators, thus the bounds can be evaluated conveniently. Secondly, the lower bounds of compromise probabilities are derived by applying the co-monotonic approximation (Dhaene et al., 2002a, b) on the risk factors that determines the compromised probabilities, while in contrast, the upper bounds are based on the independence approximation. Thirdly, the compromise probabilities' lower bounds possess an intuitive interpretation. Namely, if a given node gets infected, then the infection must be caused by either an outside attack or an inside attack launched from another compromised node. Thereby, the compromise probabilities must be bounded below by the maximum of the infection probabilities due to one of these causes. We again refer to Table 1 for the description of components contained in Equations (6) and (8).

The succeeding hypothetical, but not so unrealistic, example demonstrates the usefulness of the interval approximation in Theorem 3.6 in an illuminated manner. It is our intention to keep the example's set-up simple for ease of exposition.

Example 3.7. *Consider a fog network as per Figure 3, in which there are four fog nodes and six end devices. Two of the end devices are of type* 1 *end nodes, and the others are of type* 2 *end nodes. Further, for* $i \in \{1, \dots, 4\}, i_1 = \{1,2\}, i_2 \in \{1, \dots, 4\}$*, assume the idiosyncratic and systemic outside attack probabilities to be* $\pi_i^{\mathcal{F}} = 0.01$ *and* $\pi^{\mathcal{F}*} = 0.05$*, respectively. Regarding the end nodes, because the inherent cybersecurity* *configuration is typically weaker, outside attacks are more likely to success and we set* $\pi^{\mathcal{E}}_{1,i_1} = 0.2$ *and* $\pi^{\mathcal{E}*}_1 = 0.25$ *for the type 1 end nodes, and* $\pi^{\mathcal{E}}_{2,i_1} = 0.2$ *and* $\pi^{\mathcal{E}*}_2 = 0.3$ *for the type 2 end nodes. The inside attack probabilities,* $q^{\mathcal{F}}_{i\to j}, q^{\mathcal{F}}_{i\to(d,i_d)}, q^{\mathcal{E}}_{(d,i_d)\to j}$ *are assumed to be identical and equal to* $q \in (0,1)$*. Similarly, we set the common vulnerability likelihoods* $\nu^{\mathcal{F}} = \nu^{\mathcal{T}}_1 = \nu^{\mathcal{T}}_2 = \nu$*. Table 2 depicts the interval estimates of compromised probabilities against the true compromised probabilities based on numerical simulation.*

Figure 3

GROUPED COLUMN CHART SAMPLE THE FOG NETWORK OF EXAMPLE 3.7.

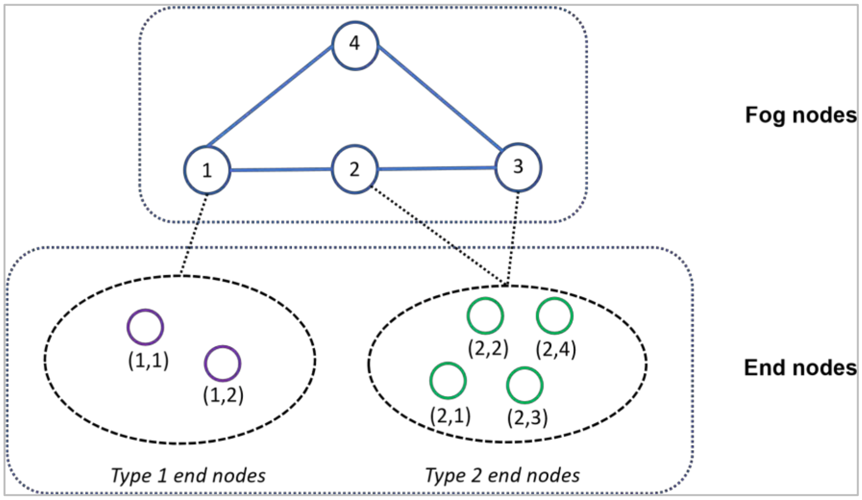


Table 2

THE INTERVAL APPROXIMATIONS AND SIMULATION-BASED CALCULATIONS OF COMPROMISE PROBABILITIES FOR THE FOG NETWORK IN EXAMPLE 3.7.

		v = 0.5			*v* = 0.6			*v* = 0.7		
		Lower	Sim.	Upper	Lower	Sim.	Upper	Lower	Sim.	Upper
q=0.1	$p^{\mathcal{F}}_1$	0.030	0.079	0.094	0.034	0.089	0.100	0.038	0.093	0.106
	$p^{\mathcal{F}}_2$	0.030	0.124	0.173	0.034	0.138	0.180	0.038	0.144	0.188
	$p^{\mathcal{F}}_3$	0.030	0.117	0.169	0.034	0.136	0.177	0.038	0.139	0.185
	$p^{\mathcal{F}}_4$	0.030	0.041	0.056	0.034	0.051	0.061	0.038	0.057	0.066
	$p^{\mathcal{E}}_{1,1}$	0.225	0.229	0.232	0.230	0.240	0.238	0.235	0.235	0.243
	$p^{\mathcal{E}}_{1,2}$	0.225	0.231	0.232	0.230	0.233	0.238	0.235	0.235	0.243
	$p^{\mathcal{E}}_{2,1}$	0.250	0.256	0.275	0.260	0.272	0.286	0.270	0.273	0.297
	$p^{\mathcal{E}}_{2,2}$	0.250	0.257	0.275	0.260	0.269	0.286	0.270	0.278	0.297
	$p^{\mathcal{E}}_{2,3}$	0.250	0.256	0.275	0.260	0.270	0.286	0.270	0.271	0.297
	$p^{\mathcal{E}}_{2,4}$	0.250	0.260	0.275	0.260	0.272	0.286	0.270	0.274	0.297
q=0.25	$p^{\mathcal{F}}_1$	0.056	0.190	0.295	0.058	0.195	0.303	0.059	0.201	0.311
	$p^{\mathcal{F}}_2$	0.063	0.268	0.511	0.065	0.275	0.520	0.068	0.277	0.528
	$p^{\mathcal{F}}_3$	0.063	0.265	0.502	0.065	0.269	0.510	0.068	0.262	0.519
	$p^{\mathcal{F}}_4$	0.030	0.120	0.223	0.034	0.130	0.230	0.038	0.126	0.238
	$p^{\mathcal{E}}_{1,1}$	0.225	0.247	0.282	0.230	0.256	0.288	0.235	0.258	0.294
	$p^{\mathcal{E}}_{1,2}$	0.225	0.251	0.282	0.230	0.259	0.288	0.235	0.250	0.294
	$p^{\mathcal{E}}_{2,1}$	0.250	0.299	0.428	0.260	0.306	0.438	0.270	0.307	0.449
	$p^{\mathcal{E}}_{2,2}$	0.250	0.298	0.428	0.260	0.309	0.438	0.270	0.307	0.449
	$p^{\mathcal{E}}_{2,3}$	0.250	0.299	0.428	0.260	0.309	0.438	0.270	0.302	0.449
	$p^{\mathcal{E}}_{2,4}$	0.250	0.298	0.428	0.260	0.307	0.438	0.270	0.300	0.449

		$v = 0.5$			$v = 0.6$			$v = 0.7$		
		Lower	Sim.	Upper	Lower	Sim.	Upper	Lower	Sim.	Upper
q=0.4	$p_1^{\mathcal{F}}$	0.090	0.318	0.673	0.092	0.317	0.679	0.094	0.316	0.684
	$p_2^{\mathcal{F}}$	0.100	0.396	0.860	0.104	0.390	0.864	0.108	0.383	0.867
	$p_3^{\mathcal{F}}$	0.100	0.388	0.854	0.104	0.381	0.858	0.108	0.374	0.861
	$p_4^{\mathcal{F}}$	0.040	0.246	0.623	0.042	0.244	0.628	0.043	0.242	0.633
	$p_{1,1}^{\mathcal{E}}$	0.225	0.295	0.434	0.230	0.297	0.439	0.235	0.298	0.444
	$p_{1,2}^{\mathcal{E}}$	0.225	0.295	0.434	0.230	0.297	0.439	0.235	0.298	0.444
	$p_{2,1}^{\mathcal{E}}$	0.250	0.369	0.676	0.260	0.365	0.682	0.270	0.362	0.687
	$p_{2,2}^{\mathcal{E}}$	0.250	0.368	0.676	0.260	0.365	0.682	0.270	0.362	0.687
	$p_{2,3}^{\mathcal{E}}$	0.250	0.369	0.676	0.260	0.365	0.682	0.270	0.362	0.687
	$p_{2,4}^{\mathcal{E}}$	0.250	0.369	0.676	0.260	0.365	0.682	0.270	0.362	0.687

Here are how the numerical results in Example 3.7 should be interpreted. Firstly, because the outside attack probabilities of end nodes are higher than that of fog nodes, the fog nodes have lower compromised probabilities than the end nodes. Due to a similar reasoning, the type one end nodes have lower compromised probabilities compared to the type two end nodes. Among the four fog nodes, it is natural to conjecture that $p_4^{\mathcal{F}} \overset{(1)}{\leq} p_1^{\mathcal{F}} \overset{(2)}{\leq} p_3^{\mathcal{F}} \overset{(3)}{\leq} p_2^{\mathcal{F}}$ where

- "$\overset{(1)}{\leq}$" holds since there is no end node directly connected to fog node 4;
- "$\overset{(2)}{\leq}$" holds since the number of end nodes directly connected to fog node 1 is smaller than that of fog nodes 2 and 3, of which the outside attack probabilities are also smaller (i.e., $\omega_{1,i_1}^{\mathcal{E}} < \omega_{2,i_2}^{\mathcal{E}}$);
- "$\overset{(3)}{\leq}$" holds since fog node 2 is closer to the type one end nodes compared with fog node 3.

It is noteworthy that both the lower and upper bounds of the interval approximations are capable of reflecting the aforementioned orders.

Secondly, we change the inside attack probabilities among $q \in \{0.1, 0.25, 0.4\}$. As shown, with all else being equal, if the fog network has a high security configuration and so the internal risk propagation probabilities are low, then the compromise probabilities are also low. In this case, since the cybersecurity risk is mainly caused by outside attacks which are well captured by the lower and upper bound formulas, the proposed interval method provides a very good estimate of the true compromise probability. However, as the inside attack probabilities increase, the effect of network dependence becomes more significant. The true network dependence is hard to be captured by the independent or co-monotonic approximation, thus the performance of the interval method decays.

Thirdly, vary the probabilities of common vulnerabilities among $v \in \{0.5, 0.6, 0.7\}$, we observe that the final compromise probabilities may get lower or higher as v increases. This is caused by the intricate interplay between the inside attacks and outside attacks in the determination of the compromise probabilities. Namely, the calculation of compromise probability can be viewed as an application of Bayes rule of two conditional compromise probabilities given whether or not a common vulnerability occurs. Depending on the order of the two conditional compromise probabilities, the increment of v may pose different directions of impacts to the unconditional compromise probabilities. The proposed interval approximation may not be capable of reflecting the direction of change in the compromise probabilities due to varying v. However, the approximation intervals are still able to capture the true compromise probabilities.

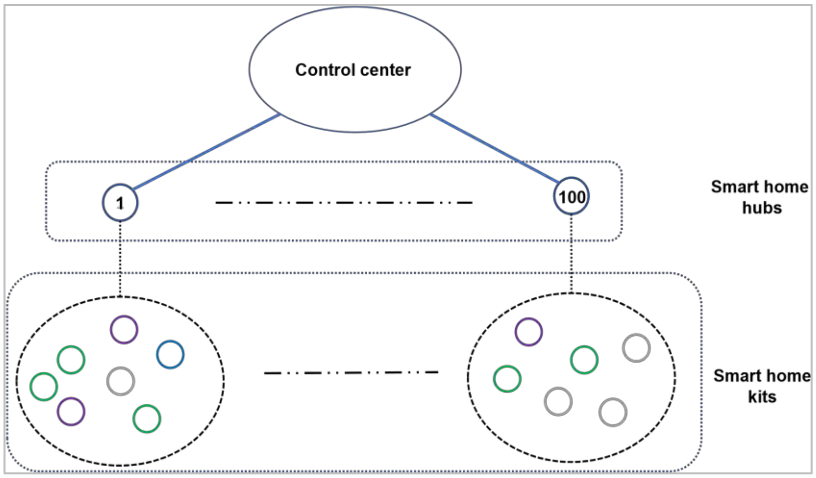
Section 4: Cybersecurity Risks in Smart Home Fog Networks

The infection models established in the previous section are rather abstract. In order to gain more insights into the proposed framework, the rest of this article is devoted to the application of the proposed infection models on a smart home system which corresponds to one of the most common fog computing based IoT architectures.

Consider the cyber infrastructure of a smart home provider, consisting of $n^{\mathcal{F}}$ individual users and $n^{\mathcal{T}}$ types of household kits (i.e., end devices). Typically, each user's smart home system is equipped with a hub or gateway (i.e., fog node), acting as a go-between for multiple smart devices and enabling automation. Further, the fog nodes among different users are interconnected through a control center maintained by the service provider. Figure 4 illustrates the network structure underling a smart home system. As shown, the hubs and household kits of the smart home system form a fog network, even though there is no direct communication between the fog nodes. Nevertheless, the compromise statuses of fog nodes may be still highly dependent due to the presences of common vulnerabilities as well as the mutual connections to the control center.

Figure 4

ILLUSTRATION OF A SMART HOME NETWORK WITH THE END DEVICES ARE GROUPED ACCORDING TO THE OWNERSHIP OF INDIVIDUAL USERS, AND DIFFERENT TYPES OF END DEVICES ARE DISPLAYED IN DIFFERENT COLORS.



The theoretical groundwork laid down in Section 3 can be utilized to model the cybersecurity risks in a smart home network. In the sequel, we will follow the same notations used in Section 3, and furthermore in order to capture the cybersecurity risks associated with the additional control center, let us introduce the following notations:

- the compromise status RV for the central control, $C^{\mathcal{C}} \in \{0,1\}$;
- the outside attack RV, $O^{\mathcal{C}} \in \{0,1\}$, with probability $\mathbb{P}(O^{\mathcal{C}} = 1) = \omega^{\mathcal{C}}$;
- the inside attack RV, $I^{\mathcal{F}}_{i \to \bullet} \in \{0,1\}$, indicates the internal infection launched from the i-th compromised fog node to the healthy central control, with $\mathbb{P}(I^{\mathcal{F}}_{i \to \bullet} = 1) = q^{\mathcal{F}}_{i \to \bullet}$, and $I^{\mathcal{C}}_{\bullet \to i} \in \{0,1\}$ indicates the internal infection launched from compromised central control to the i-th healthy fog node, with $\mathbb{P}(I^{\mathcal{C}}_{\bullet \to i} = 1) = q^{\mathcal{C}}_{\bullet \to i}$.

In the study of smart home, it is more convenient for us group the end devices based on the ownership of individual users. As demonstrated in Figure 4, each smart home device is directly connected to a single fog node, so for a specific end device, inside attack can be only launched from/to the particular connected fog node. For ease of exposition, fix $d = 1, \dots, n^{\mathcal{T}}$, $i = 1, \dots, n^{\mathcal{F}}$, we further introduce

$$\mathbb{D}_{d,i} = \{\, j_d \in \{1, \dots, n_d^{\mathcal{E}}\}: q^{\mathcal{F}}_{i\to(d,j_d)} > 0 \quad \text{or} \quad q^{\mathcal{E}}_{(d,j_d)\to i} > 0\}$$

to the denote the set of type d end devices possessed by the i-th smart home user, among which inside attacks may occur.

The state equation underlying the compromise status RV of the central control can be specified as

$$C^{\mathcal{C}} = 1 - (1 - O^{\mathcal{C}}) \prod_{j=1}^{n^{\mathcal{F}}} \left(1 - C_j^{\mathcal{F},[\bullet]} \times I^{\mathcal{F}}_{j\to\bullet}\right), \tag{10}$$

Where

$$C_j^{\mathcal{F},[\bullet]} = 1 - (1 - O_j^{\mathcal{F}}) \prod_{d=1}^{n^{\mathcal{T}}} \prod_{i_d \in \mathbb{D}_{d,j}} (\,1 - O^{\mathcal{E}}_{d,i_d} \times I^{\mathcal{E}}_{(d,i_d)\to j}), \quad j = 1, \dots, n^{\mathcal{F}}, \tag{11}$$

corresponds to the compromise status of the j-th fog node when the central control is excluded, or equivalently, assumed to be originally healthy. Moreover, state equations (1) - (3) can be adapted to capture the fog network structure of the smart home system. Namely, for the j-th fog nodes, $j = 1, \dots, n^{\mathcal{F}}$, we have

$$\begin{aligned} C_j^{\mathcal{F}} &= 1 - \left(1 - O_j^{\mathcal{F}}\right)\left(1 - C^{\mathcal{C},[j]} \times I^{\mathcal{C}}_{\bullet\to j}\right) \prod_{d=1}^{n^{\mathcal{T}}} \prod_{i_d \in \mathbb{D}_{d,j}} (\,1 - O^{\mathcal{E}}_{d,i_d} \times I^{\mathcal{E}}_{(d,i_d)\to j} \\ &= 1 - \left(1 - C^{\mathcal{C},[j]} \times I^{\mathcal{C}}_{\bullet\to j}\right)\left(1 - C_j^{\mathcal{F},[\bullet]}\right), \end{aligned} \tag{12}$$

where

$$C^{\mathcal{C},[j]} = 1 - (1 - O^{\mathcal{C}}) \prod_{i=1, i\neq j}^{n^{\mathcal{F}}} (1 - C_i^{\mathcal{F},[\bullet]} \times I^{\mathcal{F}}_{i\to\bullet}),$$

is the state equation associated with the central control but with the j-th fog node excluded from the system. The state equation for the (d, j_d)-th end device belonging to the i-th user can be specified as

$$C^{\mathcal{E}}_{d,j_d} = 1 - \left(1 - O^{\mathcal{E}}_{d,j_d}\right)\left(1 - C_i^{\mathcal{F}} I^{\mathcal{F}}_{i\to(d,j_d)}\right), \quad j_d \in \mathbb{D}_{d,i} \text{ with } d = 1, \dots, n^{\mathcal{T}}, i = 1, \dots, n^{\mathcal{F}}. \tag{13}$$

Thanks to the more specific network topology underlying the smart home platform, we manage to compute the compromised probabilities in explicit forms. At first, let us begin with a simpler situation in which the central control is highly secure, and thus the associated cybersecurity risk due to the control center can be excluded from the consideration.

Proposition 4.1. *Consider the smart home network as illustrated in Figure 4, and further, assume that the control center is highly secure with zero compromise probability, i.e.,* $p^{\mathcal{C}} := \mathbb{P}(C^{\mathcal{C}} = 1) = 0$*. For a given set of* m *fog nodes, indexed by* $\Xi = (\xi_1, \dots, \xi_m) \subseteq \{1, \dots, n^{\mathcal{F}}\}$*, their joint compromise probabilities can be computed via*

$$\tilde{p}_{\Xi}^{\mathcal{F}} := \mathbb{P}\left(\bigcap_{j\in\Xi} C_j^{\mathcal{F},[\bullet]} = 1\right) = 1 - \sum_{k=1}^{m}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi} h(\Xi_k)\,, \tag{14}$$

where $\Xi_k \in \mathbb{N}^k$ denotes any k-dimensional subset of Ξ, $k = 1, \dots, m$, and

$$h(\Xi_k) = \left[(1-\nu^{\mathcal{F}})\prod_{j\in\Xi_k}\left(1-\pi_j^{\mathcal{F}}\right) + \nu^{\mathcal{F}}(1-\pi^{\mathcal{F}*})\right]\prod_{d=1}^{n^{\mathcal{T}}} g(d,\Xi_k)$$

with

$$g(d,\Xi_k) = \left(1-\nu_d^{\mathcal{T}}\right)\prod_{j\in\Xi_k}\prod_{i_d\in\mathbb{D}_{d,j}}\left(1-\pi_{d,i_d}^{\mathcal{E}} q_{(d,i_d)\to j}^{\mathcal{E}}\right) + \nu_d^{\mathcal{T}}\left(1-\pi_d^{\mathcal{E}*} + \pi_d^{\mathcal{E}*}\prod_{j\in\Xi_k}\prod_{i_d\in\mathbb{D}_{d,j}}\left(1-q_{(d,i_d)\to j}^{\mathcal{E}}\right)\right) \tag{15}$$

measures the frequency of inside attacks launched from the type d end devices to the fog nodes within Ξ_k.

Remark 4.2. *Formula (14) is reminiscent of the inclusion-exclusion principle in combinatorics. Specifically, it is observed that,*

$$\mathbb{P}\left(\bigcap_{j\in\Xi} C_j^{\mathcal{F},[\bullet]} = 1\right) = 1 - \mathbb{P}\left(\bigcup_{j\in\Xi} C_j^{\mathcal{F},[\bullet]} = 0\right) = 1 - \sum_{k=1}^{m}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi}\mathbb{P}\left(\bigcap_{j\in\Xi_k} C_j^{\mathcal{F},[\bullet]} = 0\right),$$

where $\mathbb{P}\left(\bigcap_{j\in\Xi_k} C_j^{\mathcal{F},[\bullet]} = 0\right)$ can be computed explicitly via $h(\Xi_k)$. Within the expression of $h(\Xi_k)$, the former component captures the external attacks while the latter caters the inside attacks launched from the connected end nodes. Since the control center is assumed to have zero compromise probability, inside attacks originated from the other fog nodes are impossible to occur.

Next, we proceed to study quantify the cybersecurity risk of smart home platform without assuming zero compromise probability for the control center. The succeeding lemma is of auxiliary importance.

Lemma 4.3. *Consider the smart home network as illustrated in Figure 4, for a given set of m fog nodes, indexed by $\Xi = (\xi_1, \dots, \xi_m) \subseteq \{1, \dots, n^{\mathcal{F}}\}$, the following formula holds for the outside attack RV of the (d, j_d)-th end node belonging to the i-th smart home user:*

$$f(j_d,\Xi) = \mathbb{E}\left[O_{d,j_d}^{\mathcal{E}}\prod_{j\in\Xi} C_j^{\mathcal{F},[\bullet]}\right] = \omega_{d,j_d}^{\mathcal{E}} - \sum_{k=1}^{m}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi} u(j_d,\Xi_k)\,, \tag{16}$$

where

$$\begin{aligned} u(j_d,\Xi_k) = h(\Xi_k) \times g(d,\Xi_k)^{-1} & \\ \times \Bigg[\left(1-\nu_d^{\mathcal{T}}\right)\pi_{d,j_d}^{\mathcal{E}}\left(1-q_{(d,j_d)\to i}^{\mathcal{E}}\right)&\prod_{j\in\Xi_k}\prod_{l_d\in\mathbb{D}_{d,j},l_d\neq j_d}\left(1-\pi_{d,l_d}^{\mathcal{E}} q_{(d,l_d)\to j}^{\mathcal{E}}\right) \\ + \nu_d^{\mathcal{T}}\pi_d^{\mathcal{E}*}\prod_{j\in\Xi_k}\prod_{l_d\in\mathbb{D}_{d,j}}\left(1-q_{(d,l_d)\to j}^{\mathcal{E}}\right)\Bigg]. & \end{aligned}$$

***Theorem 4.4**. Consider the smart home network as illustrated in Figure 4, the comprise probability for the control center can be computed via*

$$p^{\mathcal{C}} := \mathbb{P}(C^{\mathcal{C}} = 1) = 1 - (1 - \omega^{\mathcal{C}}) \left[1 - \sum_{k=1}^{n^{\mathcal{F}}} (-1)^{k-1} \sum_{\Xi_k \subseteq \Xi^{\mathcal{F}}} \tilde{p}^{\mathcal{F}}_{\Xi_k} \prod_{i \in \Xi_k} q^{\mathcal{F}}_{i \to \bullet} \right],$$

where Ξ_k is any subset of $\Xi^{\mathcal{F}} = (1, \dots, n^{\mathcal{F}})$, and $\tilde{p}^{\mathcal{F}}_{\Xi_k}$ is the joint compromise probability for the fog nodes in Ξ_k which can be computed via (14), $k = 1, \dots, n^{\mathcal{F}}$.

Moreover, the underlying fog network has compromise probability for the j-th fog nodes, $i = 1, \dots, n^{\mathcal{F}}$:

$$p^{\mathcal{F}}_i = 1 - (1 - \tilde{p}^{\mathcal{F}}_i)(1 - q^{\mathcal{C}}_{\bullet \to i}) - q^{\mathcal{C}}_{\bullet \to i}(1 - \omega^{\mathcal{C}}) \left[1 - \sum_{k=1}^{n^{\mathcal{F}}} (-1)^{k-1} \sum_{\Xi_k \subseteq \Xi^{\mathcal{F}}} \tilde{p}^{\mathcal{F}}_{\Xi_k} \prod_{j \in \Xi_k, j \neq i} q^{\mathcal{F}}_{j \to \bullet} \right],$$

and the compromise probability for the j_d-th type d end node is given by, for $d = 1, \dots, n^{\mathcal{T}}, j_d = 1, \dots, n^{\mathcal{E}}_d$,

$$p^{\mathcal{E}}_{d,j_d} = \omega^{\mathcal{E}}_{d,j_d} + p^{\mathcal{F}}_i q^{\mathcal{F}}_{i \to (d,j_d)} - \omega^{\mathcal{E}}_{d,j_d} q^{\mathcal{F}}_{i \to (d,j_d)} + q^{\mathcal{F}}_{i \to (d,j_d)}(1 - q^{\mathcal{C}}_{\bullet \to j}) \times t_1 + q^{\mathcal{F}}_{i \to (d,j_d)} q^{\mathcal{C}}_{\bullet \to i}(1 - \omega^{\mathcal{C}}) \times t_2,$$

where

$$t_1 = \omega^{\mathcal{E}}_{d,j_d} - f(j_d, i),$$

and

$$t_2 = \omega^{\mathcal{E}}_{d,j_d} - \sum_{k=1}^{n^{\mathcal{F}}} (-1)^{k-1} \sum_{\Xi_k \subseteq \Xi^{\mathcal{F}}} f(j_d, \Xi_k) \prod_{j \in \Xi_k, j \neq i} q^{\mathcal{F}}_{j \to \bullet}.$$

Herein, the functions g and f are given in (15) and (16), respectively.

We illustrate the accuracy and effectiveness of the explicit formulas in Theorem 4.4 via the following example.

Example 4.5. *For the sake of exposition, let's consider a smaller smart home network with three users. However, we remark that results established in Theorem 4.4 can be applied to study smart home networks consisting of arbitrary number of users. Suppose that there are two types of smart home end devices which are illustrated in different colors in Figure 5. The outside attack probabilities associated with the fog nodes and end devices are summarized in Table 3. Based on the setting, we can conclude that the type 2 end devices are more vulnerable to outside attacks than the type 1 end devices in the sense that both the idiosyncratic and systemic attacks may occur more frequently. However, the smart hub is safer than the end devices against outside cyber-attacks. We also assume that the inside attack probabilities among end nodes and fog nodes are identical and equal to* 0.25.

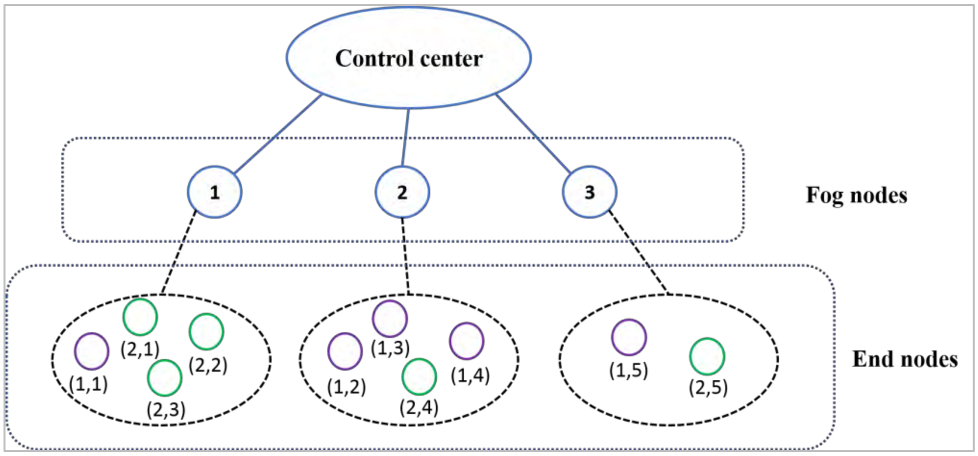
Table 3

OUTSIDE ATTACK PROBABILITIES OF EXAMPLE 4.5

	Smart hub	Type 1 home kits	Type 2 home kits
Idiosyncratic attack	0.1	0.2	0.3
Systemic attack	0.05	0.1	0.2
Common vulnerability	0.1	0.1	0.2

Typically, the control center would possess a higher level of security configuration, so we assume a lower outside attack probability $\omega^{\mathcal{C}} = 0.01$ *and inside infection probabilities* $q^{\mathcal{F}}_{j\to\bullet} = q^{\mathcal{C}}_{\bullet\to j} = 0.05, j = 1, 2, 3.$

Figure 5
CYBERSECURITY RISKS FACED BY A SMART HOME PLATFORM.



In our paper, the compromise probabilities can be computed using the explicit formulas in Theorem 4.4 or using MC simulations based on (1) – (3). Table 4 compares the performance of the precise calculation method proposed in this current section against the Monte Carlo simulation method for evaluating the compromise probabilities in the smart home infrastructure specified in Example 4.5. For each fixed sample size in the simulation study, the same experiment is repeated 1,000 times in order to construct the probability distribution of the empirical estimators for the compromise probabilities. The computation time of each simulation trial is reported at the end of Table 4.

Here are how the numerical results should be interpreted. First, the compromised probabilities computed using Theorem 4.4 coincide with the means of the estimated compromise probabilities based on Monte Carlo simulation, and the minor discrepancies are caused by the simulation fluctuations. As the same size n increases, the standard deviations of the empirical estimators of comprise probabilities decay at a rate of approximately $\sqrt{n}$, which complies with the large-sample theory of empirical means. Second, the explicit formulas proposed in Theorem 4.4 compute the compromise probabilities in much faster speeds than the simulation method. Third, the computed compromise probabilities make very intuitive sense. For instance, the control center has the lowest compromise probability among all the devices since its associated outside and insider infection probabilities are assumed to be lower. Among the three smart home users, the first user's smart hub has the highest compromise probability because the user possesses more type 2 end devices which are more vulnerable to outside cyber-attacks compared with the type 1 end devices. In contrast, the third smart home user has the least number of end devices, so the compromise probability is lowest. Within the same type of end devices, say type 1, the (1,5)-th device has the lowest compromise probability because the third user has only two end devices and insider attacks are less likely to occur. End devices (1,2), (1,3) and (1,4) have the same compromise probability because they belong to the same smart home user, while the outside and insider attack probabilities are assumed to be the same across the same type of end devices. The (1,1)-th end device has the highest compromise probability among the type 1 end devices. This can be explained by the fact that the first smart home user owns more number of the type 2 end devices which have higher cybersecurity risks, as mentioned earlier, and once infected, they may launch inside attacks to the (1,1)-th device. A similar argument can be adopted to explain the order in the compromise probabilities for the type 2 end devices.

Table 4
COMPARISON BETWEEN THE MONTE CARLO SIMULATION ESTIMATES OF THE COMPROMISE PROBABILITIES AND THE PRECISE CALCULATION ACCORDING TO THEOREM 4.4 FOR THE SMART HOME PLATFORM SPECIFIED AS PER EXAMPLE 4.5. FOR EACH SAMPLE SIZE $n \in \{1{,}000, 5{,}000, 25{,}000\}$, THE SIMULATION IS REPEATED 1,000 TIMES TO CALCULATE THE MEAN AND STANDARD DEVIATION (SD) OF THE COMPROMISE PROBABILITIES ESTIMATES. THE COMPUTATION SPEED REPORTED IN TERMS OF SECONDS, IS BASED ON A LAPTOP COMPUTER WITH THE 64 BIT WINDOWS 7 OPERATIONAL SYSTEM, AND A 3.3 GHZ CPU WITH 4 THREADS.

	Simulation						Explicit calculation
	n = 1,000		n = 5,000		n = 25,000		
	Mean	SD	Mean	SD	Mean	SD	
$p^{\mathcal{C}}$	0.046	0.007	0.047	0.003	0.048	0.001	0.048
$p_1^{\mathcal{F}}$	0.303	0.013	0.302	0.007	0.303	0.003	0.303
$p_2^{\mathcal{F}}$	0.272	0.014	0.274	0.006	0.272	0.002	0.272
$p_3^{\mathcal{F}}$	0.202	0.013	0.201	0.006	0.199	0.003	0.2
$p_{1,1}^{\mathcal{E}}$	0.244	0.013	0.244	0.006	0.244	0.003	0.244
$p_{1,2}^{\mathcal{E}}$	0.235	0.014	0.238	0.006	0.237	0.003	0.237
$p_{1,3}^{\mathcal{E}}$	0.237	0.014	0.237	0.006	0.237	0.002	0.237
$p_{1,4}^{\mathcal{E}}$	0.237	0.014	0.237	0.007	0.237	0.003	0.237
$p_{1,5}^{\mathcal{E}}$	0.223	0.015	0.223	0.006	0.222	0.003	0.222
$p_{2,1}^{\mathcal{E}}$	0.323	0.013	0.322	0.006	0.323	0.003	0.322
$p_{2,2}^{\mathcal{E}}$	0.324	0.014	0.322	0.006	0.322	0.003	0.322
$p_{2,3}^{\mathcal{E}}$	0.322	0.015	0.323	0.006	0.322	0.003	0.322
$p_{2,4}^{\mathcal{E}}$	0.319	0.014	0.319	0.007	0.319	0.003	0.319
$p_{2,5}^{\mathcal{E}}$	0.307	0.013	0.305	0.006	0.305	0.003	0.305
Time (sec.)	2.26		6.62		45.97		0.69

Section 5: Applications to Cybersecurity Insurance Pricing

Our discussion thus far focuses on modeling the frequency of compromise events in a given fog network. Namely, during a unit time period (e.g., one month/quarter/year), the compromised RV's, $\boldsymbol{C}$, indicates whether or not a specific node is compromised.

In the context of insurance pricing, the aggregate financial losses caused by the compromised nodes are of central interest. To this end, we resort to the frequency-severity approach which has evolved as an industry standard in pricing insurance risk generally, and cybersecurity risks particularly (see, Jevtić and Lanchier, 2020; Xu and Hua, 2019). To be specific, let $X^{\mathcal{C}} > 0, X_i^{\mathcal{F}} > 0$, and $X_{d,i_d}^{\mathcal{E}} > 0$ be the financial losses caused by an infection of the control center, the i-th fog node, and the (d, i_d)-th end node, respectively. It is assumed that these severity RV's $X^{\mathcal{C}}, X_i^{\mathcal{F}}$, and $X_{d,i_d}^{\mathcal{E}}$ are mutually independent, and also independent of the compromise status RV's $C^{\mathcal{C}}, C_i^{\mathcal{F}}$, and $C_{d,i_d}^{\mathcal{E}}, i = 1, \dots, n^{\mathcal{F}}, i_d = 1, .., n_d^{\mathcal{E}}, d = 1, \dots, n^{\mathcal{T}}$. The aggregate loss for the entire fog network of a smart home platform can be evaluated via

$$L = C^{\mathcal{C}} X^{\mathcal{C}} + \sum_{i=1}^{n^{\mathcal{F}}} C_i^{\mathcal{F}} X_i^{\mathcal{F}} + \sum_{d=1}^{n^{\mathcal{T}}} \sum_{i_d=1}^{n_d^{\mathcal{E}}} C_{d,i_d}^{\mathcal{E}} X_{d,i_d}^{\mathcal{E}}, \tag{17}$$

in which the three components cater the cybersecurity losses due to the compromises of control center, fog nodes and end nodes, respectively. It is noteworthy that although our discussion in this section is specialized for the smart home application, aggregate model (17) does not rely on any specific network topology, so in principle, it can be used to study any fog network.

To price the cybersecurity insurance, prevalent actuarial pricing principles include

Expectation principle: $$\varrho_1(L) = (1+\theta)\mathbb{E}[L]; \tag{18}$$

Standard deviation principle: $$\varrho_2(L) = \mathbb{E}[L] + \theta\sqrt{\mathrm{Var}(L)}; \tag{19}$$

Gini mean difference principle: $$\varrho_3(L) = \mathbb{E}[L] + \theta \mathrm{GMD}(L). \tag{20}$$

In the above pricing principles, $\theta > 0$ is the loading parameter, and for a pair of independent copies of L,

$$\mathrm{GMD}(L) = \mathbb{E}[|L_1 - L_2|]$$

given that the expectation exists, denotes the Gini mean difference (GMD) which is known to be a robust alternative of the standard deviation as a statistics measure of variability (see more detailed discussion in, e.g., Yitzhaki et al., 2003; Furman et al., 2017, 2019).

For the sake of illustration, in what follows, let us consider the insurance pricing for the smart home system considered in Example 4.5. Some additional assumptions related to the loss severity RV's are needed. It is natural that the compromise of the control center is more likely to result in more severe financial losses than the individual fog nodes and end nodes, so we assume

$$X^{\mathcal{C}} \sim \text{Lomax}(\alpha, \beta), \quad \alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+,$$

follows the heavy-tailed Lomax distribution, and

$$X_i^{\mathcal{F}} \sim \text{LN}(\mu, \sigma^2), \quad \mu \in \mathbb{R}, \sigma \in \mathbb{R}_+,$$

follows the moderately heavy-tailed log normal distribution, and

$$X_{d,i_d}^{\mathcal{E}} \sim \text{Exp}(\lambda_d), \lambda_d \in \mathbb{R}_+,$$

follows the exponential distribution which has a light tail.

In the evaluation of the expectation principle, it is straightforward to check that

$$\mathbb{E}[L] = p^{\mathcal{C}} \mu^{\mathcal{C}} + \sum_{i=1}^{n^{\mathcal{F}}} p_i^{\mathcal{F}} \mu_i^{\mathcal{F}} + \sum_{d=1}^{n^{\mathcal{T}}} \sum_{i_d=1}^{n_d^{\mathcal{E}}} p_{d,i_d}^{\mathcal{E}} \mu_{d,i_d}^{\mathcal{E}},$$

where $\mu^{\mathcal{C}} = \mathbb{E}[X^{\mathcal{C}}], \mu_i^{\mathcal{F}} = \mathbb{E}[X_i^{\mathcal{F}}], \mu_{d,i_d}^{\mathcal{E}} = \mathbb{E}[X_{d,i_d}^{\mathcal{E}}], i = 1, \dots, n^{\mathcal{F}}, i_d = 1, \dots, n_d^{\mathcal{E}}, d = 1, \dots, n^{\mathcal{T}}$, and the compromise probabilities $p^{\mathcal{C}}, p_i^{\mathcal{F}}$ and $p_{d,i_d}^{\mathcal{E}}$ can be computed explicitly according to Theorem 4.4. However, in the evaluation of the standard deviation principle and the GMD principle, both the variance and GMD of the aggregate loss L cannot (or otherwise, are very challenging to) be computed explicitly, so numerical computation via simulation is adopted.

In the succeeding numerical study, we choose the attack probabilities specified in Example 4.5 as the baseline parameters for the frequency model. The baseline parameters for the severity model are summarized in Table 5. This parameters setting indicates the followings. First, if a compromise occurs, then the control center has the highest average cybersecurity loss with the most heavy-tailed distribution. Second, the loss distributions for the fog nodes are assumed to be identical. Third, as specified earlier in the setup of Example 4.5, the type 2 home kits are more vulnerable to cyber-attacks than the type 1 home kits because the associated cybersecurity losses are lower.

Table 5
THE BASELINE PARAMETERS FOR THE SEVERITY MODELS OF DIFFERENT NETWORK ELEMENTS WITH THE SUMMARY STATISTICS OF THE ASSOCIATED DISTRIBUTIONS. THE DECIMALS ARE DROPPED FOR BRIEFNESS.

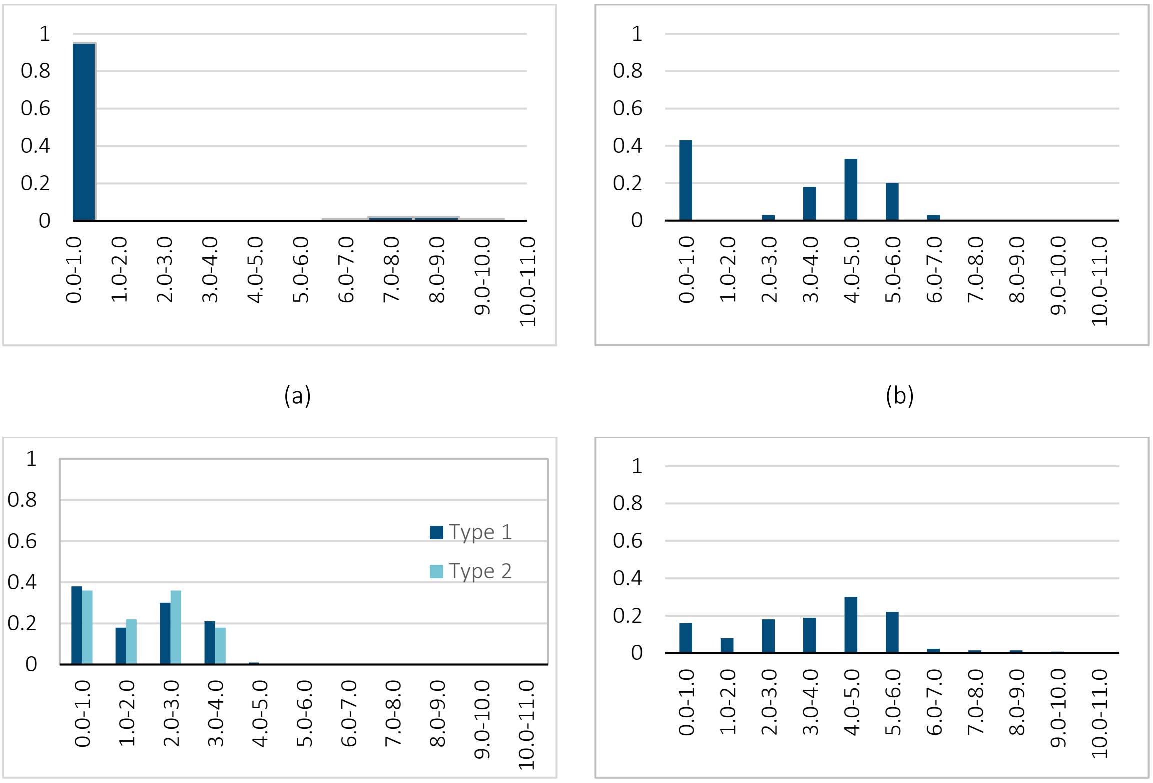
					Percentiles		
	Parameter	Mean	SD	GMD	25%	50%	75%
Control center	$(\alpha,\beta)=(5\times10^4,11)$	5000	5528	5238	1325	3252	6716
Smart hub	$(\mu,\sigma^2)=(4.26,0.83^2)$	100	100	88	93	99	106
Type 1 home kits	$\lambda_1=0.1$	10	10	10	3	7	14
Type 2 home kits	$\lambda_2=0.2$	5	5	5	1	3	7

Under the baseline parameters, Figure 6 displays the histograms of the shifted log transform of the cybersecurity loss for different components in the smart home network based on 105 times of simulation. As shown, the end devices have the most frequent cybersecurity losses because their security configurations are low and the associated compromise probabilities are high. The smart hubs / fog nodes have lower frequency of cybersecurity losses. The control center has the lowest rate of occurrence of cyber events, but once they occur, the financial loss can be very severe. As a consequence, the aggregate loss of the whole network also features a highly right skewed distribution, shedding light on the importance of having a fine risk management program in place for the cyber insurance provider to administer these tail losses.

A sensitivity analysis is conducted so as to identify the key parameters driving the insurance prices. In each scenario of the analysis, we shock a set of similar parameters by 50% while keeping the other baseline parameters unchanged, and then the variations in the cyber insurance prices are assessed. Table 6 depicts the sensitivities of the cyber insurance prices according to different types of cyber-attacks. We find that among the three actuarial pricing principles, the standard deviation principle yields the highest premium yet the expectation principle yields the lowest. The order is intuitive because, as shown earlier in Figure 6, the aggregate cybersecurity loss of the smart home network is highly right skewed, while as statistics measures of variations, the standard deviation penalizes large deviation harsher than the GMD. The insurance prices become lower in the downside case of the sensitivity analysis, because the probabilities of inside and outside attacks are lower, corresponding to a safer network. Among different attack types, the inside attack parameters have the most substantial influences to the insurance prices, which is again intuitive. Namely, in an unsecured network with high inside attack successful rates, even a single outside attack can be propagated in network and infect many other devices. Compared between the idiosyncratic attacks and systemic attacks, the insurance prices are more sensitive to the idiosyncratic attacks. The reason is that in this example, we assume relatively low occurrence rates of common vulnerabilities, so idiosyncratic attacks play a more dominating role in the determination of the compromise probability. In another unreported analysis where the common vulnerabilities probabilities are assume to be high, then we observe that the aforementioned order is reversed.

Figure 6

HISTOGRAMS OF THE SHIFTED LOG TRANSFORM (I.E., $f(l) = \log(l + 1)$, $l \geq 0$) OF THE FINANCIAL LOSSES DUE TO THE COMPROMISES OF CONTROL CENTER (TOP-LEFT), FOG NODES (TOP-RIGHT), AND END NODES (BOTTOM-LEFT), AS WELL AS THE AGGREGATE LOSS FOR THE ENTIRE NETWORK (BOTTOM-RIGHT).



(a)

(b)

(c)

(d)

Finally, the insurance price sensitivities in accordance with the compromise likelihoods of different network elements are examined. Based on Table 7, we observe that the fog nodes have the most noticeable impacts on the insurance prices, with a 50% increase in the compromise probabilities rises the insurance prices by about 25%. This is probably because the fog nodes have relatively high compromise frequency while the consequent financial losses are also higher than that of the end nodes. The control center possesses a high level of network security configuration, so the associated compromise frequency is very low, and the insurance prices are less sensitive to the changes in its compromise probability.

Collectively, our sensitivity analysis shows that the inside attack vulnerabilities and the economic losses due to compromised fog nodes are the key drivers in the smart home insurance premium calculation. However, the readers must also note that this conclusion are drawn based on the network topology specified in this section. The conclusion should not be directly generalized to all fog networks. Nevertheless, the readers can adopt the framework proposed in this present paper to identify the most sensitive components in their pricing problems.

Table 6

SENSITIVITY ANALYSIS OF THE CYBER INSURANCE PRICES IN RESPONSE TO THE CHANGES IN THE COMPROMISE PROBABILITIES AMONG DIFFERENT TYPES OF CYBER ATTACKS. EACH SET OF PARAMETERS

ARE SHOCKED BY −50% IN THE DOWNSIDE CASE AND +50% IN THE UPSIDE CASE. THE PRICES ϱ_1, ϱ_2 AND ϱ_3 ARE COMPUTED BASED ON THE PRICING PRINCIPLES SPECIFIED IN EQUATIONS (18) - (20) WITH $\theta =$ 0.1 THE PERCENTAGE OF CHANGE IN THE INSURANCE PRICE COMPARED WITH THE BASELINE PRICE IS REPORTED IN THE BRACKETS AFTER EACH SHOCKED PRICE.

Shocked parameters	Case	ϱ_1	ϱ_2	ϱ_3
	Baseline	355.388	479.845	378.589
Idiosyncratic attacks	Down	244.637(-31%)	370.289(-23%)	263.141(-30%)
$(\pi_i^{\mathcal{F}}$ or $\pi_{(d,i_d)}^{\mathcal{E}})$	Up	458.778(29%)	586.190(22%)	484.901(28%)
Systemic attacks	Down	346.962(-2%)	469.643(-2%)	369.654(-2%)
$(\pi^{\mathcal{F}*}, \pi_d^{\mathcal{E}*})$	Up	387.636(9%)	527.621(10%)	412.887(9%)
Common vulnerabilities	Down	351.698(-1%)	472.729(-1%)	374.038(1%)
$(\ v^{\mathcal{F}}, v_d^{\mathcal{E}})$	Up	368.413 (4%)	503.146(5%)	392.713(4%)
Inside attacks	Down	202.418(-43%)	292.745(-39%)	215.057(-43%)
$(q_{j\to\cdot}^{\mathcal{F}}, q_{\cdot\to j}^{\mathcal{C}}, q_{j\to(d,i)}^{\mathcal{F}}, q_{(d,i)\to j}^{\mathcal{E}})$	Up	565.294(59%)	713.606(49%)	601.435(59%)

Table 7

SENSITIVITY ANALYSIS OF THE CYBER INSURANCE PRICES IN RESPONSE TO THE CHANGES IN THE COMPROMISE PROBABILITIES AMONG DIFFERENT TYPES OF NODES. THE SET-UP OF THE SENSITIVITY ANALYSIS IS SAME AS THAT OF TABLE 6.

Shocked parameters	Case	ϱ_1	ϱ_2	ϱ_3
	Baseline	355.388	479.845	378.589
Control center	Down	341.397(-4%)	460.020(-4%)	363.028(-4%)
$(\omega^{\mathcal{C}})$	Up	369.010(4%)	492.982(3%)	393.138(4%)
Fog nodes	Down	310.522(-13%)	432.914(-10%)	331.293(-12%)
$(\pi_i^{\mathcal{F}}, \pi^{\mathcal{F}*}, v^{\mathcal{F}})$	Up	439.996(24%)	585.632(22%)	468.503(24%)
Type 1 end nodes	Down	317.642(-11%)	434.87(-9%)	339.120(-10%)
$(\pi_{(1,i_1)}^{\mathcal{E}}, \pi_1^{\mathcal{E}*}, v_1^{\mathcal{E}})$	Up	403.906(14%)	540.448(13%)	429.029(13%)
Type 2 end nodes	Down	317.737(-11%)	458.763(-9%)	339.587(-11%)
$(\pi_{(2,i_2)}^{\mathcal{E}}, \pi_2^{\mathcal{E}*}, v_2^{\mathcal{E}})$	Up	406.4879(14%)	537.1767(12%)	431.4025(14%)

Section 6: Further discussion

Admittedly, cyber insurance data available for academic research are very scarce. The focus on the emerging concept of fog computing further limits the data availability. Due to the absence of available data, we are not able to conduct a back-testing to validate the accuracy of the modeling framework. In practice, parameters such as the outside and inside attack probabilities and the ones that characterize the loss distributions of compromise components, can be only chosen according to expert knowledge. However, we do hope that our work can draw more attentions from the actuarial community to this interesting research area, so that more data may be generated and collected in the near future. When fog network attack data and insurance loss data become available, actuaries can obtain more accurate estimation the model's parameters based on the real life data. The modeling framework and the pricing approach suggested in this current paper will remain useful.

Section 7: Conclusions

In this paper, we proposed a class of mathematical models to describe the cybersecurity risk propagation mechanisms in a general fog network. We investigated the associations among a variety of risk contributors in the determination of a fog network's cybersecurity risk. For a general fog network, we suggested an interval approximation method to assess the compromise probabilities of individual network elements. For the fog network underlying a smart home platform, we obtained a set of explicit formulas to calculation the compromise probabilities precisely. A quantitative framework based on actuarial pricing principles has been proposed to price the cyber insurance contract for the smart home applications. It was discovered that the impacts of heterogeneity and interdependency should never be overlooked in the fog networks. The inherent common vulnerabilities is also crucial in determining the risk and related pricing strategies.

The study on the fog computing from the risk management and actuarial perspectives are still in its infancy. The main challenges are caused by its multi-tenant and resource-sharing architectures, which results in a considerably large attack surface. The current work makes a significant first step towards tackling the problem of modeling and pricing the cybersecurity risk in fog networks. Moving forward, we are interested in the following interesting yet challenging issues: i) *Dynamic cyber risk*. In our current study, the compromise probabilities are assumed to be static. In some practical instances, the dynamic probabilities may be desired. Therefore, a proper dynamic epidemic spreading model can be developed for this purpose; see, e.g., Xu and Hua (2019). ii) *Cascading effects*. The cascading failure can be incorporated into the modeling process, which refers to the failure of one or several nodes triggering the failure of other nodes. Note that although the cascading failure and cyber risk propagation are similar, the cascading failure mainly focuses on the physical layer, while the cyber risk propagation focuses on the communication/network layer. Since the cascading failure is not uncommon in the fog computing, it can be considered as the other risk factor. One may refer to Xing (2020) for a recent review on cascading failure in the IoT. iii) *Cybersecurity risk aggregation*. Since the fog computing is widely deployed for a variety of IoT applications, an insurance company can have several businesses lines, e.g., smart home, fog servers, and smart cars, constituting a cyber insurance portfolio. To realize the diversification benefit and properly understand the systemic risk inherent in the portfolio, the calculation of aggregate risk capital is of interest for the insurance company, which should be carefully investigated.

Section 8: Acknowledgments

The researchers' deepest gratitude goes to those without whose efforts this project could not have come to fruition: The Project Oversight Group for their diligent work reviewing and editing this report for accuracy and relevance.

Project Oversight Group members:

Syed Danish Ali

Jan Hou Chong

Joseph Hayes

Kelvin Lam

Shan Liu

Andrea Marcovici

Lisa Martell

Rasa McKean

Julie Meadows

Gabriel Penagos

Tamara Wilt

Zhijing Xu

At the Society of Actuaries:

Rob Montgomery

Erika Schulty

Appendix A: Technical proofs

Proof of Proposition 3.5. We begin by proving that the outside attack RV's are positively associated. Focus on fog nodes first and consider the RV $\boldsymbol{O}^{\mathcal{F}} = (O_1^{\mathcal{F}}, \dots, O_{n^{\mathcal{F}}}^{\mathcal{F}})$, it is straightforward to check that

$$\begin{aligned}\mathbb{P}\left(O_i^{\mathcal{F}} > o_i \middle| O_j^{\mathcal{F}} = o_j, j = 1, \dots, i-1\right) = (1 - v^{\mathcal{F}})\mathbb{P}\left(O_i^{\mathcal{F}} > o_i \middle| V^{\mathcal{F}} = 0\right) \\ + v^{\mathcal{F}}\mathbb{P}(O_i^{\mathcal{F}} > o_i | V^{\mathcal{F}} = 1, O_j^{\mathcal{F}} = o_j, j = 1, \dots, i-1)\end{aligned}$$

is nondecreasing in $(o_1, \dots, o_{i-1}) \in \{0,1\}^{i-1}$ for all $o_i \in \mathbb{R}$, $i = 2, \dots, n^{\mathcal{F}}$. So the RV $\boldsymbol{O}^{\mathcal{F}}$ is conditionally increasing in sequence, which implies it is positively associated (see, Theorem 2.4 in Joe, 1997). Repeated applications of the aforementioned argument to $\boldsymbol{O}_d^{\mathcal{E}}$, $d = 1, \dots, n^{\mathcal{T}}$, yield that each of them is positively associated. Because of Assumption 3.1, the RV $\boldsymbol{O}^{\mathcal{F}}, \boldsymbol{O}_1^{\mathcal{E}}, \dots, \boldsymbol{O}_{n^{\mathcal{T}}}^{\mathcal{E}}$ are mutually independent, so $\boldsymbol{O}$ is positively associated.

By Assumption 3.2, $\boldsymbol{I}$ is independent hence positively associated. Moreover, $\boldsymbol{I}$ and $\boldsymbol{O}$ are independent. Thereby, RV $(\boldsymbol{I}, \boldsymbol{O})$ is positively associated.

Next, note that state equations (1) and (3) can be expressed as

$$C_i^{\mathcal{F}} = h_i(\boldsymbol{I}, \boldsymbol{O}) \quad \text{and} \quad C_{d,i_d}^{\mathcal{E}} = h_{d,i_d}(\boldsymbol{I}, \boldsymbol{O}),$$

respectively, for some coordinate-wise non-decreasing functions $h_i(\cdot)$ and $h_{d,i_d}(\cdot), i = 1, \dots, n^{\mathcal{F}}, d = 1, \dots, n^{\mathcal{T}}, i_d = 1, \dots, n_d^{\mathcal{E}}$. Evoking Lemma 3.4, we can conclude that $(\boldsymbol{C}, \boldsymbol{I}, \boldsymbol{O})$ is positively associated. This completes the proof. ■

Proof of Theorem 3.6. At the beginning, note that the compromise probabilities associated with state equations (1) and (3) can be computed via*,*

$$p_j^{\mathcal{F}} = 1 - \mathbb{P}\left(O_j^{\mathcal{F}} \le 0, \bigcap_{i=1, i\neq j}^{n^{\mathcal{F}}} C_i^{\mathcal{F}} I_{i\to j}^{\mathcal{F}} \le 0, \bigcap_{d=1}^{n^{\mathcal{T}}} \bigcap_{i_d=1}^{n_d^{\mathcal{E}}} C_{d,i_d}^{\mathcal{E},[j]} I_{(d,i_d)\to j}^{\mathcal{E}} \le 0\right), \quad \text{for } j = 1, \dots, n^{\mathcal{F}}, \tag{21}$$

and

$$p_{d,j_d}^{\mathcal{E}} = 1 - \mathbb{P}\left(O_{d,j_d}^{\mathcal{E}} \le 0, \bigcap_{i=1, i\neq j}^{n^{\mathcal{F}}} C_i^{\mathcal{F}} I_{i\to(d,j_d)}^{\mathcal{F}} \le 0\right), \quad \text{for } d = 1, \dots, n^{\mathcal{T}}, j_d = 1, \dots, n_d^{\mathcal{E}}. \tag{22}$$

Thus the task in this proof boils down to identifying the lower and upper bounds for the cumulative distribution functions (CDF) of dependent binary RV's in Equations (21) and (22).

First, we consider the compromise RV's $C_{d,i_d}^{\mathcal{E},[j]}$ defined in Equation (2). We have proved in Proposition 3.5 that $(\boldsymbol{C}, \boldsymbol{I}, \boldsymbol{O})$ is positively associated. On the one hand, because positive association implies positive lower orthant dependence (Shaked, 1982), it holds that

$$\mathbb{E}\left[C_{d,i_d}^{\mathcal{E},[j]}\right] = 1 - \mathbb{P}\left(O_{d,i_d}^{\mathcal{E}} \le 0, \bigcap_{i=1, i\neq j}^{n^{\mathcal{F}}} C_i^{\mathcal{F}} I_{i\to(d,i_d)}^{\mathcal{F}} \le 0\right)$$

$$\leq 1-\mathbb{P}\left(O_{d,i_d}^{\mathcal{E}} \leq 0\right) \prod_{i=1,i\neq j}^{n^{\mathcal{F}}} \mathbb{P}\left(C_i^{\mathcal{F}} I_{i\to(d,i_d)}^{\mathcal{F}} \leq 0\right)$$
$$= 1-\left[1-\nu_d^{\mathcal{T}}\pi_d^{\mathcal{E}*}-\left(1-\nu_d^{\mathcal{T}}\right)\pi_{d,i_d}^{\mathcal{E}}\right] \prod_{i=1,i\neq j}^{n^{\mathcal{F}}} \left(1-p_i^{\mathcal{F}} q_{i\to(d,i_d)}^{\mathcal{F}}\right).$$

On the other hand, by Fréchet inequalities (Fréchet, 1951), we readily obtain

$$\mathbb{P}\left(O_{d,i_d}^{\mathcal{E}} \leq 0, \bigcap_{i=1,i\neq j}^{n^{\mathcal{F}}} C_i^{\mathcal{F}} I_{i\to(d,i_d)}^{\mathcal{F}} \leq 0\right) \leq \min\left(\mathbb{P}\left(O_{d,i_d}^{\mathcal{E}} \leq 0\right), \bigwedge_{i=1,i\neq j}^{n^{\mathcal{F}}} \mathbb{P}\left(C_i^{\mathcal{F}} I_{i\to(d,i_d)}^{\mathcal{F}} \leq 0\right)\right)$$
$$= 1-\max\left(\nu_d^{\mathcal{T}}\pi_d^{\mathcal{E}*}+\left(1-\nu_d^{\mathcal{T}}\right)\pi_{d,i_d}^{\mathcal{E}}, \bigvee_{i=1,i\neq j}^{n^{\mathcal{F}}} p_i^{\mathcal{F}} q_{i\to(d,i_d)}^{\mathcal{F}}\right).$$

So we get

$$\max\left(\omega_{d,i_d}^{\mathcal{E}}, \bigvee_{i=1,i\neq j}^{n^{\mathcal{F}}} p_i^{\mathcal{F}} q_{i\to(d,i_d)}^{\mathcal{F}}\right) \leq \mathbb{E}\left[C_{d,i_d}^{\mathcal{E},[j]}\right] \leq 1-\left(1-\omega_{d,i_d}^{\mathcal{E}}\right) \prod_{i=1,i\neq j}^{n^{\mathcal{F}}} \left(1-p_i^{\mathcal{F}} q_{i\to(d,i_d)}^{\mathcal{F}}\right). \tag{23}$$

Next, we turn to the compromise probabilities of fog nodes in Equation (21). Another application of the property of positive association yields

$$p_j^{\mathcal{F}} = 1-\mathbb{P}\left(O_j^{\mathcal{F}} \leq 0, \bigcap_{i=1,i\neq j}^{n^{\mathcal{F}}} C_i^{\mathcal{F}} I_{i\to j}^{\mathcal{F}} \leq 0, \bigcap_{d=1}^{n^{\mathcal{T}}} \bigcap_{i_d=1}^{n_d^{\mathcal{E}}} C_{d,i_d}^{\mathcal{E},[j]} I_{(d,i_d)\to j}^{\mathcal{E}} \leq 0\right)$$
$$\leq 1-\mathbb{P}\left(O_j^{\mathcal{F}} \leq 0\right) \prod_{i=1,i\neq j}^{n^{\mathcal{F}}} \mathbb{P}\left(C_i^{\mathcal{F}} I_{i\to j}^{\mathcal{F}} \leq 0\right) \prod_{d=1}^{n^{\mathcal{T}}} \prod_{i_d=1}^{n_d^{\mathcal{E}}} \mathbb{P}\left(C_{d,i_d}^{\mathcal{E},[j]} I_{(d,i_d)\to j}^{\mathcal{E}} \leq 0\right)$$
$$= 1-\left(1-\omega_j^{\mathcal{F}}\right) \prod_{i=1,i\neq j}^{n^{\mathcal{F}}} \left(1-p_i^{\mathcal{F}} q_{i\to j}^{\mathcal{F}}\right) \prod_{d=1}^{n^{\mathcal{T}}} \prod_{i_d=1}^{n_d^{\mathcal{E}}} \left[1-\mathbb{E}\left[C_{d,i_d}^{\mathcal{E},[j]}\right] q_{(d,i_d)\to j}^{\mathcal{E}}\right]. \tag{24}$$

Evoke the upper bound derived in Equation (23), we get

$$1-\mathbb{E}\left[C_{d,i_d}^{\mathcal{E},[j]}\right] q_{(d,i_d)\to j}^{\mathcal{E}} \geq 1-q_{(d,i_d)\to j}^{\mathcal{E}}+q_{(d,i_d)\to j}^{\mathcal{E}}\left(1-\omega_{d,i_d}^{\mathcal{E}}\right) \prod_{i=1,i\neq j}^{n^{\mathcal{F}}} \left(1-p_i^{\mathcal{F}} q_{i\to(d,i_d)}^{\mathcal{F}}\right)$$
$$\geq 1\text{-}q_{(d,i_d)\to j}^{\mathcal{E}}+q_{(d,i_d)\to j}^{\mathcal{E}}\left(1-\omega_{d,i_d}^{\mathcal{E}}\right) \prod_{i=1,i\neq j}^{n^{\mathcal{F}}} \left(1-q_{i\to(d,i_d)}^{\mathcal{F}}\right). \tag{25}$$

Combining the inequalities derived in (24) and (25) leads to

$$p_j^{\mathcal{F}} \leq 1-\gamma_j \prod_{i=1,i\neq j}^{n^{\mathcal{F}}} \left(1-p_i^{\mathcal{F}} q_{i\to j}^{\mathcal{F}}\right) \overset{(1)}{\leq} 1-\gamma_j\left(1-\sum_{i=1,i\neq j}^{n^{\mathcal{F}}} p_i^{\mathcal{F}} q_{i\to j}^{\mathcal{F}}\right),$$

where inequality $\overset{(1)}{=}$ holds by Weierstrass product inequality. Define an $n^{\mathcal{F}}$ by $n^{\mathcal{F}}$ zero diagonal matrix, $\boldsymbol{A}$ with off-diagonal elements $a_{ij} = \gamma_i q^{\mathcal{F}}_{j\to i}$ for $i \neq j = 1, \dots, n^{\mathcal{F}}$. Then the upper bound of the compromise probabilities for fog nodes, $\boldsymbol{u}^{\mathcal{F}} = (u^{\mathcal{F}}_1, \dots, u^{\mathcal{F}}_{n^{\mathcal{F}}})^\top$, solves the matrix equation $\boldsymbol{u}^{\mathcal{F}} = 1 - \boldsymbol{\gamma} + \boldsymbol{A}\boldsymbol{u}^{\mathcal{F}}$, or equivalently, $\boldsymbol{u}^{\mathcal{F}} = (\boldsymbol{1} - \boldsymbol{A})^{-1}(1 - \boldsymbol{\gamma})$, where $\boldsymbol{1}$ denotes an identify matrix of appropriate dimension.

Contrastingly,

$$\begin{aligned} p^{\mathcal{F}}_j &= 1 - \mathbb{P}\left(O^{\mathcal{F}}_j \leq 0, \bigcap_{i=1, i\neq j}^{n^{\mathcal{F}}} C^{\mathcal{F}}_i I^{\mathcal{F}}_{i\to j} \leq 0\,, \bigcap_{d=1}^{n^{\mathcal{T}}} \bigcap_{i_d=1}^{n^{\mathcal{E}}_d} C^{\mathcal{E},[j]}_{d,i_d} I^{\mathcal{E}}_{(d,i_d)\to j} \leq 0 \right) \\ &\overset{(1)}{\geq} \max\left(\omega^{\mathcal{F}}_j, \bigvee_{i=1, i\neq j}^{n^{\mathcal{F}}} p^{\mathcal{F}}_i q^{\mathcal{F}}_{i\to j}\,, \bigvee_{d=1}^{n^{\mathcal{T}}} \bigvee_{i_d=1}^{n^{\mathcal{E}}_d} \mathbb{E}\left[C^{\mathcal{E},[j]}_{d,i_d} \right] q^{\mathcal{E}}_{(d,i_d)\to j} \right) \\ &\overset{(2)}{\geq} \max\left(\omega^{\mathcal{F}}_j, \bigvee_{i=1, i\neq j}^{n^{\mathcal{F}}} p^{\mathcal{F}}_i q^{\mathcal{F}}_{i\to j}\,, \bigvee_{d=1}^{n^{\mathcal{T}}} \bigvee_{i_d=1}^{n^{\mathcal{E}}_d} \max\left(\omega^{\mathcal{E}}_{d,i_d}, p^{\mathcal{F}}_i q^{\mathcal{F}}_{i\to(d,i_d)} q^{\mathcal{E}}_{(d,i_d)\to j} \right) \right) \\ &\geq \max\left(\omega^{\mathcal{F}}_j, \bigvee_{i=1, i\neq j}^{n^{\mathcal{F}}} \beta_i q^{\mathcal{F}}_{i\to j}\,, \bigvee_{d=1}^{n^{\mathcal{T}}} \bigvee_{i_d=1}^{n^{\mathcal{E}}_d} \max\left(\omega^{\mathcal{E}}_{d,i_d}, \bigvee_{i=1, i\neq j}^{n^{\mathcal{F}}} \beta_i q^{\mathcal{F}}_{i\to(d,i_d)} \right) q^{\mathcal{E}}_{(d,i_d)\to j} \right), \end{aligned}$$

where inequalities “$\overset{(1)}{=}$” and “$\overset{(2)}{=}$” hold because of Fréchet inequalities and the lower bound derived in Equation (23), and

$$\beta_j = \max\left(\omega^{\mathcal{F}}_j, \bigvee_{i=1, i\neq j}^{n^{\mathcal{F}}} \omega^{\mathcal{F}}_i q^{\mathcal{F}}_{i\to j}\,, \bigvee_{d=1}^{n^{\mathcal{T}}} \bigvee_{i_d=1}^{n^{\mathcal{E}}_d} \max\left(\omega^{\mathcal{E}}_{d,i_d}, \bigvee_{i=1, i\neq j}^{n^{\mathcal{F}}} \omega^{\mathcal{F}}_i q^{\mathcal{F}}_{i\to(d,i_d)} \right) q^{\mathcal{E}}_{(d,i_d)\to j} \right).$$

We have now obtained the lower bounds for the compromise probabilities of fog nodes.

Applying the same argument as in the derivation of inequalities (23) yields

$$\max\left(\omega^{\mathcal{E}}_{d,j_d}, \bigvee_{i=1}^{n^{\mathcal{F}}} p^{\mathcal{F}}_i q^{\mathcal{F}}_{i\to(d,j_d)} \right) \leq p^{\mathcal{E}}_{d,j_d} \leq 1\text{-}(1 - \omega^{\mathcal{E}}_{d,j_d}) \prod_{i=1}^{n^{\mathcal{F}}} \left(1 - p^{\mathcal{F}}_i q^{\mathcal{F}}_{i\to(d,j_d)}\right),$$

for $d = 1, \dots, n^{\mathcal{T}}$ and $j_d = 1, \dots, n^{\mathcal{E}}_d$. Finally, substitute the lower and upper bounds for $\boldsymbol{p}^{\mathcal{F}}$ into the inequalities above, the interval approximations for the end nodes' compromise probabilities are readily obtained.

Now the proof is finished. ■

Proof of Proposition 4.1. Recall that for $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $\mathcal{N} = \{1, \dots, n\}$, the following equation holds:

$$\prod_{i=1}^{n} (1 - x_i) = 1 - \sum_{k=1}^{n} (-1)^{k-1} \sum_{\mathcal{N}_k \subseteq \mathcal{N}} h(\mathcal{N}_k)\,, \qquad (26)$$

where $h(\mathcal{N}_k)=\prod_{i\in\mathcal{N}_k} x_i$ for $\mathcal{N}_k\in\mathbb{N}^k$ is any d-dimensional subset of $\mathcal{N}$. Together with Equation (11), we have

$$\begin{aligned}\tilde{p}_{\Xi}^{\mathcal{F}} &= \mathbb{E}\left[\prod_{j\in\Xi} C_j^{\mathcal{F},[\bullet]}\right] = \mathbb{E}\left[\prod_{j\in\Xi}\left[1-(1-O_j^{\mathcal{F}})\prod_{d=1}^{n^{\mathcal{T}}}\prod_{i_d\in\mathbb{D}_{d,j}}(1-O_{d,i_d}^{\mathcal{E}}\times I_{(d,i_d)\to j}^{\mathcal{E}})\right]\right]\\ &= 1-\sum_{k=1}^{m}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi} h(\Xi_k)\end{aligned}$$

where $m=|\Xi|$ is the cardinality of Ξ, $\Xi_k\in\mathbb{N}^k$ denotes any k-dimensional subset of Ξ, and

$$\begin{aligned}h(\Xi_k) &= \mathbb{E}\left[\prod_{j\in\Xi_k}(1-O_j^{\mathcal{F}})\times\prod_{d=1}^{n^{\mathcal{T}}}\prod_{i_d\in\mathbb{D}_{d,j}}(1-O_{d,i_d}^{\mathcal{E}}\times I_{(d,i_d)\to j}^{\mathcal{E}})\right]\\ &= \mathbb{E}\left[\prod_{j\in\Xi_k}(1-O_j^{\mathcal{F}})\right]\times\mathbb{E}\left[\prod_{j\in\Xi_k}\prod_{d=1}^{n^{\mathcal{T}}}\prod_{i_d\in\mathbb{D}_{d,j}}(1-O_{d,i_d}^{\mathcal{E}}\times I_{(d,i_d)\to j}^{\mathcal{E}})\right].\end{aligned}$$

The expectations above can be further computed via

$$\begin{aligned}\mathbb{E}\left[\prod_{j\in\Xi_k}(1-O_j^{\mathcal{F}})\right] &= (1-\nu^{\mathcal{F}})\mathbb{E}\left[\prod_{j\in\Xi_k}(1-O_j^{\mathcal{F}})\,|V^{\mathcal{F}}=0\right]+\nu^{\mathcal{F}}\mathbb{E}\left[\prod_{j\in\Xi_k}(1-O_j^{\mathcal{F}})\,|V^{\mathcal{F}}=1\right]\\ &= (1-\nu^{\mathcal{F}})\prod_{j\in\Xi_k}(1-\pi_j^{\mathcal{F}})+\nu^{\mathcal{F}}(1-\pi^{\mathcal{F}*}),\end{aligned}$$

as well as

$$\begin{aligned}&\mathbb{E}\left[\prod_{j\in\Xi_k}\prod_{d=1}^{n^{\mathcal{T}}}\prod_{i_d\in\mathbb{D}_{d,j}}(1-O_{d,i_d}^{\mathcal{E}}\times I_{(d,i_d)\to j}^{\mathcal{E}})\right]\\ &=\prod_{d=1}^{n^{\mathcal{T}}}\mathbb{E}\left[\prod_{j\in\Xi_k}\prod_{i_d\in\mathbb{D}_{d,j}}(1-O_{d,i_d}^{\mathcal{E}}\times I_{(d,i_d)\to j}^{\mathcal{E}})\right]\\ &=\prod_{d=1}^{n^{\mathcal{T}}}\left\{(1-\nu_d^{\mathcal{T}})\mathbb{E}\left[\prod_{j\in\Xi_k}\prod_{i_d\in\mathbb{D}_{d,j}}(1-O_{d,i_d}^{\mathcal{E}}\times I_{(d,i_d)\to j}^{\mathcal{E}})|V_d^{\mathcal{T}}=0\right]\right.\\ &\qquad\left.+\nu_d^{\mathcal{T}}\mathbb{E}\left[\prod_{j\in\Xi_k}\prod_{i_d\in\mathbb{D}_{d,j}}(1-O_{d,i_d}^{\mathcal{E}}\times I_{(d,i_d)\to j}^{\mathcal{E}})|V_d^{\mathcal{T}}=1\right]\right\}\\ &=\prod_{d=1}^{n^{\mathcal{T}}}\left[(1-\nu_d^{\mathcal{T}})\prod_{j\in\Xi_k}\prod_{i_d\in\mathbb{D}_{d,j}}\left(1-\pi_{d,i_d}^{\mathcal{E}}q_{(d,i_d)\to j}^{\mathcal{E}}\right)+\nu_d^{\mathcal{T}}\left(1-\pi_d^{\mathcal{E}*}+\pi_d^{\mathcal{E}*}\prod_{j\in\Xi_k}\prod_{i_d\in\mathbb{D}_{d,j}}(1-q_{(d,i_d)\to j}^{\mathcal{E}})\right)\right]\\ &=\prod_{d=1}^{n^{\mathcal{T}}}g(d,\Xi_k).\end{aligned}\tag{27}$$

We have now obtained the desired result, and the proof is completed. ■

Proof of Lemma 4.3. The proof is somewhat similar to that of Proposition 4.1, so we skip certain details for brevity. It holds that

$$
\begin{aligned}
f(j_d,\Xi) &= \mathbb{E}\left[O_{d,j_d}^{\mathcal{E}}\prod_{j\in\Xi}C_j^{\mathcal{F},[\bullet]}\right] = \mathbb{E}\left[O_{d,j_d}^{\mathcal{E}}\prod_{j\in\Xi}\left[1-(1-O_j^{\mathcal{F}})\prod_{s=1}^{n^{\mathcal{T}}}\prod_{l_s\in\mathbb{D}_{s,j}}(1-O_{s,l_s}^{\mathcal{E}}\times I_{(s,l_s)\to j}^{\mathcal{E}})\right]\right] \\
&= \omega_{d,j_d}^{\mathcal{E}}-\sum_{k=1}^{m}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi}u(j_d,\Xi_k),
\end{aligned}
$$

where

$$
\begin{aligned}
u(j_d,\Xi_k) &= \mathbb{E}\left[O_{d,j_d}^{\mathcal{E}}\prod_{j\in\Xi_k}(1-O_j^{\mathcal{F}})\prod_{s=1}^{n^{\mathcal{T}}}\prod_{l_s\in\mathbb{D}_{s,j}}(1-O_{s,l_s}^{\mathcal{E}}\times I_{(s,l_s)\to j}^{\mathcal{E}})\right] \\
&= \left[(1-\nu^{\mathcal{F}})\prod_{j\in\Xi_k}(1-\pi_j^{\mathcal{F}})+\nu^{\mathcal{F}}(1-\pi^{\mathcal{F}*})\right]\times\mathbb{E}\left[O_{d,j_d}^{\mathcal{E}}\prod_{j\in\Xi_k}\prod_{s=1}^{n^{\mathcal{T}}}\prod_{l_s\in\mathbb{D}_{s,j}}(1-O_{s,l_s}^{\mathcal{E}}\times I_{(s,l_s)\to j}^{\mathcal{E}})\right].
\end{aligned}
$$

The expectation above is computed via

$$
\begin{aligned}
&\mathbb{E}\left[O_{d,j_d}^{\mathcal{E}}\prod_{j\in\Xi_k}\prod_{s=1}^{n^{\mathcal{T}}}\prod_{l_s\in\mathbb{D}_{s,j}}(1-O_{s,l_s}^{\mathcal{E}}I_{(s,l_s)\to j}^{\mathcal{E}})\right] \\
&= \mathbb{E}\left[O_{d,j_d}^{\mathcal{E}}\prod_{j\in\Xi_k}\prod_{l_d\in\mathbb{D}_{d,j}}(1-O_{l_d}^{\mathcal{E}}I_{(d,l_d)\to j}^{\mathcal{E}})\right]\prod_{s=1,s\neq d}^{n^{\mathcal{T}}}\mathbb{E}\left[\prod_{j\in\Xi_k}\prod_{l_s\in\mathbb{D}_{s,j}}(1-O_{s,l_s}^{\mathcal{E}}I_{(s,l_s)\to j}^{\mathcal{E}})\right] \\
&= \left[(1-\nu_d^{\mathcal{T}})\pi_{d,j_d}^{\mathcal{E}}(1-q_{(d,j_d)\to i}^{\mathcal{E}})\prod_{j\in\Xi_k}\prod_{l_d\in\mathbb{D}_{d,j},l_d\neq j_d}(1-\pi_{d,l_d}^{\mathcal{E}}q_{(d,l_d)\to j}^{\mathcal{E}})+\nu_d^{\mathcal{T}}\pi_d^{\mathcal{E}*}\prod_{j\in\Xi_k}\prod_{l_d\in\mathbb{D}_{d,j}}(1-q_{(d,l_d)\to j}^{\mathcal{E}})\right] \\
&\qquad\times\prod_{s=1,s\neq d}^{n^{\mathcal{T}}}g(s,\Xi_k).
\end{aligned}
$$

This yields the desired result, and the proof is completed. ■

Proof of Theorem 4.4. From the state equation of the control center as per (10), we have

$$
\begin{aligned}
p^{\mathcal{C}} &= 1-\mathbb{E}[(1-O^{\mathrm{C}})]\times\mathbb{E}\left[\prod_{i=1}^{n^{\mathcal{F}}}\left(1-C_i^{\mathcal{F},[\bullet]}\times I_{i\to\bullet}^{\mathcal{F}}\right)\right] \\
&= 1-(1-\omega^{\mathrm{C}})[1-\sum_{k=1}^{n^{\mathcal{F}}}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi^{\mathcal{F}}}\tilde{p}_{\Xi_k}^{\mathcal{F}}\prod_{i\in\Xi_k}q_{i\to\bullet}^{\mathcal{F}}]
\end{aligned}
$$

where the last equation holds because of the product formula in (26).

Turing to the study of fog nodes, elaborate the state equation (12) as

$$
C_i^{\mathcal{F}} = 1-\left(1-C_i^{\mathcal{F},[\bullet]}\right)\left(1-I_{\bullet\to i}^{\mathcal{C}}\right)-\left(1-C_i^{\mathcal{F},[\bullet]}\right)I_{\bullet\to i}^{\mathcal{C}}(1-O^{\mathcal{C}})\prod_{j=1,j\neq i}^{n^{\mathcal{F}}}\left(1-C_j^{\mathcal{F},[\bullet]}I_{j\to\bullet}^{\mathcal{F}}\right), \tag{28}
$$

and hence

$$p_i^{\mathcal{F}} = \mathbb{E}[C_i^{\mathcal{F}}] = 1 - (1 - \tilde{p}_i^{\mathcal{F}})(1 - q_{\bullet\to i}^{\mathcal{C}}) - q_{\bullet\to i}^{\mathcal{C}}(1 - \omega^{\mathcal{C}})\mathbb{E}\left[(1 - C_i^{\mathcal{F},[\bullet]}) \prod_{j=1,j\neq i}^{n^{\mathcal{F}}} (1 - C_j^{\mathcal{F},[\bullet]} I_{j\to\bullet}^{\mathcal{F}})\right].$$

To compute the expectation above, for $j = 1, \dots, n^{\mathcal{F}}$, define

$$\tilde{I}_{j\to\bullet}^{\mathcal{F}} \equiv \begin{cases} I_{j\to\bullet}^{\mathrm{F}}, & \text{if } j \neq i; \\ 1, & \text{if } j = i. \end{cases}$$

We have

$$\begin{aligned}
\mathbb{E}\left[\left(1 - C_i^{\mathcal{F},[\bullet]}\right) \prod_{j=1,j\neq i}^{n^{\mathcal{F}}} \left(1 - C_j^{\mathcal{F},[\bullet]} I_{j\to\bullet}^{\mathcal{F}}\right)\right] &= \mathbb{E}\left[\prod_{j=1}^{n^{\mathcal{F}}} \left(1 - C_j^{\mathcal{F},[\bullet]} \tilde{I}_{j\to\bullet}^{\mathcal{F}}\right)\right] \\
&= 1 - \sum_{k=1}^{\mathcal{F}} (-1)^{k-1} \sum_{\Xi_k \subseteq \Xi^{\mathcal{F}}} \mathbb{E}\left[\prod_{j\in\Xi_k} C_j^{\mathcal{F},[\bullet]}\right] \mathbb{E}\left[\prod_{j\in\Xi_k} \tilde{I}_{j\to\bullet}^{\mathcal{F}}\right] \\
&= 1 - \sum_{k=1}^{n^{\mathcal{F}}} (-1)^{k-1} \sum_{\Xi_k \subseteq \Xi^{\mathcal{F}}} \tilde{p}_{\Xi_k}^{\mathcal{F}} \prod_{j\in\Xi_k, j\neq i} q_{j\to\bullet}^{\mathcal{F}}.
\end{aligned}$$

Finally, let us consider the compromise probability for the j_d-th end node that is of type d, $d = 1, \dots, n^{\mathcal{T}}$, $j_d = 1 \dots, n_d^{\mathcal{E}}$, and it has a direct connection to the i-th fog node, i.e., $j_d \in \mathbb{D}_{d,i}$. According to the state equation (13), we get

$$\begin{aligned}
p_{d,j_d}^{\mathcal{E}} &= \mathbb{E}\left[1 - \left(1 - O_{d,j_d}^{\mathcal{E}}\right)\left(1 - C_i^{\mathcal{F}} I_{i\to(d,j_d)}^{\mathcal{F}}\right)\right] \\
&= \mathbb{E}\left[O_{d,j_d}^{\mathcal{E}} + C_i^{\mathcal{F}} I_{i\to(d,j_d)}^{\mathcal{F}} - O_{d,j_d}^{\mathcal{E}} C_i^{\mathcal{F}} I_{i\to(d,j_d)}^{\mathcal{F}}\right] \\
&= \omega_{d,j_d}^{\mathcal{E}} + p_i^{\mathcal{F}} q_{i\to(d,j_d)}^{\mathcal{F}} - \mathbb{E}[O_{d,j_d}^{\mathcal{E}} C_i^{\mathcal{F}} I_{i\to(d,j_d)}^{\mathcal{F}}].
\end{aligned}$$

Evoking the state equation for fog node in (28), the expectation above can be computed via

$$\begin{aligned}
\mathbb{E}\left[O_{d,j_d}^{\mathcal{E}} C_i^{\mathcal{F}} I_{i\to j_d}^{\mathcal{F}}\right] = \mathbb{E}\left[O_{d,j_d}^{\mathcal{E}} I_{i\to(d,j_d)}^{\mathcal{F}}\right] &- \mathbb{E}\left[O_{d,j_d}^{\mathcal{E}} I_{i\to(d,j_d)}^{\mathcal{F}} \left(1 - C_i^{\mathcal{F},[\bullet]}\right)\left(1 - I_{\bullet\to i}^{\mathcal{C}}\right)\right] \\
&- \mathbb{E}\left[O_{d,j_d}^{\mathcal{E}} I_{i\to(d,j_d)}^{\mathcal{F}} \left(1 - C_i^{\mathcal{F},[\bullet]}\right) I_{\bullet\to i}^{\mathcal{C}} (1 - O^{\mathcal{C}}) \prod_{j=1,j\neq i}^{n^{\mathcal{F}}} \left(1 - C_j^{\mathcal{F},[\bullet]} I_{j\to\bullet}^{\mathcal{F}}\right)\right] \\
&= \omega_{d,j_d}^{\mathcal{E}} q_{i\to(d,j_d)}^{\mathcal{F}} - q_{i\to(d,j_d)}^{\mathcal{F}} (1 - q_{\bullet\to j}^{\mathcal{C}}) \times t_1 - q_{i\to(d,j_d)}^{\mathcal{F}} q_{\bullet\to i}^{\mathcal{C}} (1 - \omega^{\mathcal{C}}) \times t_2
\end{aligned}$$

in which, by evoking Lemma 4.3,

$$t_1 = \mathbb{E}[O_{d,j_d}^{\mathcal{E}}(1 - C_i^{\mathcal{F},[\bullet]})] = \mathbb{E}[O_{d,j_d}^{\mathcal{E}}] - \mathbb{E}[O_{d,j_d}^{\mathcal{E}} C_i^{\mathcal{F},[\bullet]}] = \omega_{d,j_d}^{\mathcal{E}} - f(j_d, i),$$

and

$$t_2 = \mathbb{E}[O_{d,j_d}^{\mathcal{E}}(1 - C_i^{\mathcal{F},[\bullet]}) \prod_{j=1,j\neq i}^{n^{\mathcal{F}}} (1 - C_j^{\mathcal{F},[\bullet]} I_{j\to\bullet}^{\mathcal{F}})].$$

We focus on the evaluation of t_2 and obtain

$$
\begin{aligned}
t_2 &= \mathbb{E}\left[O_{j_d}^{\mathcal{E}}\left(1 - C_i^{\mathcal{F},[\bullet]}\right)\prod_{j=1, j\neq i}^{n^{\mathcal{F}}}\left(1 - C_j^{\mathcal{F},[\bullet]} I_{j\to\bullet}^{\mathcal{F}}\right)\right] \\
&= \mathbb{E}\left[O_{j_d}^{\mathcal{E}}\prod_{j=1}^{n^{\mathcal{F}}}\left(1 - C_j^{\mathcal{F},[\bullet]} \tilde{I}_{j\to\bullet}^{\mathcal{F}}\right)\right] \\
&= \mathbb{E}\left[O_{j_d}^{\mathcal{E}}\left[1 - \sum_{k=1}^{n^{\mathcal{F}}}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi^{\mathcal{F}}}\left[\prod_{j\in\Xi_k} C_j^{\mathcal{F},[\bullet]}\right]\left[\prod_{j\in\Xi_k}\tilde{I}_{j\to\bullet}^{\mathcal{F}}\right]\right]\right] \\
&= \omega_{j_d}^{\mathcal{E}} - \sum_{k=1}^{n^{\mathcal{F}}}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi^{\mathcal{F}}} f(j_d, \Xi_k)\prod_{j\in\Xi_k, j\neq i} q_{j\to\bullet}^{\mathcal{F}},
\end{aligned}
$$

where $f(j_d, \Xi_k) = \mathbb{E}[O_{j_d}^{\mathcal{E}} \prod_{j\in\Xi_k} C_j^{\mathcal{F},[\bullet]}]$ which can be computed by evoking Lemma 4.3.

■

Appendix B: Summary of the notation system

We summarize the notation system used in this present article herein.

Table 8

SUMMARY OF THE NOTATION SYSTEM.

Notation	Description
$n^{\mathcal{F}}$	Number of fog nodes
$n^{\mathcal{T}}$	Number of types of end nodes
$n_d^{\mathcal{E}}$	Number of type d end nodes
$C^{\mathcal{C}}, C_i^{\mathcal{F}}, C_{d,i_d}^{\mathcal{E}}$	Compromise statuses of the control center, fog nodes and end nodes
$p^{\mathcal{C}}, p_i^{\mathcal{F}}, p_{d,i_d}^{\mathcal{E}}$	Compromise probabilities of the control center, fog nodes and end nodes
$C^{\mathcal{C},[j]}$	Compromise status of the control center with the j-fog node excluded
$C_j^{\mathcal{F},[\bullet]}$	Compromise status of the j-th fog node with control center excluded
$C_{d,i_d}^{\mathcal{E},[j]}$	Compromise status of the (d, i_d)-th end node with the j-th fog node excluded
$O^{\mathcal{C}}, O_i^{\mathcal{F}}, O_{d,i_d}^{\mathcal{E}}$	Outside attack statuses of the control center, fog nodes and end nodes
$\omega^{\mathcal{C}}, \omega_i^{\mathcal{F}}, \omega_{d,i_d}^{\mathcal{E}}$	Outside compromise probabilities of the control center, fog nodes and end nodes
$I_{\bullet\to i}^{\mathcal{C}}$	Indicators of inside attack launched from the control center to the fog nodes
$q_{\bullet\to i}^{\mathcal{C}}$	Probabilities of inside attack launched from the control center to the fog nodes
$I_{i\to\bullet}^{\mathcal{F}}, I_{i\to j}^{\mathcal{F}}, I_{i\to(d,i_d)}^{\mathcal{F}}$	Indicators of inside attacks launched from the i-th fog node
$q_{i\to\bullet}^{\mathcal{F}}, q_{i\to j}^{\mathcal{F}}, q_{i\to(d,i_d)}^{\mathcal{F}}$	Probabilities of inside attacks launched from the i-th fog node
$I_{(d,i_d)\to i}^{\mathcal{E}}$	Indicators of inside attacks launched from the (d, i_d)-th end node to fog nodes
$q_{(d,i_d)\to i}^{\mathcal{E}}$	Probabilities of inside attacks launched from the (d, i_d)-th end node to fog nodes
$V^{\mathcal{F}}, V_d^{\mathcal{E}}$	Indicators of common vulnerabilities among fog nodes and the type d end nodes
$v^{\mathcal{F}}, v_d^{\mathcal{E}}$	Probabilities of common vulnerability among fog nodes and the type d end nodes
$\pi_i^{\mathcal{F}}, \pi_{d,i_d}^{\mathcal{E}}$	Systemic outside attack probabilities among the fog nodes and type d end nodes
$\pi^{\mathcal{F}*}, \pi_d^{\mathcal{E}*}$	Idiosyncratic outside attack probabilities for the fog nodes and end nodes

References

Baccarelli, E., Naranjo, P. G. V., Scarpiniti, M., Shojafar, M., and Abawajy, J. H. (2017). Fog of everything: Energy-efficient networked computing architectures, research challenges, and a case study. *IEEE Access*, 5:9882–9910.

Biener, C., Eling, M., and Wirfs, J. H. (2015). Insurability of cyber risk: An empirical analysis. *Geneva Papers on Risk and Insurance: Issues and Practice*, 40(1):131–158.

Böhme, R. (2005). Cyber-Insurance Revisited. *WEIS*.

Darwish, T. S. and Bakar, K. A. (2018). Fog based intelligent transportation big data analytics in the internet of vehicles environment: Motivations, architecture, challenges, and critical issues. *IEEE Access*, 6:15679–15701.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R., and Vyncke, D. (2002a). The concept of comonotonicity in actuarial science and finance: Applications. *Insurance: Mathematics and Economics*, 31(2):133–161.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R., and Vyncke, D. (2002b). The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics*, 31(1):3–33.

Eling, M. and Jung, K. (2018). Copula approaches for modeling cross-sectional dependence of data breach losses. *Insurance: Mathematics and Economics*, 82:167–180.

Eling, M. and Loperfido, N. (2017). Data breaches: Goodness of fit, pricing, and risk measurement. *Insurance: Mathematics and Economics*, 75:126–136.

Eling, M. and Wirfs, J. (2019). What are the actual costs of cyber risk events? *European Journal of Operational Research*, 272(3):1109–1119.

Fahrenwaldt, M. A., Weber, S., and Weske, K. (2018). Pricing of cyber insurance contracts in a network model. *ASTIN Bulletin*, 48(3):1175–1218.

Fang, Z., Xu, M., Xu, S., and Hu, T. (2021). A framework for predicting data breach risk: leveraging dependence to cope with sparsity. *IEEE Transactions on Information Forensics and Security*, 16:2186–2201.

Feng, S., Xiong, Z., Niyato, D., Wang, P., and Leshem, A. (2018). Evolving risk management against advanced persistent threats in fog computing. In *2018 IEEE 7th International Conference on Cloud Networking (CloudNet)*, pages 1–6. IEEE.

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont donn´ees. *Annals de l'Universit´e de Lyon*, 9:53–77.

Furman, E., Kye, Y., and Su, J. (2019). Computing the gini index: A note. *Economics Letters*, 185:108753.

Furman, E., Wang, R., and Zitikis, R. (2017). Gini-type measures of risk and variability: Gini shortfall, capital allocations, and heavy-tailed risks. *Journal of Banking & Finance*, 83:70–84.

Jevtić, P. and Lanchier, N. (2020). Dynamic structural percolation model of loss distribution for cyber risk of small and medium-sized enterprises for tree-based LAN topology. *Insurance: Mathematics and Economics*, 91:209–223.

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. Chapman and Hall, London.

Khan, S., Parkinson, S., and Qin, Y. (2017). Fog computing security: A review of current applications and security solutions. *Journal of Cloud Computing*, 6(1):19.

Kraemer, F. A., Braten, A. E., Tamkittikhun, N., and Palma, D. (2017). Fog computing in healthcare–a review and discussion. *IEEE Access*, 5:9206–9222.

Lynn, T., Endo, P., Ribeiro, A., Barbosa, G., and Rosati, P. (2020). The Internet of Things: Definitions, Key Concepts, and Reference Architectures. In *The Cloud-to-Thing Continuum*, pages 1–22. Palgrave Macmillan, Cham.

McShane, M. and Nguyen, T. (2020). Time-varying effects of cyberattacks on firm value. *Geneva Papers on Risk and Insurance: Issues and Practice*, 45(4):580–615.

NAIC (2019). 2019 Report on the Cybersecurity Insurance and Identity Theft Coverage Supplement. Technical report, National Association of Insurance Commissioners.

Peng, C., Xu, M., Xu, S., and Hu, T. (2018). Modeling multivariate cybersecurity risks. *Journal of Applied Statistics*, 45(15):2718–2740.

Puliafito, C., Mingozzi, E., Longo, F., Puliafito, A., and Rana, O. (2019). Fog computing for the Internet of Things: A survey. *ACM Transactions on Internet Technology (TOIT)*, 19(2):1–41.

Shaked, M. (1982). A general theory of some positive dependence notions. *Journal of Multivariate Analysis*, 12(2):199–218.

Sohal, A. S., Sandhu, R., Sood, S. K., and Chang, V. (2018). A cybersecurity framework to identify malicious edge device in fog computing and cloud-of-things environments. *Computers & Security*, 74:340–354.

Sun, H., Xu, M., and Zhao, P. (2020). Modeling malicious hacking data breach risks. *North American Actuarial Journal*, pages 1–19.

Wheatley, S., Maillart, T., and Sornette, D. (2016). The extreme risk of personal data breaches and the erosion of privacy. *European Physical Journal B*, 89(1):1–12.

Xing, L. (2020). Cascading failures in internet of things: review and perspectives on reliability and resilience. *IEEE Internet of Things Journal*, 8(1):44–64.

Xu, M., Da, G., and Xu, S. (2015). Cyber epidemic models with dependences. *Internet Mathematics*, 11(1):62–92.

Xu, M. and Hua, L. (2019). Cybersecurity insurance: Modeling and pricing. *North American Actuarial Journal*, 23(2):220–249.

Yitzhaki, S. et al. (2003). Gini’s mean difference: A superior measure of variability for nonnormal distributions. *Metron*, 61(2):285–316.

Yu, T., Sekar, V., Seshan, S., Agarwal, Y., and Xu, C. (2015). Handling a trillion (unfixable) flaws on a billion devices: Rethinking network security for the internet-of-things. In *Proceedings of the 14th ACM Workshop on Hot Topics in Networks*, page 5. ACM.

About The Society of Actuaries

With roots dating back to 1889, the Society of Actuaries (SOA) is the world's largest actuarial professional organization with more than 31,000 members. Through research and education, the SOA's mission is to advance actuarial knowledge and to enhance the ability of actuaries to provide expert advice and relevant solutions for financial, business and societal challenges. The SOA's vision is for actuaries to be the leading professionals in the measurement and management of risk.

The SOA supports actuaries and advances knowledge through research and education. As part of its work, the SOA seeks to inform public policy development and public understanding through research. The SOA aspires to be a trusted source of objective, data-driven research and analysis with an actuarial perspective for its members, industry, policymakers and the public. This distinct perspective comes from the SOA as an association of actuaries, who have a rigorous formal education and direct experience as practitioners as they perform applied research. The SOA also welcomes the opportunity to partner with other organizations in our work where appropriate.

The SOA has a history of working with public policymakers and regulators in developing historical experience studies and projection techniques as well as individual reports on health care, retirement and other topics. The SOA's research is intended to aid the work of policymakers and regulators and follow certain core principles:

Objectivity: The SOA's research informs and provides analysis that can be relied upon by other individuals or organizations involved in public policy discussions. The SOA does not take advocacy positions or lobby specific policy proposals.

Quality: The SOA aspires to the highest ethical and quality standards in all of its research and analysis. Our research process is overseen by experienced actuaries and nonactuaries from a range of industry sectors and organizations. A rigorous peer-review process ensures the quality and integrity of our work.

Relevance: The SOA provides timely research on public policy issues. Our research advances actuarial knowledge while providing critical insights on key policy issues, and thereby provides value to stakeholders and decision makers.

Quantification: The SOA leverages the diverse skill sets of actuaries to provide research and findings that are driven by the best available data and methods. Actuaries use detailed modeling to analyze financial risk and provide distinct insight and quantification. Further, actuarial standards require transparency and the disclosure of the assumptions and analytic approach underlying the work.

Society of Actuaries
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org