European Journal of Operational Research 305 (2023) 1273–1291

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

Structural models for fog computing based internet of things architectures with insurance and risk management applications

Xiaoyu Zhang[a], Maochao Xu[b], Jianxi Su[c,*], Peng Zhao[d]

[a] Department of Statistics and Finance, University of Science and Technology of China, Hefei, Anhui 230052, China
[b] Department of Mathematics, Illinois State University, Normal, IL 61790, USA
[c] Department of Statistics, Purdue University, West Lafayette, IN 47906, USA
[d] School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou, Jiangsu 221116, China

ARTICLE INFO

Article history:
Received 21 June 2021
Accepted 21 July 2022
Available online 25 July 2022

Keywords:
Applied probability
Actuarial pricing principles
Comonotonic approximation
Risk propagation
Smart home

ABSTRACT

Cybersecurity risk modeling and pricing are becoming a spotlight in actuarial science and operational research. This paper pertains to the analysis of the cybersecurity risks involved in the fog computing technology which has been intensively deployed in assorted Internet of Things (IoT) applications. To this end, a class of structural models are established to study the inherent cyber risk propagation process. Under the smart home applications, we manage to compute the compromise probabilities of individual nodes explicitly. Applications of the proposed structural models in the context of cyber insurance pricing are thoroughly explored. Finally, we propose an interval method for estimating the compromise probabilities of fog network's elements, which can be used to efficiently identify weak nodes for cybersecurity risk management.

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 2020. As such, studies on how to model and manage cybersecurity risks have attracted much scholar attention from the operational research discipline (e.g., Cheung & Bell, 2021; Eling & Wirfs, 2019; Khouzani, Liu, & Malacaria, 2019; Nagurney & Shukla, 2017; Paul & Zhang, 2021; Simon & Omar, 2020). Although much effort has been made on the design of network infrastructure so as to enhance the internet security, vulnerabilities cannot be fully eliminated in actual practice. To manage the residual cybersecurity risks, network service providers and users are often advised to seek insurance as a robust financial protection in case of cyber events (Biener, Eling, & Wirfs, 2015; Böhme, 2005). Following the high 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 more than 500 cyber insurance providers to date, with direct written premium amounted to three billion (NAIC, 2020).

During the recent years, as the proliferation and consumerization of Internet of Things (IoT) technology[2] continue to evolve at a rampant pace, a growing number of electronic devices are interconnected through multiple intricate networks, which generate enormous data that 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 in cloud computing (Puliafito, Mingozzi, Longo, Puliafito, & Rana, 2019). Specifically, fog computing is a decentralized computing infrastructure that extends the cloud service to

* Corresponding author.
E-mail address: jianxi@purdue.edu (J. Su).

[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.

[2] There is no universal definition on IoT (see, Lynn, Endo, Ribeiro, Barbosa, & Rosati, 2020, for a variety of descriptions from either the technical or sociotechnical perspectives). One may be heuristically understood as a network infrastructure of interconnected devices (i.e., things) for processing information from the physical and the virtual world.

https://doi.org/10.1016/j.ejor.2022.07.033
0377-2217/© 2022 Elsevier B.V. All rights reserved.

the edges of networks, hence computational resources are closer to the positions where the data are generated and used. Compared with the traditional cloud computing, fog computing features superior user-experience and failure tolerance. Thereby, fog computing has been widely deployed in a variety of IoT applications including smart home (Puliafito et al., 2019), health data management (Kraemer, Braten, Tamkittikhun, & Palma, 2017), intelligent transportation system (Darwish & Bakar, 2018), public services such as power grid, military defense, and critical national infrastructure (Baccarelli, Naranjo, Scarpiniti, Shojafar, & Abawajy, 2017). A typical fog network structure possesses the multi-tenant (e.g., computers, laptops, smart devices, automated cars, traffic lights) and resource-sharing (e.g., the connected fog nodes) features.

Although fog computing is emerging as a scalable, reliable and cost effective solution for big data analytic in the IoT domain, its multi-tenant and resource-sharing architectures induce an unprecedented degree of cybersecurity risks to the technology's users (Khan, Parkinson, & Qin, 2017). The Ponemon Institute[3] estimated that the percentage of organizations who reported data breaches due to unsecured IoT devices/applications has climbed from 15 percent in 2017 to 26 percent in 2019. However, the actual percentage may be even much higher since most organizations are not aware of the threats in their network environment. The figures underscore the acute needs for the IoT stakeholders to carefully manage the cybersecurity risks, and cyber insurance—as a natural tool for cybersecurity risk transfer and mitigation—should play a pivotal role in the evolution.

An effective implementation of cyber insurance in the IoT risk management process requires the inherent cybersecurity risks to be properly understood and the insurance products to be fairly priced. That said, quantitative frameworks for modeling and pricing the IoT cybersecurity risks seem to have gathered little scholar attention so far. Earlier studies related to fog networks are from the IT perspective, focusing on the development and deployment of the technology for IoT applications. The only relevant work that we are aware of is Feng, Xiong, Niyato, Wang, & Leshem (2018), where the risk management process for a fog network 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 to tackle the problem and aim at putting forth a class of structural models for modeling and pricing the cybersecurity risks of fog networks underlying the IoT applications. The proposed models capture the unique cybersecurity features of fog networks, so they are significantly different from the other network models considered in the cybersecurity risk management literature (e.g., Eling & Wirfs, 2019; Fahrenwaldt, Weber, & Weske, 2018; Jevtić & Lanchier, 2020; Xu, Da, & Xu, 2015; Xu & Hua, 2019). In addition, it is noteworthy that the problem of modeling cyber risk propagation shares several similarities with the default contagion problems considered in the areas of quantitative finance (e.g., Agosto & Ahelegbey, 2022; Capponi & Jarrow, 2021; Detering, Meyer-Brandis, Panagiotou, & Ritter, 2019; Veraart, 2020). Thereby, the results of our paper may go beyond studying cybersecurity risks and can be also useful in other applications such as economics, finance and insurance.

Here is a preview of the contributions of our paper:

- From the mathematical modeling and insurance pricing perspectives, we identify the key cybersecurity risk drivers underlying fog networks.
- We propose a general class of structural models for studying the compromise statuses of different types of nodes in fog networks.
- Focusing on smart home applications, we obtain explicit formulas for evaluating the compromise probabilities of control center, smart home hubs, and end devices.
- We consider the applications of three state-of-the-art actuarial principles for pricing the cybersecurity risks of a smart home network. Our numerical analysis favors the use of standard deviation principle from both the conservative ratemaking and robustness perspectives.
- We establish an interval method for approximating the compromise probabilities for a general fog network. The approximation method is useful for efficient identification of weak nodes.

It is worth mentioning that the study of weak node identification falls into the domain of network vulnerability identification and security measurement in the computer science field. One strand of the literature measures the network security by using state-based stochastic models to represent the operational statuses of network devices. For instance, Almasizadeh & Azgomi (2013); Dacier, Deswarte, & Kaâniche (1996); Ortalo, Deswarte, & Kaâniche (1999) applied the Markov states model to study the time to failure as a network security metric. Another strand of the literature quantifies the network security via the attack graph approach which describes the set of vulnerabilities and their dependencies within a network. To name a few related works, Wang, Zhang, & Kadobayashi (2013) applied the dependency attack graph approach to model a network's vulnerabilities in a cost-benefit analysis of security hardening. A similar vulnerability analysis problem was investigated by Poolsappasit, Dewri, & Ray (2011) via Bayesian attack graphs. Wang, Chen, Zhao, Di, & Liu (2018) proposed a quantitative framework for assessing a network's vulnerability degree by using the network's attack graph and maximum flow. For a detailed literature survey, we refer the readers to Ramos, Lazar, Holanda Filho, & Rodrigues (2017). These existing methods aim to precisely evaluate some security metrics such as the time to failure, the likelihood of vulnerability exploitation, the shortest attack path, etc., which can be onerous to implement for a complicated network structure. Focusing on IoT applications, our paper contributes to the related literature with a new probability-based interval approximation approach for identifying weak nodes, which possesses satisfactory performance and is efficient to implement.

The rest of the paper is organized as follows. Beginning with a non-technical discussion about the unique characteristics of the cybersecurity risks involved in fog networks in Section 2, we propose a quantitative framework for modeling the cybersecurity risk propagation process in Section 3. To exemplify the applications of the proposed network models, in Section 4, we study the cybersecurity risks in a smart home system which corresponds to one of the most popular IoT applications these days. Cybersecurity risk pricing is considered in Section 5 with numerical illustrations. In Section 6, we propose an interval approximation method to estimate the compromise probabilities for the individual network elements in a general fog network, which can be used to efficiently identify weak nodes. Section 7 concludes the paper. In order to facilitate the readings, Appendix A contains the technical proofs for the paper's main results, and Appendix B contains a summary of the notation system used throughout the paper.

2. Characterizing the cybersecurity risks in fog networks

In the development of cyber risk models for fog networks, a set of salient characteristics associated with the fog network structure must be accommodated thoroughly. Firstly, the components in 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

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

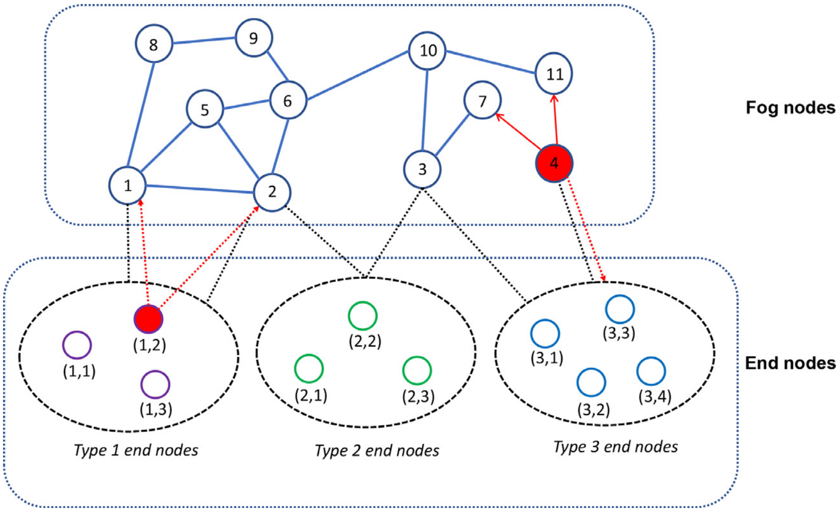

Fig. 1. An illustration of the cybersecurity risks faced by a fog network in the IoT application. In this hypothetical network, there are 11 fog nodes, and 3 distinct types of end nodes/devices marked in different colors. The label (d, i_d) indicates the i_dth type d end nodes. The compromised nodes and the surrounding propagation paths are indicated in red color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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. In a traditional centralized network, 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 lightweights of operating systems and the relatively low computational capabilities of the IoT devices (Yu, Sekar, Seshan, Agarwal, & Xu, 2015). The commonly used security protocol of fog network attempts to authenticate each edge device before providing data or computation resources, so the vulnerabilities hidden in the IoT devices will become 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 via DDoS attacks. Worse still, common vulnerabilities often exist in fog networks since similar computing nodes and end devices are operated under the same security configuration. These common vulnerabilities trigger the build-up of systemic cybersecurity risks. If a common vulnerability is identified and utilized by attackers, then devastating damages may occur to the entire network. In the cybersecurity risk management and pricing, it is critical to account for such high severity scenarios.

Thirdly, fog networks are also vulnerable to inside attacks. The inside attacks are caused by the compromised fog nodes or end devices that are already inside the network. Once penetrated into the network, the attacker can gain certain privileges in the network and advance toward the other fog nodes or edge devices easily without being discovered (Sohal, Sandhu, Sood, & Chang, 2018).

For illustrative purposes, an abstract fog network is displayed in Fig. 1, in which there is one compromised fog node and one compromised end node. The cybersecurity risks may be propagated via the network. Specifically, fog node 4 is compromised by an outside attack, which can propagate the risk to its neighboring nodes 7 and 11 via inside attacks. It can also propagate the risk to its connected end nodes, i.e., the type 3 end nodes. Similarly, the second type 1 end node is also compromised, and it can propagate the risk to its connected fog nodes 1 and 2. If fog node 2 is compromised, it can further compromise its connected fog nodes and end nodes, i.e., the type 1 and type 2 end nodes.

3. Modeling the cyber risk propagation in fog networks

In the study of cybersecurity risk, the typical first step is to model the occurrences of cyber attacks as well as the infection propagation process within the network system. To this end, we put forth a class of structural models for modeling the compromise frequencies among different components of a fog network. In particular, the proposed infection models accommodate all the indispensable characteristics outlined in Section 2.

For a given fog network, let $n^{\mathscr{F}}$ be the number of fog nodes, $n^{\mathscr{T}}$ be the number of end node types, and $n_d^{\mathscr{E}}$ be the number of end nodes that are of type $d \in \{1, \ldots, n^{\mathscr{T}}\}$. Here and in the sequel, the superscripts "$\mathscr{F}$", "$\mathscr{T}$", and "$\mathscr{E}$" indicate that a specific object of interest is related to fog nodes, the types of end nodes, and end nodes, respectively. To illuminate, in the hypothetical fog network displayed in Fig. 1, we have

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

For the ith fog node, $i \in \{1, \ldots, n^{\mathscr{F}}\}$, define the compromise status random variable (RV), $C_i^{\mathscr{F}} \in \{0, 1\}$, with $C_i^{\mathscr{F}} = 1$ means that the node is compromised, and zero otherwise. Similarly, $C_{d,i_d}^{\mathscr{E}} \in \{0, 1\}$ indicates the compromise status for the (d, i_d)th end node, $d \in \{1, \ldots, n^{\mathscr{T}}\}$, $i_d \in \{1, \ldots, n_d^{\mathscr{E}}\}$. Herein and throughout, the label (d, i_d) represents the i_dth type d end node. Whenever an indicator RV is used, the value one represents a positive risk status and zero represents a negative risk status. We also shorthand the compromise RV's by

$$\boldsymbol{C} = \left(\boldsymbol{C}^{\mathscr{F}}, \boldsymbol{C}_1^{\mathscr{E}}, \ldots, \boldsymbol{C}_{n^{\mathscr{T}}}^{\mathscr{E}}\right),$$

$$\text{where } \boldsymbol{C}^{\mathscr{F}} = (C_1^{\mathscr{F}}, \ldots, C_{n^{\mathscr{F}}}^{\mathscr{F}}),\ \boldsymbol{C}_d^{\mathscr{E}} = (C_{d,1}^{\mathscr{E}}, \ldots, C_{d,n_d^{\mathscr{E}}}^{\mathscr{E}}),\ d \in \{1, \ldots, n^{\mathscr{T}}\}.$$

The study of $\boldsymbol{C}$ further depends on the frequencies of outside attacks and inside attacks which we are going to discuss next.

Let $O_i^{\mathscr{F}} \in \{0, 1\}$ and $O_{d,i_d}^{\mathscr{E}} \in \{0, 1\}$ be respectively the outside attack status RV's for the ith fog node, $i = 1, \ldots, n^{\mathscr{F}}$, and the (d, i_d)th end node, $d \in \{1, \ldots, n^{\mathscr{T}}\}$, $i_d \in \{1, \ldots, n_d^{\mathscr{E}}\}$. To account for the presence of systemic cybersecurity risk discussed in Section 2, we consider two types of vulnerabilities that can be exploited by outside attackers. Outside attacks through a common vulnerability may imperil all the nodes that are of the same type, potentially causing systemic failures among a cohort of network components. In contrast, an idiosyncratic vulnerability may only exist in a particular node, through which an outside attack will only infect the node individually. Let Bernoulli RV's $V^{\mathscr{F}} \sim \mathrm{Ber}(\nu^{\mathscr{F}})$ and $V_d^{\mathscr{E}} \sim \mathrm{Ber}(\nu_d^{\mathscr{E}})$ with

$$\nu^{\mathscr{F}} := \mathbb{P}\big(V^{\mathscr{F}} = 1\big) \quad \text{and} \quad \nu_d^{\mathscr{E}} := \mathbb{P}\big(V_d^{\mathscr{E}} = 1\big), \quad d = 1, \ldots, n^{\mathscr{T}}, \tag{1}$$

indicate whether or not a common vulnerability among the fog nodes and the type d end nodes is harnessed by an outside attacker, respectively. The above discussions about network vulnerabilities lead to the following stochastic formulation for the outside attack RV associated with the ith fog node:

$$O_i^{\mathscr{F}} = (1 - V^{\mathscr{F}})Y_i^{\mathscr{F}} + V^{\mathscr{F}} Z^{\mathscr{F}},$$

where Bernoulli RV's $Y_i^{\mathscr{F}} \sim \mathrm{Ber}(\pi_i^{\mathscr{F}})$ and $Z^{\mathscr{F}} \sim \mathrm{Ber}(\pi^{\mathscr{F}*})$ indicate respectively whether or not an outside attack succeeds via an idiosyncratic vulnerability with probability $\pi_i^{\mathscr{F}}$ and a common vulnerability with probability $\pi^{\mathscr{F}*}$. In the probability notations above, the superscript "$*$" emphasizes that the compromise is caused by a common vulnerability. Similarly, the outside attack RV of the (d, i_d)th end node is modeled via

$$O_{d,i_d}^{\mathscr{E}} = (1 - V_d^{\mathscr{E}})Y_{d,i_d}^{\mathscr{E}} + V_d^{\mathscr{E}} Z_d^{\mathscr{E}},$$

where $Y_{d,i_d}^{\mathscr{E}} \sim \mathrm{Ber}(\pi_{d,i_d}^{\mathscr{E}})$ and $Z_d^{\mathscr{E}} \sim \mathrm{Ber}(\pi_d^{\mathscr{E}*})$ for $i_d = 1, \ldots, n_d^{\mathscr{E}}$ and $d = 1, \ldots, n^{\mathscr{T}}$.

Assumption 3.1. Because outside attacks are launched randomly, it is natural for us to assume $V^{\mathscr{F}}, V_d^{\mathscr{E}}, Y_i^{\mathscr{F}}, Y_{d,i_d}^{\mathscr{E}}, Z^{\mathscr{F}}$ and $Z_d^{\mathscr{F}}$ to be independent, where $i = 1, \ldots, n^{\mathscr{F}}$, $i_d = 1, \ldots, n_d^{\mathscr{E}}$ and $d = 1, \ldots, n^{\mathscr{T}}$.

For notational convenience, denote by $\boldsymbol{O} = \big(\boldsymbol{O}^{\mathscr{F}}, \boldsymbol{O}_1^{\mathscr{E}}, \ldots, \boldsymbol{O}_{n^{\mathscr{T}}}^{\mathscr{E}}\big)$ the set of outside attack RV's with

$$\boldsymbol{O}^{\mathscr{F}} = \big(O_1^{\mathscr{F}}, \ldots, O_{n^{\mathscr{F}}}^{\mathscr{F}}\big), \quad \boldsymbol{O}_d^{\mathscr{E}} = \big(O_{d,1}^{\mathscr{E}}, \ldots, O_{d,n_d^{\mathscr{E}}}^{\mathscr{E}}\big), \quad d = 1, \ldots, n^{\mathscr{T}}.$$

Speaking plainly, the above setting implies that if a common vulnerability is exploited, then all the nodes of the same type either get infected simultaneously if the attack succeeds, or all remain healthy if the attack fails. When $V^{\mathscr{F}} = 0$ (resp., $V_d^{\mathscr{E}} = 0$, $d \in \{1, \ldots, n^{\mathscr{T}}\}$), then the coordinates of $\boldsymbol{O}^{\mathscr{F}}$ (resp., $\boldsymbol{O}_d^{\mathscr{E}}$, $d \in \{1, \ldots, n^{\mathscr{T}}\}$) are assumed to be independent, since in this case, infections are caused individually by different attacks. Assumption 3.1 implies that $\boldsymbol{O}^{\mathscr{F}}$, $\boldsymbol{O}_1^{\mathscr{E}}$, …, $\boldsymbol{O}_{n^{\mathscr{T}}}^{\mathscr{E}}$ are mutually independent. Nevertheless, the outside attack RV's belonging to the same node type are generally dependent because of the presence of common vulnerabilities. To illustrate the dependencies, let $\boldsymbol{s}$ be a vector of appropriate dimension which contains elements equal to either zero or one. Then we have

$$\begin{aligned}\mathbb{P}\big(\boldsymbol{O}^{\mathscr{F}} = \boldsymbol{s} \,\big|\, V^{\mathscr{F}} = 1\big) &= \mathbb{1}_{\{\boldsymbol{s}=\boldsymbol{0}\}}\, \mathbb{P}\big(Z^{\mathscr{F}} = 0\big) + \mathbb{1}_{\{\boldsymbol{s}=\boldsymbol{1}\}}\, \mathbb{P}\big(Z^{\mathscr{F}} = 1\big) \\ &= \mathbb{1}_{\{\boldsymbol{s}=\boldsymbol{0}\}}\, (1 - \pi^{\mathscr{F}*}) + \mathbb{1}_{\{\boldsymbol{s}=\boldsymbol{1}\}}\, \pi^{\mathscr{F}*},\end{aligned}$$

where $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function, and

$$\mathbb{P}\big(\boldsymbol{O}^{\mathscr{F}} = \boldsymbol{s} \,\big|\, V^{\mathscr{F}} = 0\big) = \prod_{i=1}^{n^{\mathscr{F}}} \mathbb{P}\big(Y_i^{\mathscr{F}} = s_i\big) = \prod_{i=1}^{n^{\mathscr{F}}} \big(\pi_i^{\mathscr{F}}\big)^{s_i}\, (1 - \pi_i^{\mathscr{F}})^{1-s_i},$$

in which $s_i \in \{0, 1\}$ is the ith element of $\boldsymbol{s}$. Similarly, the following probability expressions hold for the type d end nodes:

$$\mathbb{P}\big(\boldsymbol{O}_d^{\mathscr{E}} = \boldsymbol{s} \,\big|\, V_d^{\mathscr{E}} = 1\big) = \mathbb{1}_{\{\boldsymbol{s}=\boldsymbol{0}\}}\, (1 - \pi_d^{\mathscr{E}*}) + \mathbb{1}_{\{\boldsymbol{s}=\boldsymbol{1}\}}\, \pi_d^{\mathscr{E}*}$$

and

$$\mathbb{P}\big(\boldsymbol{O}_d^{\mathscr{E}} = \boldsymbol{s} \,\big|\, V_d^{\mathscr{E}} = 0\big) = \prod_{i=1}^{n_d^{\mathscr{E}}} \big(\pi_{d,i}^{\mathscr{E}}\big)^{s_i}\, (1 - \pi_{d,i}^{\mathscr{E}})^{1-s_i}, \quad d \in \{1, \ldots, n^{\mathscr{T}}\}.$$

Next, we turn to the cybersecurity risks due to inside attacks which are launched by the existing compromised nodes through network connections. Here are the notations needed for catering the possible paths of risk contagions. For the ith fog node, denote by $I_{i\to j}^{\mathscr{F}} \in \{0, 1\}$ and $I_{i\to(d,i_d)}^{\mathscr{F}} \in \{0, 1\}$ the activation status of the link to the jth fog node and the (d, i_d)th end node, respectively, with value one means active, zero means inactive, $d \in \{1, \ldots, n^{\mathscr{T}}\}$, $i \neq j \in \{1, \ldots, n^{\mathscr{F}}\}$, and $i_d \in \{1, \ldots, n_d^{\mathscr{E}}\}$. If a link is active, then a compromised node can launch an inside attack to infect a healthy node via the link. The corresponding infection rates are denoted by

$$\begin{aligned}&\mathbb{P}\big(I_{i\to j}^{\mathscr{F}} = 1 \,|\, C_i^{\mathscr{F}} = 1\big) =: q_{i\to j}^{\mathscr{F}} \in [0, 1], \\ &\mathbb{P}\big(I_{i\to(d,i_d)}^{\mathscr{F}} = 1 \,|\, C_i^{\mathscr{F}} = 1\big) =: q_{i\to(d,i_d)}^{\mathscr{F}} \in [0, 1].\end{aligned}$$

Regarding 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 communications between end nodes and fog nodes, but not from end nodes to end nodes. Let $I_{(d,i_d)\to j}^{\mathscr{E}}$ represent the activation status of the link from the (d, i_d)th end node to the jth fog node, with

$$\mathbb{P}\Big(I_{(d,i_d)\to j}^{\mathscr{E}} = 1 \,\Big|\, C_{d,i_d}^{\mathscr{E}} = 1\Big) =: q_{(d,i_d)\to j}^{\mathscr{E}} \in [0, 1],$$

for $d \in \{1, \ldots, n^{\mathscr{T}}\}$, $i_d \in \{1, \ldots, n_d^{\mathscr{E}}\}$, $j \in \{1, \ldots, n^{\mathscr{F}}\}$. 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 Fig. 1, we have

$$q_{4\to j}^{\mathscr{F}} = \begin{cases} >0, & j = 7,\ 11; \\ = 0, & \text{otherwise}, \end{cases}$$

$$q_{4\to(d,i_d)}^{\mathscr{F}} = \begin{cases} >0, & (d, i_d) = (3, 1),\ (3, 2),\ (3, 3),\ (3, 4); \\ = 0, & \text{otherwise}. \end{cases}$$

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

$$q_{(1,2)\to j}^{\mathscr{E}} = \begin{cases} >0, & j = 1, 2; \\ = 0, & \text{otherwise}. \end{cases}$$

Assumption 3.2. Denote by

$\boldsymbol{I} = \big(\boldsymbol{I}^{\mathscr{F}}, \boldsymbol{I}_1^{\mathscr{E}}, \ldots, \boldsymbol{I}_{n^{\mathscr{T}}}^{\mathscr{E}}\big),$

where $\boldsymbol{I}^{\mathscr{F}} = (I_1^{\mathscr{F}}, \ldots, I_{n^{\mathscr{F}}}^{\mathscr{F}})$, $\boldsymbol{I}_d^{\mathscr{E}} = (I_{d,1}^{\mathscr{E}}, \ldots, I_{d,n_d^{\mathscr{E}}}^{\mathscr{E}})$, $d \in \{1, \ldots, n^{\mathscr{T}}\}$,

the set of all link status RV's for modeling the inside attacks. 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 practically reasonable to assume that the outside attack RV, $\boldsymbol{O}$, and inside attack RV, $\boldsymbol{I}$, are independent.

With the outside attack and inside attack RV's defined, we now set out to establish a system of state equations for modeling the compromise statuses of different components in fog networks. For $j = 1, \ldots, n^{\mathscr{F}}$, the state equation associated with the jth fog node is given by

$$C_j^{\mathscr{F}} = 1 - \underbrace{\big(1 - O_j^{\mathscr{F}}\big)}_{①} \underbrace{\prod_{i=1, i\neq j}^{n^{\mathscr{F}}} \big(1 - C_i^{\mathscr{F}}\, I_{i\to j}^{\mathscr{F}}\big)}_{②} \underbrace{\prod_{d=1}^{n^{\mathscr{T}}} \prod_{i_d=1}^{n_d^{\mathscr{E}}} \big(1 - C_{d,i_d}^{\mathscr{E},[j]}\, I_{(d,i_d)\to j}^{\mathscr{E}}\big)}_{③}, \tag{2}$$

Table 1
Descriptions of the different elements in state Eqs. (2)–(4).

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

where

$$C_{d,i_d}^{\mathscr{E},[j]} = 1 - \underbrace{\left(1 - O_{d,i_d}^{\mathscr{E}}\right)}_{①} \underbrace{\prod_{i=1, i\neq j}^{n^{\mathscr{F}}} \left(1 - C_i^{\mathscr{F}} I_{i\to(d,i_d)}^{\mathscr{F}}\right)}_{②} \tag{3}$$

is the state equation of the (d, i_d)th end node while assuming that the jth fog node is originally healthy (equivalently, excluding the jth fog node from the state equation). The state equation for the (d, j_d)th end node is

$$C_{d,j_d}^{\mathscr{E}} = 1 - \underbrace{\left(1 - O_{d,j_d}^{\mathscr{E}}\right)}_{①} \underbrace{\prod_{i=1}^{n^{\mathscr{F}}} \left(1 - C_i^{\mathscr{F}} I_{i\to(d,j_d)}^{\mathscr{F}}\right)}_{②}, \quad \text{for } d = 1, \ldots, n^{\mathscr{T}},\ j_d = 1, \ldots, n_d^{\mathscr{E}}. \tag{4}$$

Table 1 outlines the descriptions of the different elements in state Eqs. (2)–(4). A concise summary of the notation system used throughout this paper is provided in Appendix B.

Remark that state Eqs. (2)–(4) not only endogenize the stochastic compromise statuses of all the fog network's nodes, but also capture the intricate risk contagion process. The coordinates of the compromise status RV $\boldsymbol{C}$ are highly dependent. One origin of the dependence comes from the embedded nature of the state equations system, with the state equation underlying each compromise RV also depends on the compromise statuses of its neighboring nodes. Another origin is via the involved outside attack RV's, which are correlated among the same type of nodes because of the common vulnerabilities. Consequently, it is considerably challenging to evaluate the compromise probabilities underlying $\boldsymbol{C}$, and to the best of our knowledge, no explicit results can be established unless certain specific network structures are assumed. Generally, in order to calculate the compromise probabilities, numerical simulations must be adopted. However, the probability of at least one infected node can be computed explicitly via $1 - \mathbb{P}\left(\boldsymbol{C} = \boldsymbol{0}\right)$, where

$$\mathbb{P}\left(\boldsymbol{C} = \boldsymbol{0}\right) = \left[\nu^{\mathscr{F}}(1 - \pi^{\mathscr{F}*}) + (1 - \nu^{\mathscr{F}}) \prod_{i=1}^{n^{\mathscr{F}}} (1 - \pi_i^{\mathscr{F}})\right]$$
$$\prod_{d=1}^{n^{\mathscr{T}}} \left[\nu_d^{\mathscr{E}}(1 - \pi_d^{\mathscr{E}*}) + (1 - \nu_d^{\mathscr{E}}) \prod_{i_d=1}^{n_d^{\mathscr{E}}} (1 - \pi_{(d,i_d)}^{\mathscr{E}})\right]. \tag{5}$$

Note that the above formula for $\mathbb{P}(\boldsymbol{C} = \boldsymbol{0})$ is independent of the inside attack probabilities, which means that a network can remain healthy as long as no outside attacks succeed. Moreover, if the number of nodes in a network increases, then there are more multiplicative terms of probabilities involved in (5), and hence $\mathbb{P}(\boldsymbol{C} = \boldsymbol{0})$ becomes smaller. In other words, the more number of nodes contained in a network, the more likely that the network would be compromised, which is natural to expect.

The infection models established in this current section are rather abstract. In order to further exemplify the usefulness of the proposed methodology, the next section is devoted to the study of smart home system which corresponds to one of the most prevalent fog computing based IoT architectures. In this specific case, we manage to evaluate the compromise probabilities for all nodes explicitly.

4. Cybersecurity risks in smart home fog networks

Consider the cyber infrastructure of a smart home provider which consists of $n^{\mathscr{F}}$ individual users and $n^{\mathscr{T}}$ types of household kits (i.e., end nodes). 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. Moreover, the fog nodes among different users are interconnected through a control center maintained by the service provider. Fig. 2 illustrates the network structure underling a typical 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. However, the compromise statuses of the fog nodes may be still highly dependent due to the presences of common vulnerabilities as well as the mutual connections to the control center.

The theoretical groundwork laid down in Section 3 can be utilized to study the cybersecurity risks in the smart home network. In addition to the notations introduced in Section 3, so as to capture the cybersecurity risks associated with the control center, let us further define

- the compromise status RV for the central control, $C^{\mathscr{C}} \in \{0, 1\}$;
- the outside attack RV, $O^{\mathscr{C}} \in \{0, 1\}$, with probability $\mathbb{P}(O^{\mathscr{C}} = 1) = \omega^{\mathscr{C}}$;
- the inside attack RV, $I_{i\to\bullet}^{\mathscr{F}} \in \{0, 1\}$, which indicates the internal attack launched from the ith compromised fog node to the healthy central control with $\mathbb{P}\left(I_{i\to\bullet}^{\mathscr{F}} = 1\right) = q_{i\to\bullet}^{\mathscr{F}}$, and $I_{\bullet\to i}^{\mathscr{C}} \in \{0, 1\}$, which indicates the internal attack launched from compromised central control to the ith healthy fog node with $\mathbb{P}\left(I_{\bullet\to i}^{\mathscr{C}} = 1\right) = q_{\bullet\to i}^{\mathscr{C}}$.

As demonstrated in Fig. 2, 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 that particular connected fog node. For this reason, in the study of smart home system, it is more convenient for us to group the end devices based on the ownership of individual users. To reflect the aforementioned feature in the smart home system, for $d = 1, \ldots, n^{\mathscr{T}}$, $i = 1, \ldots, n^{\mathscr{F}}$, we further introduce

$$\mathbb{D}_{d,i} = \left\{j_d \in \{1, \ldots, n_d^{\mathscr{E}}\} : \ q_{i\to(d,j_d)}^{\mathscr{F}} > 0 \quad \text{or} \quad q_{(d,j_d)\to i}^{\mathscr{E}} > 0\right\}$$

to denote the set of type d end devices possessed by the ith smart home user.

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

$$C^{\mathscr{C}} = 1 - (1 - O^{\mathscr{C}}) \prod_{j=1}^{n^{\mathscr{F}}} (1 - C_j^{\mathscr{F},[\bullet]} \times I_{j\to\bullet}^{\mathscr{F}}), \tag{6}$$

where

$$C_j^{\mathscr{F},[\bullet]} = 1 - (1 - O_j^{\mathscr{F}}) \prod_{d=1}^{n^{\mathscr{T}}} \prod_{i_d\in\mathbb{D}_{d,j}} \left(1 - O_{d,i_d}^{\mathscr{E}} \times I_{(d,i_d)\to j}^{\mathscr{E}}\right), \qquad j = 1, \ldots, n^{\mathscr{F}}, \tag{7}$$

is the compromise status of the jth fog node with the central control excluded, or equivalently, assumed to be originally healthy. State Eqs. (2)–(4) can be adapted to study the fog network in the

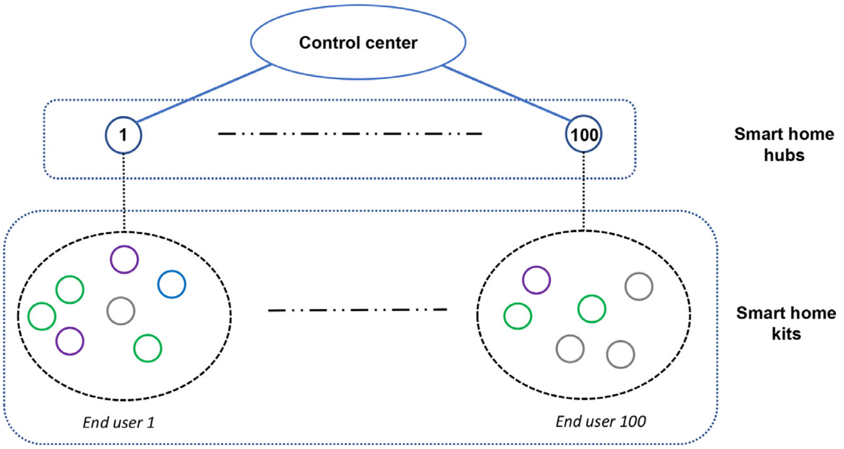

Fig. 2. An illustration of smart home network with the end devices are grouped according to the ownership of individual users, and different types of end devices are displayed in different colors.

smart home system. Namely, for the jth fog nodes, $j=1,\dots,n^{\mathscr{F}}$, we have

$$\begin{aligned} C_j^{\mathscr{F}} &= 1-(1-O_j^{\mathscr{F}})\,(1-C^{\mathscr{C},[j]}\times I^{\mathscr{C}}_{\bullet\to j})\prod_{d=1}^{n^{\mathscr{T}}}\prod_{i_d\in\mathbb{D}_{d,j}}\left(1-O^{\mathscr{E}}_{d,i_d}\times I^{\mathscr{E}}_{(d,i_d)\to j}\right) \\ &= 1-(1-C^{\mathscr{C},[j]}\times I^{\mathscr{C}}_{\bullet\to j})\left(1-C_j^{\mathscr{F},[\bullet]}\right), \end{aligned} \tag{8}$$

where

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

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

$$\begin{aligned} &C^{\mathscr{E}}_{d,j_d}=1-\left(1-O^{\mathscr{E}}_{d,j_d}\right)\left(1-C_i^{\mathscr{F}}\,I^{\mathscr{F}}_{i\to(d,j_d)}\right), \\ &\qquad j_d\in\mathbb{D}_{d,i}\ \text{with}\ d=1,\dots,n^{\mathscr{T}},\ i=1,\dots,n^{\mathscr{F}}. \end{aligned} \tag{9}$$

Thanks to the more specific network topology in the smart home system, 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 Fig. 2, and further, assume that the control center is highly secure with zero compromise probability, i.e., $p^{\mathscr{C}}:=\mathbb{P}(C^{\mathscr{C}}=1)=0$. For a given set of m fog nodes, indexed by $\Xi=(\xi_1,\dots,\xi_m)\subseteq\{1,\dots,n^{\mathscr{F}}\}$, their joint compromise probabilities can be computed via*

$$\tilde{p}^{\mathscr{F}}_{\Xi}:=\mathbb{P}\Big(\bigcap_{j\in\Xi}C_j^{\mathscr{F},[\bullet]}=1\Big)=1-\sum_{k=1}^{m}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi}h(\Xi_k), \tag{10}$$

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

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

with

$$\begin{aligned} g(d,\Xi_k)&=(1-\nu_d^{\mathscr{E}})\prod_{j\in\Xi_k}\prod_{i_d\in\mathbb{D}_{d,j}}(1-\pi^{\mathscr{E}}_{d,i_d}\,q^{\mathscr{E}}_{(d,i_d)\to j}) \\ &\quad+\nu_d^{\mathscr{E}}\left(1-\pi_d^{\mathscr{E}*}+\pi_d^{\mathscr{E}*}\prod_{j\in\Xi_k}\prod_{i_d\in\mathbb{D}_{d,j}}\left(1-q^{\mathscr{E}}_{(d,i_d)\to j}\right)\right) \end{aligned} \tag{11}$$

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

Remark 4.2. Formula (10) is reminiscent of the inclusion-exclusion principle in combinatorics. Specifically, it is observed that

$$\begin{aligned} &\mathbb{P}\Big(\bigcap_{j\in\Xi}C_j^{\mathscr{F},[\bullet]}=1\Big)=1-\mathbb{P}\Big(\bigcup_{j\in\Xi}C_j^{\mathscr{F},[\bullet]}=0\Big) \\ &=1-\sum_{k=1}^{m}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi}\mathbb{P}\Big(\bigcap_{j\in\Xi_k}C_j^{\mathscr{F},[\bullet]}=0\Big), \end{aligned}$$

where $\mathbb{P}\big(\bigcap_{j\in\Xi_k}C_j^{\mathscr{F},[\bullet]}=0\big)$ can be computed explicitly via $h(\Xi_k)$. Within the expression of $h(\Xi_k)$, the former block of calculations capture the external attacks while the latter cater 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.

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

Denote the unconditional outside attack probabilities by

$$\begin{aligned} \omega_j^{\mathscr{F}}&:=\mathbb{P}(O_j^{\mathscr{F}}=1)=\mathbb{P}(V^{\mathscr{F}}=0)\,\mathbb{P}(O_j^{\mathscr{F}}=1|V^{\mathscr{F}}=0) \\ &\quad+\mathbb{P}(V^{\mathscr{F}}=1)\,\mathbb{P}(O_j^{\mathscr{F}}=1|V^{\mathscr{F}}=1) \\ &=\pi_j^{\mathscr{F}}+\nu^{\mathscr{F}}\left(\pi^{\mathscr{F}*}-\pi_j^{\mathscr{F}}\right),\quad j=1,\dots,n^{\mathscr{F}}, \end{aligned}$$

and similarly,

$$\begin{aligned} &\omega^{\mathscr{E}}_{d,j_d}:=\mathbb{P}(O^{\mathscr{E}}_{d,j_d}=1)=\pi^{\mathscr{E}}_{d,j_d}+\nu_d^{\mathscr{E}}\left(\pi_d^{\mathscr{E}*}-\pi^{\mathscr{E}}_{d,j_d}\right), \\ &\qquad d=1,\dots,n^{\mathscr{T}},\ j_d=1,\dots,n_d^{\mathscr{E}}. \end{aligned}$$

Lemma 4.3. *Consider the smart home network as illustrated in Fig. 2. For a given set of m fog nodes, indexed by $\Xi=(\xi_1,\dots,\xi_m)\subseteq\{1,\dots,n^{\mathscr{F}}\}$, the following formula holds for the outside attack RV of the (d,j_d)th end node belonging to the ith smart home user:*

$$\mathbb{E}\Big[O^{\mathscr{E}}_{d,j_d}\prod_{j\in\Xi}C_j^{\mathscr{F},[\bullet]}\Big]=\omega^{\mathscr{E}}_{d,j_d}-\sum_{k=1}^{m}(-1)^{k-1}\sum_{\Xi_k\subseteq\Xi}u(j_d,\Xi_k)=:f(j_d,\Xi), \tag{12}$$

where

$$\begin{aligned} u(j_d,\Xi_k)&=h(\Xi_k)\times g(d,\Xi_k)^{-1} \\ &\times\Big[(1-\nu_d^{\mathscr{E}})\pi^{\mathscr{E}}_{d,j_d}(1-q^{\mathscr{E}}_{(d,j_d)\to i})\prod_{j\in\Xi_k}\prod_{l_d\in\mathbb{D}_{d,j},\,l_d\neq j_d}(1-\pi^{\mathscr{E}}_{d,l_d}\,q^{\mathscr{E}}_{(d,l_d)\to j}) \\ &\quad+\nu_d^{\mathscr{E}}\,\pi_d^{\mathscr{E}*}\prod_{j\in\Xi_k}\prod_{l_d\in\mathbb{D}_{d,j}}(1-q^{\mathscr{E}}_{(d,l_d)\to j})\Big]. \end{aligned}$$

Table 2
Outside attack probabilities for the smart home network in Example 4.5.

	Smart hubs	Type 1 home kits	Type 2 home kits
Idiosyncratic attack (i.e., $\pi_i^{\mathscr{F}}$, $\pi_{(d,i_d)}^{\mathscr{E}}$)	0.1	0.2	0.3
Systemic attack (i.e., $\pi^{\mathscr{F}*}$, $\pi_d^{\mathscr{E}*}$)	0.05	0.1	0.2
Common vulnerability (i.e., $\nu^{\mathscr{F}}$, $\nu_d^{\mathscr{E}}$)	0.1	0.1	0.2

In the sequel, we denote the compromise probability of the control center by $p^{\mathscr{C}} = \mathbb{P}(C^{\mathscr{C}} = 1)$, and the compromise probabilities of fog nodes and end nodes are respectively denoted by

$$\boldsymbol{p}^{\mathscr{F}} = (p_1^{\mathscr{F}}, \ldots, p_{n^{\mathscr{F}}}^{\mathscr{F}})^\top, \quad \text{with } p_i^{\mathscr{F}} := \mathbb{P}(C_i^{\mathscr{F}} = 1),\ i = 1, \ldots, n^{\mathscr{F}}, \tag{13}$$

and

$$\boldsymbol{p}_d^{\mathscr{E}} = \left(p_{d,1}^{\mathscr{E}}, \ldots, p_{d,n_d^{\mathscr{E}}}^{\mathscr{E}}\right)^\top,$$

$$\text{with } p_{d,i_d}^{\mathscr{E}} := \mathbb{P}(C_{d,i_d}^{\mathscr{E}} = 1),\ d = 1, \ldots, n^{\mathscr{T}},\ i_d = 1, \ldots, n_d^{\mathscr{E}}. \tag{14}$$

Theorem 4.4. *Consider the smart home network as illustrated in Fig. 2. The compromise probability for the control center can be computed via*

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

where Ξ_k is any subset of $\Xi^{\mathscr{F}} = (1, \ldots, n^{\mathscr{F}})$, and $\tilde{p}_{\Xi_k}^{\mathscr{F}}$ is the joint compromise probability for the fog nodes in Ξ_k which can be computed via (10), $k = 1, \ldots, n^{\mathscr{F}}$.

Moreover, the compromise probability for the ith fog node, $i = 1, \ldots, n^{\mathscr{F}}$, can be computed via

$$p_i^{\mathscr{F}} = 1 - (1 - \tilde{p}_i^{\mathscr{F}})(1 - q_{\bullet\to i}^{\mathscr{C}}) - q_{\bullet\to i}^{\mathscr{C}}(1 - \omega^{\mathscr{C}})$$

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

and the compromise probability for the j_dth type d end node which belongs to the ith smart home user, is given by

$$\begin{aligned} p_{d,j_d}^{\mathscr{E}} &= \omega_{d,j_d}^{\mathscr{E}} + p_i^{\mathscr{F}}\, q_{i\to(d,j_d)}^{\mathscr{F}} - \omega_{d,j_d}^{\mathscr{E}}\, q_{i\to(d,j_d)}^{\mathscr{F}} \\ &\quad + q_{i\to(d,j_d)}^{\mathscr{F}}\,(1 - q_{\bullet\to j}^{\mathscr{C}}) \times t_1 + q_{i\to(d,j_d)}^{\mathscr{F}}\, q_{\bullet\to i}^{\mathscr{C}}\,(1 - \omega^{\mathscr{C}}) \times t_2, \end{aligned}$$

for $d = 1, \ldots, n^{\mathscr{T}}$, $j_d \in \mathbb{D}_{d,i}$, $i = 1, \ldots, n^{\mathscr{F}}$, with

$$t_1 = \omega_{d,j_d}^{\mathscr{E}} - f(j_d, i),$$

and

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

Herein, the functions g and f are given in (11) and (12), respectively.

We illustrate the accuracy and efficiency of the explicit formulas in Theorem 4.4 via the following example. Although Theorem 4.4 can be applied to study smart home networks with an arbitrary number of users, for explanatory purposes, let us consider a simplified situation in which there are three users only.

Example 4.5. Suppose that there are two types of smart home end devices which are marked in different colors in Fig. 3. The outside attack probabilities associated with the fog nodes and end devices are summarized in Table 2. Based on the settings, 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 hubs are safer than the end devices against outside cyber attacks. We also assume that the inside attack probabilities among the fog nodes and end nodes are identical and equal to 0.25.

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

Table 3 compares the performance of the precise calculations proposed in Theorem 4.4 against the Monte Carlo simulation method for evaluating the compromise probabilities of the smart home infrastructure specified in Example 4.5. For each fixed sample size in the simulation study, the same experiment is repeated 1000 times in order to estimate the means and SD's of the compromise probability estimates. The computation time is reported at the end of Table 3.

What can be concluded from Table 3 are as follows. Firstly, the compromise probabilities computed using the explicit formulas from Theorem 4.4 coincide with the means of the simulated compromise probabilities (the minor discrepancies are only caused by the simulation errors). As the sample size n increases, the SD's of the simulation-based comprise probability estimators decay at a rate of approximately $\sqrt{n}$, which complies with the large-sample theory for the empirical means. Secondly, the explicit formulas compute the compromise probabilities in faster speeds than the simulation method. Thirdly, the orders of the compromise probabilities make very intuitive sense. For instance, the control center has the lowest compromise probability among all the network components since its outside and inside infection probabilities are assumed to be the lowest. Among the three smart home users, the first user's smart hub has the highest compromise probability because the user possesses the most 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 corresponding compromise probability is the lowest. Within the same type of end devices, say type 1, the (1,5)th end device has the lowest compromise probability because the third user has only two end devices, so inside 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, and the same type of end devices have the same outside and inside attack probabilities. The (1,1)th end device has the highest compromise probability among the type 1 end devices. This is because the first smart home user owns more type 2 end devices which have higher cybersecurity risks. Once they are infected, they may launch inside attacks to the (1,1)th device via the common connections to the smart home hub. A similar argument can be adopted to explain the orders of the compromise probabilities among the type 2 end devices.

Remark 4.6. The compromise probability formulas reported in Theorem 4.4 may be onerous to implement for an excessively large network. In particular, evaluating these formulas require us to consider the joint compromise probabilities for every subset of the fog nodes while assuming that the control center has zero compromise probability (see, Proposition 4.1). Notwithstanding, we still argue that the explicit approach is superior over the simulation approach in cyber modeling from the following perspectives. Firstly, for nodes having compromise probabilities closed to zero (which

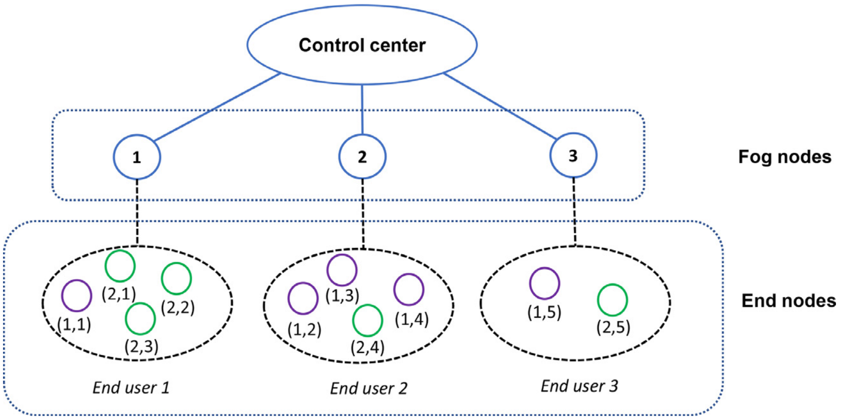

Fig. 3. The smart home network for Example 4.5.

Table 3
Comparisons between the Monte Carlo simulation estimates of the compromise probabilities and the precise calculations according to Theorem 4.4 for the smart home platform specified in Example 4.5. For each sample size $n \in \{1000, 5000, 25,000\}$, the simulation is repeated 1000 times to estimate the mean and standard deviation (SD) of the compromise probability 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
	$n = 1000$		$n = 5000$		$n = 25,000$		
	Mean	SD	Mean	SD	Mean	SD	
$p^{\mathscr{C}}$	0.046	0.007	0.047	0.003	0.048	0.001	0.048
$p_1^{\mathscr{F}}$	0.303	0.013	0.302	0.007	0.303	0.003	0.303
$p_2^{\mathscr{F}}$	0.272	0.014	0.274	0.006	0.272	0.002	0.272
$p_3^{\mathscr{F}}$	0.202	0.013	0.201	0.006	0.199	0.003	0.200
$p_{1,1}^{\mathscr{E}}$	0.244	0.013	0.244	0.006	0.244	0.003	0.244
$p_{1,2}^{\mathscr{E}}$	0.235	0.014	0.238	0.006	0.237	0.003	0.237
$p_{1,3}^{\mathscr{E}}$	0.237	0.014	0.237	0.006	0.237	0.002	0.237
$p_{1,4}^{\mathscr{E}}$	0.237	0.014	0.237	0.007	0.237	0.003	0.237
$p_{1,5}^{\mathscr{E}}$	0.223	0.015	0.223	0.006	0.222	0.003	0.222
$p_{2,1}^{\mathscr{E}}$	0.323	0.013	0.322	0.006	0.323	0.003	0.322
$p_{2,2}^{\mathscr{E}}$	0.324	0.014	0.322	0.006	0.322	0.003	0.322
$p_{2,3}^{\mathscr{E}}$	0.322	0.015	0.323	0.006	0.322	0.003	0.322
$p_{2,4}^{\mathscr{E}}$	0.319	0.014	0.319	0.007	0.319	0.003	0.319
$p_{2,5}^{\mathscr{E}}$	0.307	0.013	0.305	0.006	0.305	0.003	0.305
Average time per set of simulation	2.26		6.62		45.97		
Total time (seconds)	2263		6621		45,968		0.69

may still cause high losses if compromised, thus they should not be simply ignored from the analysis), it may require a large number of simulations in order to obtain reliable estimates of their compromise probabilities. The explicit approach is immune to sampling errors, and the results are precise. Secondly, in some applications, one may be only interested in the compromise probabilities for certain nodes. For example, an analyst wants to study the cybersecurity risks for a subgroup of smart homes, or test the sensitivity of the compromise probability of the control center in response to the changing outside attack probabilities of certain types of end devices. If the simulation approach is used, then we still have to conduct repeated simulations of the entire network even we are only interested in a subgroup of nodes. Differently, the explicit formulas in Theorem 4.4 can be directly adopted to study the subset of nodes of interest rather than analyzing the entire network.

For a smart home network, the probability of zero compromise can be computed in the same manner as (5), but it has an additional multiplication with $\mathbb{P}\left(O^{\mathscr{C}} = 0\right) = 1 - w^{\mathscr{C}}$, which is the probability of no outside attack occurred to the control center. The probability of at least one infected node for the smart home network considered in Example 4.5 is calculated to be 0.916. This result is reasonable, because the outside attack probabilities specified in Table 3 are not small and there are a considerable number of nodes contained in the network.

5. Applications to cybersecurity insurance pricing

Our study hitherto focuses on the frequencies of compromise events in a given fog network. That is, during a unit time period (e.g., one month/quarter/year), the compromise RV, $\boldsymbol{C}$, indicates whether or not certain nodes are compromised.

In the context of insurance pricing, the financial losses caused by the compromised nodes are of central interest. To this purpose, we resort to the frequency-severity approach which has evolved as an industry standard for pricing insurance risks generally, and cybersecurity risks particularly (see, e.g., Jevtić & Lanchier, 2020; Xu & Hua, 2019). To be specific, let $X^{\mathscr{C}} > 0$, $X_i^{\mathscr{F}} > 0$, and $X_{d,i_d}^{\mathscr{E}} > 0$ represent the financial losses caused by an infection of the control center, the ith fog node, and the (d, i_d)th end node, respectively. It is assumed that these severity RV's are mutually independent, and they are also independent of the compromise status RV $\boldsymbol{C}$. The aggregate loss for the entire smart home platform can be evaluated via

$$L = C^{\mathscr{C}} X^{\mathscr{C}} + \sum_{i=1}^{n^{\mathscr{F}}} C_i^{\mathscr{F}} X_i^{\mathscr{F}} + \sum_{d=1}^{n^{\mathscr{F}}} \sum_{i_d=1}^{n_d^{\mathscr{E}}} C_{d,i_d}^{\mathscr{E}} X_{d,i_d}^{\mathscr{E}}, \tag{15}$$

Table 4
The baseline parameters for the severity models of the different network elements with the summary statistics for the associated distributions.

	Parameter	Mean	SD	GMD	Percentiles		
					25%	50%	75%
Control center	$(\alpha, \beta) = (5 \times 10^4, 11)$	5000	5528	5238	1325	3252	6716
Smart hubs	$(\mu, \sigma^2) = (4.26, 0.83^2)$	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

in which the three blocks of calculations in the above formula cater the losses due to the compromised control center, fog nodes, and end nodes, respectively. It is noteworthy that although our discussion herein is inspired by the smart home application, loss model (15) does not rely on a specific network topology, so in principle, it can be used to study any fog networks.

To price the cybersecurity risks, prevalent actuarial pricing principles can be applied. Examples include

$$\text{Expectation principle:} \quad \varrho_1(L) = (1+\theta)\,\mathbb{E}[L]; \tag{16}$$

$$\text{Standard deviation principle:} \quad \varrho_2(L) = \mathbb{E}[L] + \theta\sqrt{\mathrm{Var}(L)}; \tag{17}$$

$$\text{Gini mean difference principle:} \quad \varrho_3(L) = \mathbb{E}[L] + \theta\,\mathrm{GMD}(L). \tag{18}$$

In the above pricing principles, $\theta > 0$ is the loading parameter which reflects the risk preferences of the actuaries. Meanwhile, let L_1 and L_2 be a pair of independent copies of L, the Gini mean difference (GMD):

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

is a statistical measure of variability, well known to be a robust alternative to the SD (see more detailed discussions in, e.g., Furman, Kye, & Su, 2019; Furman, Wang, & Zitikis, 2017; Yitzhaki et al., 2003).

In what follows, let us consider the cyber insurance pricing for the smart home system studied 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. As a result, we assume

$$X^{\mathscr{C}} \sim \mathrm{Lomax}(\alpha, \beta), \qquad \alpha \in \mathbb{R}_+,\ \beta \in \mathbb{R}_+,$$

which is a heavy-tailed distribution, and

$$X_i^{\mathscr{F}} \sim \mathrm{LN}(\mu, \sigma^2), \qquad \mu \in \mathbb{R},\ \sigma \in \mathbb{R}_+,\ i = 1, 2, 3,$$

which is a moderately heavy-tailed distribution, and

$$X_{d,i_d}^{\mathscr{E}} \sim \mathrm{Exp}(\lambda_d), \qquad \lambda_d \in \mathbb{R}_+,\ d = 1, 2,$$

which is a light-tailed distribution.

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

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

where $\mu^{\mathscr{C}} = \mathbb{E}\big[X^{\mathscr{C}}\big]$, $\mu_i^{\mathscr{F}} = \mathbb{E}\big[X_i^{\mathscr{F}}\big]$, $\mu_{d,i_d}^{\mathscr{E}} = \mathbb{E}\big[X_{d,i_d}^{\mathscr{E}}\big]$, $i = 1, \ldots, n^{\mathscr{F}}$, $i_d = 1, \ldots, n_d^{\mathscr{E}}$, $d = 1, \ldots, n^{\mathscr{T}}$, and the compromise probabilities $p^{\mathscr{C}}$, $p_i^{\mathscr{F}}$ and $p_{d,i_d}^{\mathscr{E}}$ can be computed explicitly according to Theorem 4.4. However, if the SD principle or the GMD principle is used, then the variance and the GMD of the aggregate loss RV L can not be computed explicitly. Numerical simulation is adopted in these cases to price the cybersecurity risks.

In the succeeding numerical study, we use the attack probabilities specified in Example 4.5 as the baseline parameters for modeling the compromise frequencies. The baseline parameters for the severity distributions are summarized in Table 4. These parameters choices reflect the followings. Firstly, the control center has the highest average cybersecurity loss. Secondly, the smart home hubs among different users are usually similar, so the loss distributions for the fog nodes are assumed to be identical. Thirdly, as specified earlier in the set-up of Example 4.5, the type 2 home kits are more vulnerable to cyber attacks than the type 1 home kits, which is because the associated cybersecurity losses are lower.

Under the aforementioned baseline parameters, Fig. 4 displays the histograms of the shifted log transforms of the cybersecurity losses for the different components in the smart home system. The loss observations for constructing the histograms are generated by 10^5 times of simulations. As shown, the end devices have the most frequent cybersecurity losses because their security configurations are low and so the associated compromise probabilities are high. The control center has the lowest frequency of cybersecurity events, but once they occur, the financial losses can be very severe. As a consequence, the aggregate loss of the entire network features a highly right-skewed distribution, shedding light on the importance of having a fine risk management program in place for the smart home users as well as the cyber insurance providers in order to administer the tail risks.

A sensitivity analysis is conducted so as to identify the key risk drivers of the insurance prices. In each scenario, we shock a set of similar parameters by 50% while keeping the other baseline parameters unchanged, then we assess the variations in the cyber insurance prices. Table 5 depicts the sensitivities of the cyber insurance prices according to different types of cyber attacks. Among the three actuarial pricing principles under investigation, we find that the SD principle yields the highest premium while the expectation principle yields the lowest. The orders are intuitive. Namely, note that the three premium principles (16)–(18) only differ up to an extra term added to the expectation. As shown earlier in Fig. 4, the loss distribution of the smart home network is highly right-skewed. Thereby, it is natural to expect that the expectation principle yields a lower premium than the SD principle and GMD principle. Moreover, the SD penalizes large variations in the loss distribution harsher than the GMD, thus the premium determined by the SD principle is higher than that by the GMD principle. In practice, we suggest insurance companies who are strongly averse to the variation of losses to use the SD principle for conservative ratemaking. Interestingly, we note that the SD principle is the most robust principle in response to the changing attack parameters. Thereby, our suggestion on the use of SD principle also makes sense from the robustness standpoint. On the contrary, the expectation principle is the most sensitive principle. The orders of sensitivities reveal that the changes in the attack parameters would more adversely impact the mean of the loss distribution than the SD.

Here are some additional findings obtained from the sensitivity analysis. Firstly, the insurance prices become lower in the downside cases of the sensitivity analysis, because lower probabilities of inside and outside attacks correspond to a safer network. Secondly, among different attack types, the inside attack parameters have the

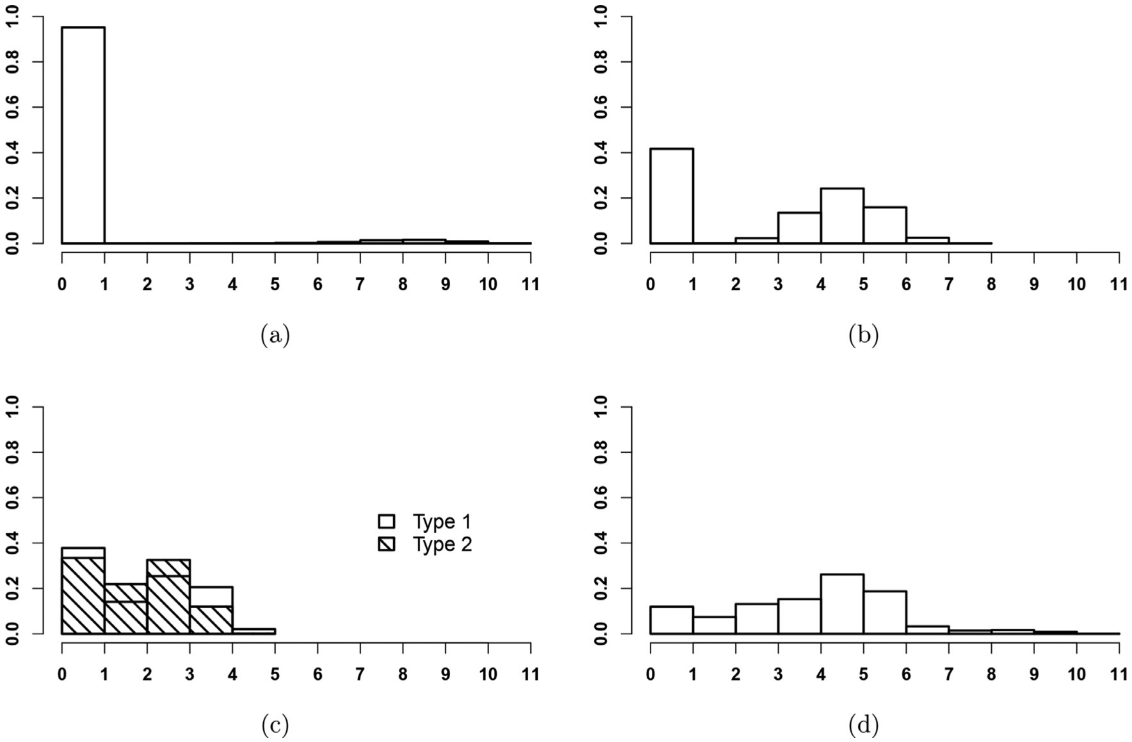

Fig. 4. Histograms of the shifted log transforms (i.e., $f(l) = \log(l+1)$, $l \geq 0$) of the financial losses caused by the compromises of the control center (top-left), fog nodes (top-right), and end nodes (bottom-left), as well as the aggregate loss for the entire network (bottom-right).

Table 5
Sensitivity analysis for the cyber insurance prices in response to the changes in the frequencies of different cyber attack types. 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 Eqs. (16)–(18) with $\theta = 0.1$. The percentage of change in the insurance price compared with the baseline price is reported in the parentheses 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%)
$\left(\pi_i^{\mathscr{F}}, \pi_{(d,i_d)}^{\mathscr{E}}\right)$	Up	458.778 (29%)	586.190 (22%)	484.901 (28%)
Systemic attacks	Down	346.962 (−2%)	469.643 (−2%)	369.654 (−2%)
$\left(\pi^{\mathscr{F}*}, \pi_d^{\mathscr{E}*}\right)$	Up	387.636 (9%)	527.621 (10%)	412.887 (9%)
Common vulnerabilities	Down	351.698 (−1%)	472.729 (−1%)	374.038 (−1%)
$\left(\nu^{\mathscr{F}}, \nu_d^{\mathscr{E}}\right)$	Up	368.413 (4%)	503.146 (5%)	392.713 (4%)
Inside attacks	Down	202.418 (−43%)	292.745 (−39%)	215.057 (−43%)
$\left(q_{j\to\bullet}^{\mathscr{F}}, q_{\bullet\to j}^{\mathscr{C}}, q_{j\to(d,i)}^{\mathscr{F}}, q_{(d,i)\to j}^{\mathscr{E}}\right)$	Up	565.294 (59%)	713.606 (49%)	601.435 (59%)

most substantial influences on the insurance prices, which is intuitive. Namely, in an unsecured network with high inside attack successful rates, even a single outside attack can be propagated across the entire network and infect many other devices. Thirdly, 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 determining the compromise probabilities. In another unreported analysis where the common vulnerabilities probabilities are assume to be high, then we observe that the aforementioned order is reversed.

Next, the insurance price sensitivities in accordance with the changes in the compromise frequencies of different network elements are examined. Based on Table 6, we observe that the fog nodes have the most noticeable impacts on the insurance prices, with an 50% increase in the compromise probabilities rises the insurance prices by about 25%. This is probably because the fog nodes have relatively high compromise frequencies 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 change in the control center's compromise probability.

Finally, we are interested in understanding the impacts of one weak node on the security of the entire smart home network. Specifically, we separately shock the idiosyncratic attack probabilities of fog node 1 and end node (1,1) from the baseline values specified in Table 2 to 0.8, and then we study the resulting impacts on the network nodes' compromise probabilities and the associated insurance prices. From Table 7, we observe that in both the shocked scenarios, the presence of a weak node leads to an increase in the compromise probability of the control center. Compared with the case of one weak end node, a weak fog node has a greater impact on the control center because for a weak fog node, it can directly communicate with the control center but not for a weak end node. A weak fog node also has more significant impacts on the compromise probabilities of the original end nodes within the same smart house, because an inside attack to an end note must be launched through a fog node. Furthermore, a weak fog node or end node would have significant influences on the other nodes within the same house, while the impacts on the

Table 6
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 5.

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%)
$\left(\omega^{\mathscr{C}}\right)$	Up	369.010 (4%)	492.982 (3%)	393.138 (4%)
Fog nodes	Down	310.522 (−13%)	432.914 (−10%)	331.293 (−12%)
$(\pi_i^{\mathscr{F}}, \pi^{\mathscr{F}*}, \nu^{\mathscr{F}})$	Up	439.996 (24%)	585.632 (22%)	468.503 (24%)
Type 1 end nodes	Down	317.642 (−11%)	434.87 (−9%)	339.120 (−10%)
$(\pi_{(1,i_1)}^{\mathscr{E}}, \pi_1^{\mathscr{E}*}, \nu_1^{\mathscr{E}})$	Up	403.906 (14%)	540.448 (13%)	429.029 (13%)
Type 2 end nodes	Down	317.737 (−11%)	458.763 (−9%)	339.587 (−11%)
$(\pi_{(2,i_2)}^{\mathscr{E}}, \pi_2^{\mathscr{E}*}, \nu_2^{\mathscr{E}})$	Up	406.4879 (14%)	537.1767 (12%)	431.4025 (14%)

Table 7
The compromise probabilities for the smart home network of interest with the outside attack probability for a fog node or a end node shocked.

	Normal case	Weak node	
		End node (1,1)	Fog node 1
$p^{\mathscr{C}}$	0.048	0.052 (10%)	0.071 (49%)
$p_1^{\mathscr{F}}$	0.303	0.402 (33%)	0.788 (160%)
$p_2^{\mathscr{F}}$	0.272	0.273 (0%)	0.273 (0%)
$p_3^{\mathscr{F}}$	0.200	0.200 (0%)	0.201 (0%)
$p_{1,1}^{\mathscr{E}}$	0.244	0.748 (206%)	0.347 (42%)
$p_{1,2}^{\mathscr{E}}$	0.237	0.237 (0%)	0.237 (0%)
$p_{1,3}^{\mathscr{E}}$	0.237	0.237 (0%)	0.237 (0%)
$p_{1,4}^{\mathscr{E}}$	0.237	0.237 (0%)	0.237 (0%)
$p_{1,5}^{\mathscr{E}}$	0.222	0.222 (0%)	0.223 (0%)
$p_{2,1}^{\mathscr{E}}$	0.322	0.342 (6%)	0.418 (30%)
$p_{2,2}^{\mathscr{E}}$	0.322	0.342 (6%)	0.418 (30%)
$p_{2,3}^{\mathscr{E}}$	0.322	0.342 (6%)	0.418 (30%)
$p_{2,4}^{\mathscr{E}}$	0.319	0.319 (0%)	0.319 (0%)
$p_{2,5}^{\mathscr{E}}$	0.305	0.305 (0%)	0.305 (0%)

other smart houses are much smaller because of the strict security measure implemented in the control center. For example, when the weak node is end node (1,1), then the compromise probability of fog node 1 increases from 0.303 to 0.402, while the compromise probabilities of the other fog nodes remain nearly unchanged. Meanwhile, the compromise probabilities for end nodes (1,1), (2,1), (2,2), (2,3) increase, but little changes in the compromise probabilities of the other end nodes were observed.

Table 8 summarizes the changes in insurance prices under the two shocked scenarios. As shown, the insurance prices become higher in both cases, with a more substantial increase observed in the weak fog node case. This is natural to expect, because as mentioned earlier, a weak fog node has more significant impacts on the compromise probabilities of the control center and the original end nodes than one weak end node. In the meantime, the financial losses due to compromised control center and fog nodes are likely to be high.

6. Approximation method for efficient identification of weak nodes

In insurance ratemaking, actuaries consider a snapshot of the cybersecurity risks at a point before the issuance of an insurance policy. After the insurance policy is issued, it is possible that a secure node may turn into a weak node due to software updates, discoveries of hidden vulnerability entries, changes of functions etc.[4] As shown at the end of last section, the presence of weak nodes may impose substantial impacts on the security of the entire network. Given today's rapidly evolving technology landscape, it is important for insurance companies and the fog network users to continuously monitor the emergence of weak nodes and properly manage the cybersecurity threats associated with them. Thereby, an efficient method for weak nodes identification is natural called upon. Ideally, the weak nodes identification method should (a) work for any type of fog networks; (b) be efficient to implement so that weak nodes can be monitored in real-time; (c) well capture the ranks of nodes in terms of compromise probabilities; (d) provide a range of compromise probability estimation to support decisions on further actions.

The interval approximation method we aim to put forth in this current section meet all the desirable criteria listed above. The derivation mainly hinges on the classical notion of positive association of RV's.

Definition 6.1. 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

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

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 the compromise statuses, inside attacks, and outside attacks, respectively. Next, we show that the aforementioned RV's are positively associated. The succeeding lemma is of auxiliary importance.

Lemma 6.2 (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 $\big(f_1(\boldsymbol{X}), \dots, f_m(\boldsymbol{X})\big)$ is also positively associated.*

Proposition 6.3. *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.*

We are now in the position to spell out the interval approximations for the compromise probabilities $\boldsymbol{p}^{\mathscr{F}}$ and $\boldsymbol{p}_d^{\mathscr{E}}$, $d = 1, \dots, n^{\mathscr{T}}$.

The following notations are needed to define the lower bounds. Let $\boldsymbol{l}^{\mathscr{F}} = (l_1^{\mathscr{F}}, \dots, l_{n^{\mathscr{F}}}^{\mathscr{F}})^\top$ with elements

$$l_j^{\mathscr{F}} = \max\Bigg(\underbrace{\omega_j^{\mathscr{F}},}_{①} \underbrace{\bigvee_{i=1, i\neq j}^{n^{\mathscr{F}}} \beta_i q_{i\to j}^{\mathscr{F}},}_{②} \underbrace{\bigvee_{d=1}^{n^{\mathscr{T}}} \bigvee_{i_d=1}^{n_d^{\mathscr{E}}} \max\Bigg(\omega_{d,i_d}^{\mathscr{E}}, \bigvee_{i=1, i\neq j}^{n^{\mathscr{F}}} \beta_i\, q_{i\to(d,i_d)}^{\mathscr{F}}\Bigg) q_{(d,i_d)\to j}^{\mathscr{E}}}_{③}\Bigg), \tag{19}$$

where

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

$$j = 1, \dots, n^{\mathscr{F}},$$

[4] Note that the secure versus weak node we mentioned herein is a relative concept. Typically, a network company would have a security enhancement budget which is allocated to the weakest set of nodes.

Table 8
The cyber insurance prices for the smart home network of interest with the outside attack probability for a fog node or a end node shocked.

		ϱ_1	ϱ_2	ϱ_3
Baseline		355.388	479.845	378.589
Weak	End node (1,1)	400.726 (13%)	526.221 (10%)	424.967 (12%)
node	Fog node 1	556.778 (57%)	708.871 (48%)	587.634 (55%)

can be viewed as a finite-hop approximation of $p_i^{\mathscr{F}}$ in which an infection due to an outside attack can only advance at most two steps forward. For $d = 1, \dots, n^{\mathscr{F}}$, $j_d = 1, \dots, n_d^{\mathscr{E}}$, define $\boldsymbol{l}_d^{\mathscr{E}} = (l_{d,1}^{\mathscr{E}}, \dots, l_{d,n_d^{\mathscr{E}}}^{\mathscr{E}})^\top$ with elements

$$l^{\mathscr{E}}_{d,j_d} = \max\Bigg(\underbrace{\omega^{\mathscr{E}}_{d,j_d}}_{\text{①}}, \underbrace{\bigvee_{i=1}^{n^{\mathscr{F}}} l_i^{\mathscr{F}}\, q^{\mathscr{F}}_{i\to(d,j_d)}}_{\text{②}}\Bigg), \tag{20}$$

where $l_j^{\mathscr{F}}$ is specified as per Eq. (19). To facilitate the reading, the definitions of upper bounds $\boldsymbol{u}^{\mathscr{F}}$ and $\boldsymbol{u}_d^{\mathscr{E}}$, $d = 1, \dots, n^{\mathscr{F}}$, are relegated to Appendix B.

Theorem 6.4. *Consider a fog network described as per Section 3, and suppose that Assumptions 3.1 and 3.2 hold. Then the compromise probability vectors satisfy the following inequalities:*

$$\boldsymbol{l}^{\mathscr{F}} \le \boldsymbol{p}^{\mathscr{F}} \le \boldsymbol{u}^{\mathscr{F}}$$

and

$$\boldsymbol{l}_d^{\mathscr{E}} \le \boldsymbol{p}_d^{\mathscr{E}} \le \boldsymbol{u}_d^{\mathscr{E}}, \quad \text{for } d = 1, \dots, n^{\mathscr{F}}.$$

Some remarks about Theorem 6.4 are worth mentioning herein. Firstly, the compromise probabilities' lower bounds and upper bounds only contain simple algebraic operators, thus they can be calculated conveniently. Secondly, the compromise probabilities' lower bounds are derived by applying the co-monotonic approximation method (Dhaene, Denuit, Goovaerts, Kaas, & Vyncke, 2002a; 2002b) on the risk factors that determine the compromise probabilities, while in contrast, independent approximation is used to obtain the upper bounds. Thirdly, the compromise probabilities' lower bounds can be interpreted intuitively. We again refer to Table 1 for the descriptions of elements contained in Eqs. (19) and (20). What Eq. (19) tells is that, 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 the aforementioned causes. Specifically, let events A_i, $i = 1, \dots, n$, represent the causes of compromise for a given node, then we have the compromise probability $\mathbb{P}(\bigcup_{i=1}^n A_i) = 1 - \mathbb{P}(\bigcap_{i=1}^n A_i^c) \ge 1 - \bigwedge_{i=1}^n \mathbb{P}(A_i^c) = \bigvee_{i=1}^n \mathbb{P}(A_i)$.

The following hypothetical example demonstrates the usefulness of the proposed interval approximation method in an illuminated manner. For ease of exposition, it is our intention to keep the set-up of the example to be as simple as possible.

Example 6.5. Consider a fog network as per Fig. 5, in which there are four fog nodes and six end devices. Two of the end devices are type 1 end nodes, and the others are type 2 end nodes. Among the four fog nodes, assume the idiosyncratic and systemic outside attack probabilities to be $\pi_i^{\mathscr{F}} = 0.01$, $i \in \{1, \dots, 4\}$, and $\pi^{\mathscr{F}*} = 0.05$, respectively. Among the end nodes, because their cybersecurity configuration is typically weaker, outside attacks are more likely to occur compared with the fog nodes. So we set $\pi_{1,i_1}^{\mathscr{E}} = 0.2$, $i_1 = \{1, 2\}$, and $\pi_1^{\mathscr{E}*} = 0.25$ for the type 1 end nodes, and $\pi_{2,i_2}^{\mathscr{E}} = 0.2$, $i_2 \in \{1, \dots, 4\}$, and $\pi_2^{\mathscr{E}*} = 0.3$ for the type 2 end nodes. The inside attack probabilities $q_{i\to j}^{\mathscr{F}}$, $q_{i\to(d,i_d)}^{\mathscr{F}}$, $q_{(d,i_d)\to j}^{\mathscr{E}}$ are assumed to be identical and equal to $q \in \{0.1, 0.25, 0.4\}$. In a similar vein, the common vulnerability probabilities are set to be $\nu^{\mathscr{F}} = \nu_1^{\mathscr{E}} = \nu_2^{\mathscr{E}} = \nu \in \{0.5, 0.6, 0.7\}$.

Table 9 depicts the approximated compromise probabilities using Theorem 6.4 for the fog network specified in Example 6.5. The true compromise probabilities are computed by numerical simulations. Here are how the numerical results should be interpreted. Firstly, because the outside attack probabilities of the end nodes are higher than that of the fog nodes, the fog nodes have lower compromised probabilities than the end nodes. Due to a similar reason, the type 1 end nodes have lower compromise probabilities compared to the type 2 end nodes. Among the four fog nodes, since their outside attack probabilities are assumed to be identical, the orders of their compromise probabilities are driven by the frequencies of inside attacks launched from the compromised end nodes. Based on the construction of the fog network, it is natural to have that

$$p_4^{\mathscr{F}} \overset{(1)}{\le} p_1^{\mathscr{F}} \overset{(2)}{\le} p_3^{\mathscr{F}} \overset{(3)}{\le} p_2^{\mathscr{F}},$$

where

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

Both the lower and upper bounds of the interval approximations capture the aforementioned orders.

Secondly, varying the inside attack probabilities among $q \in \{0.1, 0.25, 0.4\}$ shows that if the fog network has a higher security configuration and so the internal risk propagation is less likely to occur, then the compromise probabilities are also lower across all the nodes. In this case, since the cybersecurity risks are mainly caused by outside attacks which are well captured by the lower and upper bound formulas, the proposed interval method provides very good estimates of the true compromise probabilities. However, as the inside attack probabilities increase, the effect of network dependence becomes more significant. The true network dependence is harder to be captured by the independent approximation or co-monotonic approximation, thus the performance of the interval method decays.

Thirdly, vary the probabilities of common vulnerabilities among $\nu \in \{0.5, 0.6, 0.7\}$, we observe that the final compromise probabilities may become lower or higher as ν increases. This is caused by the complexity involved in the calculations of compromise probabilities, which can be viewed as an application of Bayes rule on two conditional compromise probabilities based on whether or not common vulnerabilities occur. Depending on the order of the two conditional compromise probabilities, the increment of ν may pose different directions of impacts to the unconditional compromise probabilities. The proposed interval approximation method may

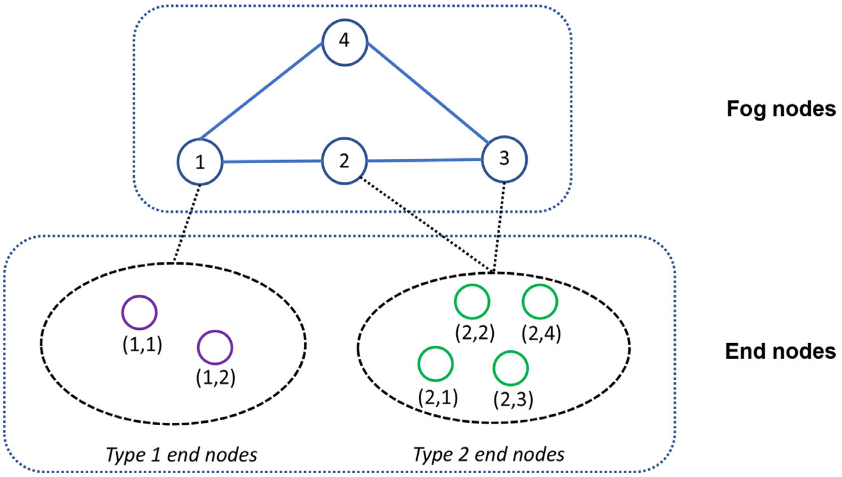

Fig. 5. The fog network for Example 6.5.

Table 9
The interval approximations and simulation-based calculations of the compromise probabilities for the fog network considered in Example 6.5, with the corresponding ranks reported in the parentheses after the probability estimates. In this table, the ranks of the lower/upper bounds are determined based on five decimal places of the estimated probabilities. The root mean square error and Pearson's correlation between the ranks of the simulated probabilities and the ranks of lower/upper bounds are denoted by "RMSE" and "Corr", respectively.

		$\nu = 0.5$			$\nu = 0.6$			$\nu = 0.7$		
		Lower	Sim.	Upper	Lower	Sim.	Upper	Lower	Sim.	Upper
$q=$	$p_1^{\mathscr{F}}$	0.03(1)	0.08(2)	0.09(2)	0.03(1)	0.09(2)	0.10(2)	0.04(1)	0.09(2)	0.11(2)
0.1	$p_2^{\mathscr{F}}$	0.03(1)	0.13(4)	0.17(4)	0.03(1)	0.14(4)	0.18(4)	0.04(1)	0.15(4)	0.19(4)
	$p_3^{\mathscr{F}}$	0.03(1)	0.12(3)	0.17(3)	0.03(1)	0.13(3)	0.18(3)	0.04(1)	0.14(3)	0.18(3)
	$p_4^{\mathscr{F}}$	0.03(1)	0.04(1)	0.06(1)	0.03(1)	0.05(1)	0.06(1)	0.04(1)	0.06(1)	0.07(1)
	$p_{1,1}^{\mathscr{E}}$	0.23(5)	0.23(5)	0.23(5)	0.23(5)	0.24(5)	0.24(5)	0.24(5)	0.24(5)	0.24(5)
	$p_{1,2}^{\mathscr{E}}$	0.23(5)	0.23(5)	0.23(5)	0.23(5)	0.24(5)	0.24(5)	0.24(5)	0.24(5)	0.24(5)
	$p_{2,1}^{\mathscr{E}}$	0.25(7)	0.26(7)	0.28(7)	0.26(7)	0.27(7)	0.29(7)	0.27(7)	0.27(7)	0.30(7)
	$p_{2,2}^{\mathscr{E}}$	0.25(7)	0.26(7)	0.28(7)	0.26(7)	0.27(7)	0.29(7)	0.27(7)	0.27(7)	0.30(7)
	$p_{2,3}^{\mathscr{E}}$	0.25(7)	0.26(7)	0.28(7)	0.26(7)	0.27(7)	0.29(7)	0.27(7)	0.27(7)	0.30(7)
	$p_{2,4}^{\mathscr{E}}$	0.25(7)	0.26(7)	0.28(7)	0.26(7)	0.27(7)	0.29(7)	0.27(7)	0.27(7)	0.30(7)
	RMSE	1.18		0	1.18		0	1.18		0
	Corr	0.94		1	0.94		1	0.94		1
$q=$	$p_1^{\mathscr{F}}$	0.06(2)	0.19(2)	0.29(4)	0.06(2)	0.20(2)	0.30(4)	0.06(2)	0.20(2)	0.31(4)
0.25	$p_2^{\mathscr{F}}$	0.06(3)	0.27(6)	0.51(10)	0.07(3)	0.28(6)	0.52(10)	0.07(3)	0.28(6)	0.53(10)
	$p_3^{\mathscr{F}}$	0.06(3)	0.26(5)	0.50(9)	0.07(3)	0.27(5)	0.51(9)	0.07(3)	0.27(5)	0.52(9)
	$p_4^{\mathscr{F}}$	0.03(1)	0.12(1)	0.22(1)	0.03(1)	0.13(1)	0.23(1)	0.04(1)	0.13(1)	0.24(1)
	$p_{1,1}^{\mathscr{E}}$	0.23(5)	0.25(3)	0.28(2)	0.23(5)	0.26(3)	0.29(2)	0.24(5)	0.26(3)	0.29(2)
	$p_{1,2}^{\mathscr{E}}$	0.23(5)	0.25(3)	0.28(2)	0.23(5)	0.26(3)	0.29(2)	0.24(5)	0.26(3)	0.29(2)
	$p_{2,1}^{\mathscr{E}}$	0.25(7)	0.30(7)	0.43(5)	0.26(7)	0.31(7)	0.44(5)	0.27(7)	0.31(7)	0.45(5)
	$p_{2,2}^{\mathscr{E}}$	0.25(7)	0.30(7)	0.43(5)	0.26(7)	0.31(7)	0.44(5)	0.27(7)	0.31(7)	0.45(5)
	$p_{2,3}^{\mathscr{E}}$	0.25(7)	0.30(7)	0.43(5)	0.26(7)	0.31(7)	0.44(5)	0.27(7)	0.31(7)	0.45(5)
	$p_{2,4}^{\mathscr{E}}$	0.25(7)	0.30(7)	0.43(5)	0.26(7)	0.31(7)	0.44(5)	0.27(7)	0.31(7)	0.45(5)
	RMSE	1.45		2.32	1.45		2.32	1.45		2.32
	Corr	0.79		0.58	0.79		0.58	0.79		0.58
$q=$	$p_1^{\mathscr{F}}$	0.09(2)	0.32(4)	0.67(4)	0.09(2)	0.32(4)	0.68(4)	0.09(2)	0.32(4)	0.68(4)
0.4	$p_2^{\mathscr{F}}$	0.10(3)	0.40(10)	0.86(10)	0.10(3)	0.39(10)	0.86(10)	0.11(3)	0.38(10)	0.87(10)
	$p_3^{\mathscr{F}}$	0.10(3)	0.39(9)	0.85(9)	0.10(3)	0.38(9)	0.86(9)	0.11(3)	0.37(9)	0.86(9)
	$p_4^{\mathscr{F}}$	0.04(1)	0.25(1)	0.62(3)	0.04(1)	0.24(1)	0.63(3)	0.04(1)	0.24(1)	0.63(3)
	$p_{1,1}^{\mathscr{E}}$	0.23(5)	0.30(2)	0.43(1)	0.23(5)	0.30(2)	0.44(1)	0.24(5)	0.30(2)	0.44(1)
	$p_{1,2}^{\mathscr{E}}$	0.23(5)	0.30(2)	0.43(1)	0.23(5)	0.30(2)	0.44(1)	0.24(5)	0.30(2)	0.44(1)
	$p_{2,1}^{\mathscr{E}}$	0.25(7)	0.37(5)	0.68(5)	0.26(7)	0.37(5)	0.68(5)	0.27(7)	0.36(5)	0.69(5)
	$p_{2,2}^{\mathscr{E}}$	0.25(7)	0.37(5)	0.68(5)	0.26(7)	0.37(5)	0.68(5)	0.27(7)	0.36(5)	0.69(5)
	$p_{2,3}^{\mathscr{E}}$	0.25(7)	0.37(5)	0.68(5)	0.26(7)	0.37(5)	0.68(5)	0.27(7)	0.36(5)	0.69(5)
	$p_{2,4}^{\mathscr{E}}$	0.25(7)	0.37(5)	0.68(5)	0.26(7)	0.37(5)	0.68(5)	0.27(7)	0.36(5)	0.69(5)
	RMSE	3.51		0.77	3.51		0.77	3.51		0.77
	Corr	0.01		0.96	0.01		0.96	0.01		0.96

not be capable of reflecting the directions of changes in the compromise probabilities due to varying ν. For instance, we see that when $q = 0.4$, as ν increases, the simulated compromise probabilities decrease but the approximated probabilities increase. Nevertheless, the approximation intervals still cover the true compromise probabilities.

Remark 6.6. We openly admit that the accuracy of the interval method declines (i.e., the approximation intervals become wider) as the inside attack probabilities increase. This is because the risk propagation routes become more complicated if there are more frequent internal infections, which can not be adequately captured by the proposed independent or comonotonic approximation.

However, the current best practice of cybersecruity is to employ a network security monitoring system that can perform packet capture, deep packet inspection, flow-based examination and so forth, so that the inside attack probabilities can be significantly reduced (Ghafir, Prenosil, Svoboda, & Hammoudeh, 2016). Under the practical circumstance of low inside attack probabilities, the proposed approximation method works well. Nevertheless, we deliberately to consider more extremal—though not so realistic—cases in Example 6.5 so that the limitations of the interval approximation method can be communicated in a fully transparent manner. If evaluating the compromise probabilities is the primary interest, then we suggest the readers to use either the explicit approach (when it is available) or simulation approach with a large number of simulation trials to adequately calculate the compromise probabilities.

In future research, it will be very interesting to explore the possibility for incorporating with partial dependence information in order to improve the performance of the interval approximation method. For example, for some network structures, we may be able to partition the nodes into a few subgroups, in which we can compute their conditional compromise probabilities explicitly within each subgroup. Then the independent or comonotonic approximation idea can be applied to calculate the unconditional compromise probabilities across the entire network.

Finally, we consider the performance of interval approximation for identifying weak nodes. Table 9 presents the RMSE and Pearson's correlation coefficient between the ranks of the simulated probabilities and the ranks of the lower/upper bounds. Intuitively, a lower value of RMSE or a higher value of correlation indicates that the ranks of the lower/upper bounds can well approximate the ranks of the compromise risks of different nodes. We observe that when the inside attack probability parameter $q = 0.1$, the ranks of the upper bounds perfectly match the ranks of the simulated probabilities. The ranks approximated by the lower bounds have a comparable performance compared with the upper bounds. As q increases to 0.25, the performances of the rank approximations by both the lower and upper bounds decline yet still reasonable. The decline in performance in this case is because the true compromise probabilities among fog nodes 2, 3, and end nodes $(2,1),\dots,(2,4)$ are rather similar, and the approximation method is inadequate to properly capture their orders. When $q = 0.4$, we find the ranks approximated by the lower bounds do not perform satisfactorily. For instance, the ranks of the lower bounds fail to capture the two weakest nodes when $q = 0.4$. However, in this case, the upper bounds still provide a very good approximation for the ranks of the compromise probabilities. We notice that varying the common vulnerability parameter v does not seem to have impacts on the performance of the rank approximation by the lower and upper bounds. In general, the ranks of the lower bounds have more ties than that of the upper bounds, which is probably one of the reasons why the lower bounds perform worse than the upper bounds. In light of the discussions above, we suggest the readers to use the ranks of upper bounds for the weak node identification purpose.

When the attack parameters are assumed to be homogeneous within the same node type, then the interval approximation method can be immediately used to identify weak node types. For instance, in Example 6.5, the true values of the compromise probabilities for type 1 devices or type 2 devices are identical, and the type 1 end nodes have lower compromise probabilities than the type 2 end nodes. Both the lower and upper bounds induced by the proposed interval method can accurately capture the fact that the security of type 2 nodes is weaker than that of the type 1 nodes. When the compromise probabilities are not identical within the same node type, e.g., the fog nodes in Example 6.5, there can be different ways for ranking the security levels of different node types. For example, the ranking can be based on the minimum, medium, mean, and maximum of the compromise probabilities within each node type. It will be a very interesting future research topic to explore what the "best" approach is for identifying weak node types such that the cost-benefit of network enhancement can be optimized, and how the interval method proxies the cybersecurity ranking based on the "best" approach.

7. Conclusions

In this paper, we proposed a class of structural models for studying the cybersecurity risk propagation mechanism within a general fog network. In the smart home application, we obtained a set of explicit formulas to calculate the compromise probabilities precisely. An actuarial framework has been proposed to price the fog network's cybersecurity risks. It was discovered that the heterogeneity and interdependency features of the fog networks should never be overlooked in studying the cybersecurity risks. The inherent common vulnerabilities are also crucial in determining the risks and related pricing strategies. Finally, we suggested an interval approximation method to estimate the compromise probabilities for the individual network elements in a general fog network, which can be used to efficiently identify weak nodes.

The studies of fog computing from the risk management and insurance perspectives are still in the very early stage. The main challenges are the associated multi-tenant and resource-sharing architectures, which result in a considerably large attack surface. The current work makes a significant first step towards modeling and pricing the cybersecurity risks in fog networks. Moving forward, we are interested in studying the following important yet challenging issues: i) *Dynamic modeling of cybersecurity risks*. In our current study, the compromise probabilities are assumed to be static. However, in reality, the attackers' behaviors may have a learning curve, so the inside attack and outside attack probabilities may change over time. We plan to extend the current model to a dynamic epidemic spreading model in future research. In this respect, we find the probability framework proposed in Xu & Hua (2019) useful. ii) *Cascading effects*. Cascading failures can be incorporated into the modeling process, which refer to the failures of one or several nodes triggering the failures of the other nodes. Note that although cascading failures and cyber risk propagation are similar, cascading failures are mainly related to the physical layer, while the cyber risk propagation focuses on the communication/network layer. Since cascading failures are not uncommon in fog computing, they can be considered as another important risk factor. One may refer to Xing (2020) for a recent review about cascading failures in the IoT domain. iii) *Cyber insurance portfolio management*. Since fog computing is widely deployed in a variety of IoT applications, an insurance company can have several businesses lines, e.g., smart home, fog servers, and smart cars, which constitute a large cyber insurance portfolio. It is important for the insurance company to realize the diversification benefit and properly understand the systemic risk inherent in the portfolio.

Acknowledgments

We are indebted to Editor Steffen Rebennack, and four anonymous referees for constructive criticism and generous suggestions. We are grateful to the Society of Actuaries for the generous financial support of our research, as well as for appointing a very knowledgeable project oversight group, whose insightful comments and thought-provoking questions resulted in a much-improved version of this paper.

Appendix A. Technical proofs

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

$$\prod_{i=1}^{n}(1 - x_i) = 1 - \sum_{k=1}^{n}(-1)^{k-1} \sum_{\mathcal{N}_k \subseteq \mathcal{N}} h(\mathcal{N}_k), \tag{21}$$

where $h(\mathcal{N}_k) = \prod_{i \in \mathcal{N}_k} x_i$ for $\mathcal{N}_k \in \mathbb{N}^k$ is any k-dimensional subset of $\mathcal{N}$. Together with Eq. (7), we have

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

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

$$\begin{aligned}
h(\Xi_k) &= \mathbb{E}\Big[\prod_{j \in \Xi_k} \big(1 - O_j^{\mathscr{F}}\big) \times \prod_{d=1}^{n^{\mathscr{T}}} \prod_{i_d \in \mathbb{D}_{d,j}} \big(1 - O^{\mathscr{E}}_{d,i_d} \times I^{\mathscr{E}}_{(d,i_d)\to j}\big)\Big] \\
&= \mathbb{E}\Big[\prod_{j \in \Xi_k} \big(1 - O_j^{\mathscr{F}}\big)\Big] \times \mathbb{E}\Big[\prod_{j \in \Xi_k}\prod_{d=1}^{n^{\mathscr{T}}} \prod_{i_d \in \mathbb{D}_{d,j}} \big(1 - O^{\mathscr{E}}_{d,i_d} \times I^{\mathscr{E}}_{(d,i_d)\to j}\big)\Big].
\end{aligned}$$

The expectations above can be further computed via

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

as well as

$$\begin{aligned}
&\mathbb{E}\Big[\prod_{j \in \Xi_k}\prod_{d=1}^{n^{\mathscr{T}}} \prod_{i_d \in \mathbb{D}_{d,j}} \big(1 - O^{\mathscr{E}}_{d,i_d} \times I^{\mathscr{E}}_{(d,i_d)\to j}\big)\Big] \\
&= \prod_{d=1}^{n^{\mathscr{T}}} \mathbb{E}\Big[\prod_{j \in \Xi_k} \prod_{i_d \in \mathbb{D}_{d,j}} \big(1 - O^{\mathscr{E}}_{d,i_d} \times I^{\mathscr{E}}_{(d,i_d)\to j}\big)\Big] \\
&= \prod_{d=1}^{n^{\mathscr{T}}} \Big\{ (1 - \nu_d^{\mathscr{E}})\,\mathbb{E}\Big[\prod_{j \in \Xi_k} \prod_{i_d \in \mathbb{D}_{d,j}} \big(1 - O^{\mathscr{E}}_{d,i_d} \times I^{\mathscr{E}}_{(d,i_d)\to j}\big)\,\Big|\, V_d^{\mathscr{E}} = 0\Big] \\
&\quad + \nu_d^{\mathscr{E}}\,\mathbb{E}\Big[\prod_{j \in \Xi_k} \prod_{i_d \in \mathbb{D}_{d,j}} \big(1 - O^{\mathscr{E}}_{d,i_d} \times I^{\mathscr{E}}_{(d,i_d)\to j}\big)\,\Big|\, V_d^{\mathscr{E}} = 1\Big]\Big\} \\
&= \prod_{d=1}^{n^{\mathscr{T}}} \Big[(1 - \nu_d^{\mathscr{E}}) \prod_{j \in \Xi_k} \prod_{i_d \in \mathbb{D}_{d,j}} (1 - \pi^{\mathscr{E}}_{d,i_d} q^{\mathscr{E}}_{(d,i_d)\to j}) \\
&\quad + \nu_d^{\mathscr{E}}\Big(1 - \pi_d^{\mathscr{E}*} + \pi_d^{\mathscr{E}*} \prod_{j \in \Xi_k} \prod_{i_d \in \mathbb{D}_{d,j}} \big(1 - q^{\mathscr{E}}_{(d,i_d)\to j}\big)\Big)\Big] \\
&= \prod_{d=1}^{n^{\mathscr{T}}} g(d, \Xi_k).
\end{aligned} \tag{22}$$

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

$$\begin{aligned}
f(j_d, \Xi) &= \mathbb{E}\Big[O^{\mathscr{E}}_{d,j_d} \prod_{j \in \Xi} C_j^{\mathscr{F},[\bullet]}\Big] \\
&= \mathbb{E}\Big[O^{\mathscr{E}}_{d,j_d} \prod_{j \in \Xi}\Big[1 - (1 - O_j^{\mathscr{F}})\prod_{s=1}^{n^{\mathscr{T}}} \prod_{l_s \in \mathbb{D}_{s,j}} \big(1 - O^{\mathscr{E}}_{s,l_s} \times I^{\mathscr{E}}_{(s,l_s)\to j}\big)\Big]\Big] \\
&= \omega^{\mathscr{E}}_{d,j_d} - \sum_{k=1}^{m}(-1)^{k-1} \sum_{\Xi_k \subseteq \Xi} u(j_d, \Xi_k),
\end{aligned}$$

where

$$\begin{aligned}
u(j_d, \Xi_k) &= \mathbb{E}\Big[O^{\mathscr{E}}_{d,j_d} \prod_{j \in \Xi_k}(1 - O_j^{\mathscr{F}})\prod_{s=1}^{n^{\mathscr{T}}} \prod_{l_s \in \mathbb{D}_{s,j}} \big(1 - O^{\mathscr{E}}_{s,l_s} \times I^{\mathscr{E}}_{(s,l_s)\to j}\big)\Big] \\
&= \Big[(1 - \nu^{\mathscr{F}}) \prod_{j \in \Xi_k}(1 - \pi_j^{\mathscr{F}}) + \nu^{\mathscr{F}}(1 - \pi^{\mathscr{F}*})\Big] \\
&\quad \times \mathbb{E}\Big[O^{\mathscr{E}}_{d,j_d} \prod_{j \in \Xi_k}\prod_{s=1}^{n^{\mathscr{T}}} \prod_{l_s \in \mathbb{D}_{s,j}} \big(1 - O^{\mathscr{E}}_{s,l_s} \times I^{\mathscr{E}}_{(s,l_s)\to j}\big)\Big].
\end{aligned}$$

The expectation above is computed via

$$\begin{aligned}
&\mathbb{E}\Big[O^{\mathscr{E}}_{d,j_d} \prod_{j \in \Xi_k}\prod_{s=1}^{n^{\mathscr{T}}} \prod_{l_s \in \mathbb{D}_{s,j}} \big(1 - O^{\mathscr{E}}_{s,l_s} I^{\mathscr{E}}_{(s,l_s)\to j}\big)\Big] \\
&= \mathbb{E}\Big[O^{\mathscr{E}}_{d,j_d} \prod_{j \in \Xi_k} \prod_{l_d \in \mathbb{D}_{d,j}} \big(1 - O^{\mathscr{E}}_{d,l_d} I^{\mathscr{E}}_{(d,l_d)\to j}\big)\Big] \\
&\quad \prod_{s=1,\, s\neq d}^{n^{\mathscr{T}}} \mathbb{E}\Big[\prod_{j \in \Xi_k} \prod_{l_s \in \mathbb{D}_{s,j}} \big(1 - O^{\mathscr{E}}_{s,l_s} I^{\mathscr{E}}_{(s,l_s)\to j}\big)\Big] \\
&= \Big[(1 - \nu_d^{\mathscr{E}})\,\pi^{\mathscr{E}}_{d,j_d}(1 - q^{\mathscr{E}}_{(d,j_d)\to i}) \prod_{j \in \Xi_k} \prod_{l_d \in \mathbb{D}_{d,j},\, l_d \neq j_d} (1 - \pi^{\mathscr{E}}_{d,l_d} q^{\mathscr{E}}_{(d,l_d)\to j}) \\
&\quad + \nu_d^{\mathscr{E}}\,\pi_d^{\mathscr{E}*} \prod_{j \in \Xi_k} \prod_{l_d \in \mathbb{D}_{d,j}} (1 - q^{\mathscr{E}}_{(d,l_d)\to j})\Big] \\
&\quad \times \prod_{s=1,\, s\neq d}^{n^{\mathscr{T}}} g(s, \Xi_k).
\end{aligned}$$

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 (6), we have

$$\begin{aligned}
p^{\mathscr{C}} &= 1 - \mathbb{E}\big[(1 - O^{\mathscr{C}})\big] \times \mathbb{E}\Big[\prod_{i=1}^{n^{\mathscr{F}}}(1 - C_i^{\mathscr{F},[\bullet]} \times I^{\mathscr{F}}_{i\to\bullet})\Big] \\
&= 1 - (1 - \omega^{\mathscr{C}})\Big[1 - \sum_{k=1}^{n^{\mathscr{F}}}(-1)^{k-1} \sum_{\Xi_k \subseteq \Xi^{\mathscr{F}}} \tilde{p}^{\mathscr{F}}_{\Xi_k} \prod_{i \in \Xi_k} q^{\mathscr{F}}_{i\to\bullet}\Big]
\end{aligned}$$

where the last equation holds because of the product formula in (21).

Turing to the study of fog nodes, elaborate the state Eq. (8) as

$$C_i^{\mathscr{F}} = 1 - \left(1 - C_i^{\mathscr{F},[\bullet]}\right)\left(1 - I_{\bullet\to i}^{\mathscr{C}}\right) - \left(1 - C_i^{\mathscr{F},[\bullet]}\right) I_{\bullet\to i}^{\mathscr{C}} \left(1 - O^{\mathscr{C}}\right) \prod_{j=1, j\neq i}^{n^{\mathscr{F}}} \left(1 - C_j^{\mathscr{F},[\bullet]} I_{j\to\bullet}^{\mathscr{F}}\right), \tag{23}$$

and hence

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

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

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

We have

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

Finally, let us consider the compromise probability for the j_dth end node that is of type d, $d = 1, \ldots, n^{\mathscr{T}}$, $j_d = 1 \ldots, n_d^{\mathscr{E}}$, and it has a direct connection to the ith fog node, i.e., $j_d \in \mathbb{D}_{d,i}$. According to the state Eq. (9), we get

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

Evoking the state equation for fog node in (23), the expectation above can be computed via

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

in which, by evoking Lemma 4.3,

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

and

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

We focus on the evaluation of t_2 and obtain

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

where $f(j_d, \Xi_k) = \mathbb{E}\left[O_{(d,j_d)}^{\mathscr{E}} \prod_{j\in\Xi_k} C_j^{\mathscr{F},[\bullet]}\right]$ which can be computed by evoking Lemma 4.3. The proof is finished. □

Proof of Proposition 6.3. We begin by proving that the outside attack RV's are positively associated. Focus on the fog nodes first and consider the RV $\boldsymbol{O}^{\mathscr{F}} = (O_1^{\mathscr{F}}, \ldots, O_{n^{\mathscr{F}}}^{\mathscr{F}})$, it is straightforward to check that

$$\begin{aligned}
&\mathbb{P}(O_i^{\mathscr{F}} > o_i \,|\, O_j^{\mathscr{F}} = o_j,\ j = 1, \ldots, i-1) \\
&= (1 - \nu^{\mathscr{F}})\, \mathbb{P}(O_i^{\mathscr{F}} > o_i \,|V^{\mathscr{F}} = 0) \\
&+ \nu^{\mathscr{F}}\, \mathbb{P}(O_i^{\mathscr{F}} > o_i \,|V^{\mathscr{F}} = 1,\ O_j^{\mathscr{F}} = o_j,\ j = 1, \ldots, i-1)
\end{aligned}$$

is nondecreasing in $(o_1, \ldots, o_{i-1}) \in \{0,1\}^{i-1}$ for all $o_i \in \{0,1\}$, $i = 2, \ldots, n^{\mathscr{F}}$. So the RV $\boldsymbol{O}^{\mathscr{F}}$ is conditionally increasing in sequence, which implies positive association (see, Theorem 2.4 in Joe, 1997). A repeated application of the aforementioned argument to $\boldsymbol{O}_d^{\mathscr{E}}$, $d = 1, \ldots, n^{\mathscr{T}}$, yields that the elements in each $\boldsymbol{O}_d^{\mathscr{E}}$ are positively associated. Because of Assumption 3.1, the RV's $\boldsymbol{O}^{\mathscr{F}}, \boldsymbol{O}_1^{\mathscr{E}}, \ldots, \boldsymbol{O}_{n^{\mathscr{T}}}^{\mathscr{E}}$ 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 Eqs. (2) and (4) can be expressed as

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

respectively, for some coordinate-wise non-decreasing functions $h_i(\cdot)$ and $h_{d,i_d}(\cdot)$, $i = 1, \ldots, n^{\mathscr{F}}$, $i_d = 1, \ldots, n_d^{\mathscr{E}}$, $d = 1, \ldots, n^{\mathscr{T}}$. Evoking Lemma 6.2, we can conclude that $(\boldsymbol{C}, \boldsymbol{I}, \boldsymbol{O})$ is positively associated. This completes the proof. □

Proof of Theorem 6.4. To start off, note that the compromise probabilities associated with state Eqs. (2) and (4) can be computed via

$$p_j^{\mathscr{F}} = 1 - \mathbb{P}\left(O_j^{\mathscr{F}} \le 0,\ \bigcap_{i=1, i\neq j}^{n^{\mathscr{F}}} C_i^{\mathscr{F}} I_{i\to j}^{\mathscr{F}} \le 0,\ \bigcap_{d=1}^{n^{\mathscr{T}}} \bigcap_{i_d=1}^{n_d^{\mathscr{E}}} C_{d,i_d}^{\mathscr{E},[j]} I_{(d,i_d)\to j}^{\mathscr{E}} \le 0\right), \quad \text{for } j = 1, \ldots, n^{\mathscr{F}}, \tag{24}$$

and

$$p_{d,j_d}^{\mathscr{E}} = 1 - \mathbb{P}\left(O_{d,j_d}^{\mathscr{E}} \le 0,\ \bigcap_{i=1, i\neq j}^{n^{\mathscr{F}}} C_i^{\mathscr{F}} I_{i\to(d,j_d)}^{\mathscr{F}} \le 0\right), \quad \text{for } d = 1, \ldots, n^{\mathscr{T}},\ j_d = 1, \ldots, n_d^{\mathscr{E}}. \tag{25}$$

Thus the task in this proof boils down to identifying the lower and upper bounds for the cumulative distribution functions of dependent binary RV's in Eqs. (24) and (25).

First, consider the compromise RV $C_{d,i_d}^{\mathscr{E},[j]}$ defined in Eq. (3). We have proved in Proposition 6.3 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

$$
\begin{aligned}
\mathbb{E}\left[C_{d,i_d}^{\mathscr{E},[j]}\right] &= 1-\mathbb{P}\left(O_{d,i_d}^{\mathscr{E}} \le 0,\ \bigcap_{i=1,i\neq j}^{n^{\mathscr{F}}} C_i^{\mathscr{F}} I_{i\to(d,i_d)}^{\mathscr{F}} \le 0\right)\\
&\le 1-\mathbb{P}\left(O_{d,i_d}^{\mathscr{E}} \le 0\right) \prod_{i=1,i\neq j}^{n^{\mathscr{F}}} \mathbb{P}\left(C_i^{\mathscr{F}} I_{i\to(d,i_d)}^{\mathscr{F}} \le 0\right)\\
&= 1-\left[1-\nu_d^{\mathscr{E}}\,\pi_d^{\mathscr{E}*}-(1-\nu_d^{\mathscr{E}})\,\pi_{d,i_d}^{\mathscr{E}}\right] \prod_{i=1,i\neq j}^{n^{\mathscr{F}}} \left(1-p_i^{\mathscr{F}}\, q_{i\to(d,i_d)}^{\mathscr{F}}\right).
\end{aligned}
$$

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

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

So we get

$$
\begin{aligned}
&\max\left(\omega_{d,i_d}^{\mathscr{E}},\ \bigvee_{i=1,i\neq j}^{n^{\mathscr{F}}} p_i^{\mathscr{F}}\, q_{i\to(d,i_d)}^{\mathscr{F}}\right)\\
&\le\ \mathbb{E}\left[C_{d,i_d}^{\mathscr{E},[j]}\right]\ \le\ 1-\left(1-\omega_{d,i_d}^{\mathscr{E}}\right) \prod_{i=1,i\neq j}^{n^{\mathscr{F}}} \left(1-p_i^{\mathscr{F}}\, q_{i\to(d,i_d)}^{\mathscr{F}}\right). \qquad (26)
\end{aligned}
$$

Next, we turn to the compromise probabilities of fog nodes in Eq. (24). Another application of the property of positive association yields

$$
\begin{aligned}
p_j^{\mathscr{F}} &= 1-\mathbb{P}\left(O_j^{\mathscr{F}} \le 0,\ \bigcap_{i=1,i\neq j}^{n^{\mathscr{F}}} C_i^{\mathscr{F}} I_{i\to j}^{\mathscr{F}} \le 0,\ \bigcap_{d=1}^{n^{\mathscr{T}}} \bigcap_{i_d=1}^{n_d^{\mathscr{E}}} C_{d,i_d}^{\mathscr{E},[j]} I_{(d,i_d)\to j}^{\mathscr{E}} \le 0\right)\\
&\le 1-\mathbb{P}\left(O_j^{\mathscr{F}} \le 0\right) \prod_{i=1,i\neq j}^{n^{\mathscr{F}}} \mathbb{P}\left(C_i^{\mathscr{F}} I_{i\to j}^{\mathscr{F}} \le 0\right) \prod_{d=1}^{n^{\mathscr{T}}} \prod_{i_d=1}^{n_d^{\mathscr{E}}} \mathbb{P}\left(C_{d,i_d}^{\mathscr{E},[j]} I_{(d,i_d)\to j}^{\mathscr{E}} \le 0\right)\\
&= 1-\left(1-\omega_j^{\mathscr{F}}\right) \prod_{i=1,i\neq j}^{n^{\mathscr{F}}} \left(1-p_i^{\mathscr{F}} q_{i\to j}^{\mathscr{F}}\right) \prod_{d=1}^{n^{\mathscr{T}}} \prod_{i_d=1}^{n_d^{\mathscr{E}}} \left[1-\mathbb{E}\left[C_{d,i_d}^{\mathscr{E},[j]}\right] q_{(d,i_d)\to j}^{\mathscr{E}}\right].
\end{aligned}
\qquad (27)
$$

Evoke the upper bound derived in Eq. (26), we get

$$
\begin{aligned}
&1-\mathbb{E}\left[C_{d,i_d}^{\mathscr{E},[j]}\right] q_{(d,i_d)\to j}^{\mathscr{E}}\\
&\ge\ 1-q_{(d,i_d)\to j}^{\mathscr{E}}+q_{(d,i_d)\to j}^{\mathscr{E}}\left(1-\omega_{d,i_d}^{\mathscr{E}}\right) \prod_{i=1,i\neq j}^{n^{\mathscr{F}}} \left(1-p_i^{\mathscr{F}}\, q_{i\to(d,i_d)}^{\mathscr{F}}\right)\\
&\ge\ 1-q_{(d,i_d)\to j}^{\mathscr{E}}+q_{(d,i_d)\to j}^{\mathscr{E}}\left(1-\omega_{d,i_d}^{\mathscr{E}}\right) \prod_{i=1,i\neq j}^{n^{\mathscr{F}}} \left(1-q_{i\to(d,i_d)}^{\mathscr{F}}\right). \qquad (28)
\end{aligned}
$$

Combining the inequalities derived in (27) and (28) leads to

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

where inequality "$\overset{(1)}{=}$" holds by Weierstrass product inequality. Define an $n^{\mathscr{F}}$ by $n^{\mathscr{F}}$ zero diagonal matrix, $\boldsymbol{A}$ with off-diagonal elements $a_{ij}=\gamma_i\, q_{j\to i}^{\mathscr{F}}$ for $i\neq j=1,\ldots,n^{\mathscr{F}}$. Then the upper bounds of the compromise probabilities for fog nodes, $\boldsymbol{u}^{\mathscr{F}}=(u_1^{\mathscr{F}},\ldots,u_{n^{\mathscr{F}}}^{\mathscr{F}})^\top$, solve the matrix equation $\boldsymbol{u}^{\mathscr{F}}=1-\boldsymbol{\gamma}+\boldsymbol{A}\boldsymbol{u}^{\mathscr{F}}$, or equivalently, $\boldsymbol{u}^{\mathscr{F}}=(\mathbf{1}-\boldsymbol{A})^{-1}(1-\boldsymbol{\gamma})$ if the spectral radius of $\boldsymbol{A}$ is less than 1, where $\mathbf{1}$ denotes an identify matrix of appropriate dimension.

Contrastingly,

$$
\begin{aligned}
p_j^{\mathscr{F}} &= 1-\mathbb{P}\left(O_j^{\mathscr{F}} \le 0,\ \bigcap_{i=1,i\neq j}^{n^{\mathscr{F}}} C_i^{\mathscr{F}} I_{i\to j}^{\mathscr{F}} \le 0,\ \bigcap_{d=1}^{n^{\mathscr{T}}} \bigcap_{i_d=1}^{n_d^{\mathscr{E}}} C_{d,i_d}^{\mathscr{E},[j]} I_{(d,i_d)\to j}^{\mathscr{E}} \le 0\right)\\
&\overset{(1)}{\ge} \max\left(\omega_j^{\mathscr{F}},\ \bigvee_{i=1,i\neq j}^{n^{\mathscr{F}}} p_i^{\mathscr{F}} q_{i\to j}^{\mathscr{F}},\ \bigvee_{d=1}^{n^{\mathscr{T}}} \bigvee_{i_d=1}^{n_d^{\mathscr{E}}} \mathbb{E}\left[C_{d,i_d}^{\mathscr{E},[j]}\right] q_{(d,i_d)\to j}^{\mathscr{E}}\right)\\
&\overset{(2)}{\ge} \max\left(\omega_j^{\mathscr{F}},\ \bigvee_{i=1,i\neq j}^{n^{\mathscr{F}}} p_i^{\mathscr{F}} q_{i\to j}^{\mathscr{F}},\ \bigvee_{d=1}^{n^{\mathscr{T}}} \bigvee_{i_d=1}^{n_d^{\mathscr{E}}} \max\left(\omega_{d,i_d}^{\mathscr{E}},\ \bigvee_{i=1,i\neq j}^{n^{\mathscr{F}}} p_i^{\mathscr{F}}\, q_{i\to(d,i_d)}^{\mathscr{F}}\right) q_{(d,i_d)\to j}^{\mathscr{E}}\right)\\
&\ge \max\left(\omega_j^{\mathscr{F}},\ \bigvee_{i=1,i\neq j}^{n^{\mathscr{F}}} \beta_i\, q_{i\to j}^{\mathscr{F}},\ \bigvee_{d=1}^{n^{\mathscr{T}}} \bigvee_{i_d=1}^{n_d^{\mathscr{E}}} \max\left(\omega_{d,i_d}^{\mathscr{E}},\ \bigvee_{i=1,i\neq j}^{n^{\mathscr{F}}} \beta_i\, q_{i\to(d,i_d)}^{\mathscr{F}}\right) q_{(d,i_d)\to j}^{\mathscr{E}}\right),
\end{aligned}
$$

where inequalities "$\overset{(1)}{=}$" and "$\overset{(2)}{=}$" hold because of Fréchet inequalities and the lower bound derived in Eq. (26), and

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

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

Applying the same argument as in the derivation of inequalities (26) yields

$$
\max\left(\omega_{d,j_d}^{\mathscr{E}},\ \bigvee_{i=1}^{n^{\mathscr{F}}} p_i^{\mathscr{F}}\, q_{i\to(d,j_d)}^{\mathscr{F}}\right)\ \le\ p_{d,j_d}^{\mathscr{E}}\ \le\ 1-\left(1-\omega_{d,j_d}^{\mathscr{E}}\right) \prod_{i=1}^{n^{\mathscr{F}}} \left(1-p_i^{\mathscr{F}}\, q_{i\to(d,j_d)}^{\mathscr{F}}\right),
$$

for $d=1,\ldots,n^{\mathscr{T}}$ and $j_d=1,\ldots,n_d^{\mathscr{E}}$. Finally, substitute the lower and upper bounds for $\boldsymbol{p}^{\mathscr{F}}$ into the inequalities above, the interval approximations for the end nodes' compromise probabilities are readily obtained. The proof is finished. □

Appendix B. Summary of notations

Here are some additional notations needed in Theorem 6.4 for constructing the upper bonds of compromise probabilities. Let

$$
\boldsymbol{u}^{\mathscr{F}} = (u_1^{\mathscr{F}},\ldots,u_{n^{\mathscr{F}}}^{\mathscr{F}})^\top = (\mathbf{1}-\boldsymbol{A})^{-1}(1-\boldsymbol{\gamma}), \qquad (29)
$$

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

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

and $\boldsymbol{A}$ is an $n^{\mathscr{F}}$ by $n^{\mathscr{F}}$ zero diagonal matrix having off-diagonal elements $a_{ij}=\gamma_i\, q_{j\to i}^{\mathscr{F}}$ for $i\neq j\in\{1,\ldots,n^{\mathscr{F}}\}$. Here, we assume that the spectral radius of $\boldsymbol{A}$ is less than 1. The upper bonds for the end nodes are $\boldsymbol{u}_d^{\mathscr{E}}=(u_{d,1}^{\mathscr{E}},\ldots,u_{d,n_d^{\mathscr{E}}}^{\mathscr{E}})^\top$ with elements

$$
u_{d,j_d}^{\mathscr{E}} = 1-\left(1-\omega_{d,j_d}^{\mathscr{E}}\right) \prod_{i=1}^{n^{\mathscr{F}}} \left(1-u_i^{\mathscr{F}}\, q_{i\to(d,j_d)}^{\mathscr{F}}\right), \qquad (30)
$$

where $u_i^{\mathscr{F}}$, $i=1,\ldots,n^{\mathscr{F}}$, is specified in Eq. (29).

Finally, Table 10 summarizes the notation system used throughout this present article.

Table 10
Summary of the notation system.

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

References

Agosto, A., & Ahelegbey, D. F. (2022). Default count-based network models for credit contagion. *Journal of the Operational Research Society, 73*(1), 139–152.

Almasizadeh, J., & Azgomi, M. A. (2013). A stochastic model of attack process for the evaluation of security metrics. *Computer Networks, 57*(10), 2159–2180.

Baccarelli, E., Naranjo, P. G. V., Scarpiniti, M., Shojafar, M., & 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., & Wirfs, J. H. (2015). Insurability of cyber risk: An empirical analysis. *Geneva Papers on Risk and Insurance: Issues and Practice, 40*(1), 131–158. https://doi.org/10.1057/gpp.2014.19.

Böhme, R. (2005). Cyber-insurance revisited. In *Weis*.

Capponi, A., & Jarrow, R. (2021). Preface to the special issue on systemic risk and financial networks. *Mathematics and Financial Economics, 15*(1), 1–3.

Cheung, K.-F., & Bell, M. G. (2021). Attacker–defender model against quantal response adversaries for cyber security in logistics management: An introductory study. *European Journal of Operational Research, 291*(2), 471–481.

Dacier, M., Deswarte, Y., & Kaâniche, M. (1996). Models and tools for quantitative assessment of operational security. In *IFIP international conference on ICT systems security and privacy protection* (pp. 177–186). Springer.

Darwish, T. S., & 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.

Detering, N., Meyer-Brandis, T., Panagiotou, K., & Ritter, D. (2019). Managing default contagion in inhomogeneous financial networks. *SIAM Journal on Financial Mathematics, 10*(2), 578–614.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R., & Vyncke, D. (2002a). The concept of comonotonicity in actuarial science and finance: Applications. *Insurance: Mathematics and Economics, 31*(2), 133–161. https://doi.org/10.1016/S0167-6687(02)00135-X.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R., & Vyncke, D. (2002b). The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics, 31*(1), 3–33. https://doi.org/10.1016/S0167-6687(02)00134-8.

Eling, M., & Wirfs, J. (2019). What are the actual costs of cyber risk events? *European Journal of Operational Research, 272*(3), 1109–1119. https://doi.org/10.1016/j.ejor.2018.07.021.

Fahrenwaldt, M. A., Weber, S., & Weske, K. (2018). Pricing of cyber insurance contracts in a network model. *ASTIN Bulletin, 48*(3), 1175–1218. https://doi.org/10.1017/asb.2018.23.

Feng, S., Xiong, Z., Niyato, D., Wang, P., & Leshem, A. (2018). Evolving risk management against advanced persistent threats in fog computing. In *2018 IEEE 7th international conference on cloud networking* (pp. 1–6).

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Annals de l'Université de Lyon, 9*, 53–77.

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

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

Ghafir, I., Prenosil, V., Svoboda, J., & Hammoudeh, M. (2016). A survey on network security monitoring systems. In *2016 IEEE 4th international conference on future internet of things and cloud workshops (FiCloudW)* (pp. 77–82). IEEE.

Jevtić, P., & 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. https://doi.org/10.1016/j.insmatheco.2020.02.005.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. London: Chapman and Hall. https://doi.org/10.1201/b13150.

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

Khouzani, M., Liu, Z., & Malacaria, P. (2019). Scalable min-max multi-objective cyber-security optimisation over probabilistic attack graphs. *European Journal of Operational Research, 278*(3), 894–903.

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

Lynn, T., Endo, P. T., Ribeiro, A. M. N. C., Barbosa, G. B. N., & Rosati, P. (2020). *The internet of things: Definitions, key concepts, and reference architectures*. In T. Lynn, J. G. Mooney, B. Lee, & P. T. Endo (Eds.) (pp. 1–22)). Cham: Springer.

Nagurney, A., & Shukla, S. (2017). Multifirm models of cybersecurity investment competition vs. cooperation and network vulnerability. *European Journal of Operational Research, 260*(2), 588–600.

NAIC (2020). 2020 Report on the cybersecurity insurance and identity theft coverage supplement. *Technical Report*. National Association of Insurance Commissioners.

Ortalo, R., Deswarte, Y., & Kaâniche, M. (1999). Experimenting with quantitative evaluation tools for monitoring operational security. *IEEE Transactions on Software Engineering, 25*(5), 633–650.

Paul, J. A., & Zhang, M. (2021). Decision support model for cybersecurity risk planning: A two-stage stochastic programming framework featuring firms, government, and attacker. *European Journal of Operational Research, 291*(1), 349–364.

Poolsappasit, N., Dewri, R., & Ray, I. (2011). Dynamic security risk management using Bayesian attack graphs. *IEEE Transactions on Dependable and Secure Computing, 9*(1), 61–74.

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

Ramos, A., Lazar, M., Holanda Filho, R., & Rodrigues, J. J. (2017). Model-based quantitative network security metrics: A survey. *IEEE Communications Surveys & Tutorials, 19*(4), 2704–2734.

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

Simon, J., & Omar, A. (2020). Cybersecurity investments in the supply chain: Coordination and a strategic attacker. *European Journal of Operational Research, 282*(1), 161–171.

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

Veraart, L. A. M. (2020). Distress and default contagion in financial networks. *Mathematical Finance, 30*(3), 705–737.

Wang, H., Chen, Z., Zhao, J., Di, X., & Liu, D. (2018). A vulnerability assessment method in industrial internet of things based on attack graph and maximum flow. *IEEE Access, 6*, 8599–8609.

Wang, S., Zhang, Z., & Kadobayashi, Y. (2013). Exploring attack graph for cost-benefit security hardening: A probabilistic approach. *Computers and Security, 32*, 158–169.

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., & Xu, S. (2015). Cyber epidemic models with dependences. *Internet Mathematics, 11*(1), 62–92. https://doi.org/10.1080/15427951.2014.902407.

Xu, M., & Hua, L. (2019). Cybersecurity insurance: Modeling and pricing. *North American Actuarial Journal, 23*(2), 220–249. https://doi.org/10.1080/10920277.2019.1566076.

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

Yu, T., Sekar, V., Seshan, S., Agarwal, Y., & 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* (pp. 1–7).