E-ISSN 2586-9027
Homepage : *https://tci-thaijo.org/index.php/SciTechAsia*
Science & Technology Asia
Vol.30 No.1 January - March 2025
Page: [177-188]
Original research article

Cybersecurity Insurance Modeling Using Archimedean Copulas

Nasya Kanya Puteri, Felivia Kusnadi*, Farah Kristiani

Center for Mathematics and Society, Faculty of Science, Parahyangan Catholic University, Bandung 40141, Indonesia

Received 18 April 2024; Received in revised form 14 December 2024
Accepted 17 January 2025; Available online 24 March 2025

ABSTRACT

While technology has brought many benefits, it has also created challenges, including a rise in cyber-crime. One major issue is the spread of malware, such as viruses, which can threaten personal, business, and national security. This can be countered by purchasing cybersecurity insurance. To accommodate the increasing demand, it is necessary to provide the adequate method to measure the risks and estimate the premium charge. In this paper, the modeling of the infection and recovery process of an electrical device that is represented as a node in a network system is discussed. The model that is discussed is a non-Markov model with dependence between cybersecurity risks, which is modeled by the Archimedean copula function, namely Clayton, Frank, and Gumbel. Furthermore, the premium charge is estimated using standard deviation and the exponential utility principle. Based on the simulation, it can be concluded that a node connected with several nodes is more likely to be infected than a node connected with fewer number of nodes. Modeling infection times between nodes using the Gumbel copula function generates higher premiums than other copula functions, therefore it is better to use the Gumbel copula function for modeling cybersecurity insurance in the first contract period because insured companies tend to be more interested in extending the contract if the premium charge of the subsequent year is less than or equal to the rate of the previous year. In addition, changes in parameter of times-to-infections from neighbors does not cause a significant difference in expected number of infections, expected losses, and premium charge.

Keywords: Archimedean copula; Computer network; Cybersecurity insurance; Dependence; Virus spread

***Corresponding author:** felivia@unpar.ac.id
doi: 10.14456/scitechasia.2025.11

1. Introduction and Motivation

The rapid development of technology in this era of globalization has brought many positive impacts in people's lives especially during the COVID-19 pandemic, such as availability of face-to-face virtual communication, e-mail exchange, online banking, etc. However, technology development also have negative impacts namely cyber-crime or cyber-attack such as ransomware which causes important company data and information hacked [4, 8, 14]. Therefore, companies need to come with smart strategies for prevention, protection, and loss countermeasures. One of the solutions is to buy a cybersecurity insurance policy. Several insurance companies have issued cybersecurity insurance, such as AXA XL Insurance, Zurich Insurance Group Ltd, and Tokio Marine Indonesia.

The number of companies that require cybersecurity risk transfer has caused the cybersecurity insurance market to expand. However, cybersecurity insurance still does not have a standard measurement system such as a standard actuarial table due to the fact that the term cybersecurity is still relatively new and the historical data related to existing cyber-attacks is not sufficient. Furthermore, companies are still afraid to report cyber-attacks because the impacts can lead to a decrease in market share, bad reputation for the company, and other losses. Therefore, insurance companies need a strategy to estimate the amount of cybersecurity insurance premiums.

Several methods to estimate the amount of cybersecurity insurance premiums have been explored. One of the methods used was discussed by Schwartz and Sastry [12] who developed cybersecurity risk management on large-scale interdependent networks. However, the model they discuss uses the assumption of a constant-value infection probability which is less suitable for use because devices that are connected to numerous other devices generally have a higher chance of infection. Another alternative is the Markov method discussed by Xu and Hua [15] in their paper. However, the Markov model cannot be used if the infection and recovery processes are not exponentially distributed. In the same paper, the non-Markov model is also discussed with an independent case in which cybersecurity risks between devices represented as nodes in the network are assumed to be independent. However, in real life, the infection of a network in a computer can spread to other networks. Model discussed by Xu and Hua [15] is further explored, utilizing the copula function that models dependency in cybersecurity risks. In particular, we use the Archimedean copulas, namely Clayton, Gumbel, and Frank. The Archimedean copulas will be used because its explicit form makes it easy to implement. The three copulas mentioned were chosen because non-symmetric high-dimensional Archimedean copulas require special constructions.

This study is conducted as follows. The next section discusses the Archimedean copulas to model dependency in cybersecurity risks and the methodology to calculate the insurance premium. The third section shows the simulation and comparison of the premiums for the three copulas. Finally, we analyze the impact of changes in one of the time-to-infections parameters on the simulation results.

2. Epidemic Spreading Model

2.1 Archimedean copulas

In this section, we discuss non-Markov with dependence between times-to-infections that is modeled using the Archimedean copulas. In this paper, a company's network system is represented in the form of a simple undirected graph. A graph notated as $\Gamma = (V; E)$ consists of $V = \{v_1, v_2, \ldots, v_n\}$, a nonempty set of vertices (or nodes) and E, a set of edges [10]. From the graph, a matrix A known as the adjacency matrix is formed. The elements of the matrix indicate whether pairs of vertices are adjacent or not in the graph, i.e.:

$$a_{i,j} = \begin{cases} 1, & \text{if there is an edge in } E \\ & \text{that connected } v_i \text{ to } v_j; \\ 0, & \text{otherwise.} \end{cases} \quad (2.1)$$

A node v holds only one status of secure, infected (but vulnerable to attacks), or infected (can potentially attack other nodes) at any time t. The status of the network system at time t can be represented as

$$(I_1(t), I_2(t), \ldots, I_n(t)),$$

where $I_v(t) = 1$ signifies node v is infected at time t, whereas $I_v(t) = 0$ signifies node v is secure at time t. Also, the infection probability vector is denoted by

$$\mathbf{p}^\top(t) = (p_1(t), p_2(t), \ldots, p_n(t)),$$

where $p_v(t) = \mathbb{P}[I_v(t) = 1]$.

A node v can be infected due to two reasons, namely internal infections caused by attacks from other nodes connected to the node v and external infections caused by external attacks. External infection can occur when the user of the device visits a malicious site, opens links with malicious malware, or connects to an infected USB [3, 14, 15].

For non-Markov model, suppose for any node v, exists D_v infected neighbors at time t:

$$D_v(t) = \sum_{j=1}^{n} a_{v,j} I_j(t),$$

and we have

$$\begin{aligned} \mathbb{E}[D_v] &= \mathbb{E}\left[\sum_{j=1}^{n} a_{v,j} I_j(t)\right] \\ &= \sum_{j=1}^{n} a_{v,j} \cdot p_j(t). \end{aligned}$$

Furthermore, define random variables representing the times-to-infections by infected neighbors, denoted as $\left(Y_{v_1}, Y_{v_2}, \ldots, Y_{v_{D_v}}\right)$. Define another random variable representing the time-to-infection by threats outside the network as Z_v. Therefore, the time to get an infection for node v is

$$T_v = \min\left\{Y_{v_1}, Y_{v_2}, \ldots, Y_{v_{D_v}}, Z_v\right\}.$$

Once node v is infected, the node enters the recovery process with recovery time needed for node v is denoted as R_v. In this paper, we assume that if the node v is infected, then the attacks would stop, and, after the node v is recovered, the attacker(s) would relaunch the attacks. The dependence between the random variables $Y_{v_1}, Y_{v_2}, \ldots, Y_{v_{d_v}}$ is a positive lower orthant dependence (PLOD), which can be understood by considering the dependence between two related nodes is stronger than two unrelated nodes. The PLOD will affect the value of the Kendall's tau coefficient and the parameters of the Archimedean copulas.

The copula function represents how the multivariate distribution function is constructed from its marginal distribution functions [7, 9, 15]. A function $C : [0,1]^k \to [0,1]$ is referred to as a copula of size k if it has the following properties:

- $C(u_1, u_2, \ldots, u_k)$ is increasing in u_z for $z \in \{1, 2, \ldots, k\}$;
- $C(u_1, \ldots, u_{z-1}, 0, u_{z+1}, \ldots, u_k) = 0$ for all $u_j \in [0,1]$ where $j \in \{1, 2, \ldots, k\}$ and $j \neq z$;

- $C(1,\dots,1,u_z,1,\dots,1) = u_z$ for all $u_z \in [0,1]$ where $z \in \{1,2,\dots,k\}$;
- C is k-increasing for all $(u_{1,1},u_{1,2},\dots,u_{1,k})$ and $(u_{2,1},u_{2,2},\dots,u_{2,k})$ in $[0,1]^k$ where $u_{1,j} \leqslant u_{2,j}$, and

$$\sum_{z_1=1}^{2}\cdots\sum_{z_k=1}^{2}(-1)^{\sum_{j=1}^{k} z_j} C(u_{z_1,1}, u_{z_2,2},\dots,u_{z_k,k}) \geqslant 0,$$

for all $j \in \{1,2,\dots,k\}$.

Suppose $X_1, X_2, \dots, X_k$ are random variables with distribution functions $F_1, F_2, \dots, F_k$, the joint distribution function $F(x_1,x_2,\dots,x_k)$ where

$$F(x_1,\dots,x_k) = \mathbb{P}[X_1 \leqslant x_1,\dots,X_k \leqslant x_k],$$

there exists a copula C such that

$$F(x_1,\dots,x_k) = C(F_1(x_1),\dots,F_k(x_k)).$$

Furthermore, the joint-survival function (denoted as $\overline{H}_{d_v}(t,\dots,t)$) is defined by

$$\overline{H}_{d_v}(t,\dots,t) = \hat{C}\left(\overline{F}_1(t),\dots,\overline{F}_{d_v}(t)\right),$$

where $\hat{C}$ is the survival copula of $(Y_{v_1}, Y_{v_2},\dots,Y_{v_{D_v}})$ and

$$\overline{F}_j(t) = \mathbb{P}[Y_{v_j} > t].$$

Archimedean copula with generator function denoted as ψ (a strictly decreasing convex continuous function) is defined by

$$C(u_1,\dots,u_n) = \psi^{-1}\left(\psi(u_1)+\cdots+\psi(u_n)\right),$$

where ψ^{-1} is called the pseudo-inverse function of the copula generator [2,9,15]. A particular case is the Clayton copula where the generator takes the form

$$\psi_\theta(u) = u^{-\theta} - 1.$$

Therefore, the Clayton copula is defined by

$$C(u_1,\dots,u_n) = \left[\sum_{j=1}^{n} u_j^{-\theta} - n + 1\right]^{-\frac{1}{\theta}} \quad (2.2)$$

for $n \geqslant 2$, $\theta > 0$. Another Archimedean copula is the Gumbel Copula where the generator takes the form

$$\psi_\epsilon(u) \;=\; (-\ln u)^\epsilon, \quad \epsilon > 1.$$

Therefore, the Gumbel copula is defined by

$$C(u_1,\dots,u_n) = e^{\left\{-\left[\sum_{j=1}^{n}(-\ln u_j)^\epsilon\right]^{\frac{1}{\epsilon}}\right\}} \quad (2.3)$$

The last Archimedean copula is the Frank Copula where the generator takes the form

$$\psi_\delta(u) \;=\; -\ln\left(\frac{\exp\{-\delta\cdot u\}-1}{\exp\{-\delta\}-1}\right), \delta > 0.$$

Therefore, the Frank copula is defined by

$$C(u_1,\dots,u_n) = -\frac{1}{\delta}\cdot\ln(1+\varphi), \quad (2.4)$$

with $\varphi = \dfrac{\prod_{i=1}^{n}(\exp\{-\delta\cdot u_i\}-1)}{(\exp\{-\delta\}-1)^{n-1}}$. To compare the results of three copulas, we use measures such as the Kendall's tau coefficient with form:

$$\tau(\theta) \;=\; 1 + 4\int_0^1 \frac{\psi_\theta(u)}{\psi'_\theta(u)}\,du.$$

The Kendall's tau coefficient for Clayton copula is defined by

$$\tau(\theta) = \frac{\theta}{\theta+2}, \quad \theta(\tau) = \tfrac{2\tau}{1-\tau}. \quad (2.5)$$

The Kendall's tau coefficient of Gumbel copula is obtained in a similar manner:

$$\tau(\epsilon) = \frac{\epsilon-1}{\epsilon}, \quad \epsilon(\tau) = \tfrac{1}{1-\tau}. \quad (2.6)$$

To obtain the Kendall's tau coefficient of the Frank copula:

$$\tau(\delta) = 1 - \frac{4}{\delta}\int_0^1 \frac{\ln\left(\frac{\exp\{-\delta\cdot u\}-1}{\exp\{-\delta\}-1}\right)}{\frac{\exp(-\delta\cdot u)}{\exp(-\delta\cdot u)-1}}\,du \quad (2.7)$$

The difference between the three copulas lies in the tail dependence of the copulas [1, 2, 9]. Tail dependence measure co-movement between the random variables, therefore selection of copula is based on the level of risk of the insured company:

Table 1. Table of the tail dependence of the three copulas.

Copula	Tail Dependence	Risk Level
Clayton	Left	High
Gumbel	Right	High
Frank	None	Moderate

2.2 Types of loss

The infection process that occurs in the company's network system potentially cause the company to suffer losses. In this paper, there are two types of losses:

1. Loss Due to Infection
The infection process that occurs on the node v can lead to loss of information, data corruption, information leakage, and first-party legal fees. The cost function is defined by

$$\eta_v\left(L_{v,i}\right) = c\cdot L_{v,i},$$

where c represents the cost rate due to infection, $L_{v,i}$ represents the value of missing information of the node v at the i-th time. The Beta distribution is used to model the loss as a proportion of the total capacity of the computer network system [11], with density function:

$$f_{L_v}(x) = \frac{1}{(w_v)^{\alpha+\beta-1}}\cdot\frac{1}{B(\alpha,\beta)}\cdot x^{\alpha-1}\cdot(w_v-x)^{\beta-1},$$

where $0 \leqslant x \leqslant w_v$ and w_v represents the initial wealth (or information) of the node v, and $\alpha, \beta > 0$ are the shape parameters of the Beta distribution [5].

2. Loss Due to Recovery Process
When node v is recovering from the infection, the node cannot operate as usual and will result in opportunity costs. This type of loss depends on the recovery period, and defined by

$$C_v\left(R_{v,i}\right) = c_1\cdot w_v + c_2\cdot R_{v,i},$$

where c_1 represents cost rate based on the initial value, c_2 represents the cost rate of the recovery process, and $R_{v,i}$ represents the recovery period of the node v at the i-th time.

Therefore, the total cost for node v is defined by

$$S_v(t) = \sum_{i=1}^{M_v(t)}\left[\eta_v\left(L_{v,i}\right) + C_v\left(R_{v,i}\right)\right],$$

where $M_v(t)$ represents number of infections occurred on the node v up to time t. Furthermore, the company's total cumulative loss up to time t is defined by

$$\begin{aligned} S(t) &= \sum_{v=1}^{n} S_v(t) \\ &= \sum_{v=1}^{n}\left(\sum_{i=1}^{M_v(t)}\left[\eta_v\left(L_{v,i}\right) + C_v\left(R_{v,i}\right)\right]\right). \end{aligned}$$

Subsequently, the premium charge is estimated using two premium principles, namely standard deviation premium principle and exponential utility premium principle [6]. Estimated premium using the standard deviation premium principle will be denoted as P_1, whereas estimated premium using the exponential utility premium principle will be denoted as P_2. Premium P_1 is defined by

$$P_1(S) = \mathbb{E}[S] + \lambda \cdot \sqrt{\mathrm{Var}(S)},$$

where S represents the cumulative loss random variable, and λ represents risk-loading factor. Premium P_2 is defined by

$$P_2(S) \quad = \quad \frac{\ln\left\{\mathbb{E}\left[e^{\gamma \cdot S}\right]\right\}}{\gamma},$$

where γ represents the degree of risk aversion. It should be noted that this principle is the principle of equivalent utility using the exponential utility function defined by

$$u(\omega) = \begin{cases} \frac{1-\exp(-\gamma \cdot \omega)}{\gamma}, & \text{if } \gamma \neq 0; \\ \omega, & \text{if } \gamma = 0. \end{cases}$$

Here is a cybersecurity insurance contract simulation algorithm with a k days contract period. Cybersecurity insurance contracts are short-term contracts, which generally use quarterly to annual periods with a k value ranging from 90–365 days.

3. Simulation and Analysis

We assume the insured company's network system is depicted in Fig. 1. The network system of the insured company and the loss model parameters are taken from Xu and Hua Model [15]:

$$\begin{aligned}&(\alpha; \beta; c; c_1; c_2; w_v)\\ &= (2; 4; 10^{-3}; 10^{-7}; 5 \cdot 10^{-5}; \text{USD1000}),\end{aligned}$$

where α and β are Beta distribution's parameters. Also, the parameters for the premium calculation are assumed to be

$$(\lambda; \gamma) \quad = \quad (0.2; 0.5)\,.$$

The chosen γ is positive because insurance companies tend to avoid risks so they do not

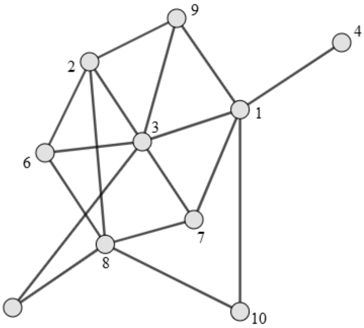

Fig. 1. A company's computer network system.

go bankrupt. We also assume that the infection processes follow Lognormal distributions [13], and the parameters are

$$\begin{aligned}&(\mu_1; \sigma_1^2; \mu_2; \sigma_2^2; \mu_v; \sigma_v^2)\\ &= (1.1094; 1; 0.1931; 1; -0.5; 1), \quad (3.1)\end{aligned}$$

where μ_1 and σ_1^2 represent Lognormal distribution's parameters for times-to-infection from neighbors, μ_2 and σ_2^2 represent Lognormal distribution's parameters for times-to-infection from outside the network, and μ_v and σ_v^2 represent Lognormal distribution's parameters for recovery periods.

The adjacency matrix is formed based on equation Eq. (2.1), as:

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

3.1 Expected number of infections

The Clayton copula parameters are assumed to be $\theta_1 = 2$, which represents a moderate positive dependence, and $\theta_2 = 20$, which indicates a much stronger positive dependence, roughly ten times the strength of the dependence compared to θ_1. Using Eq. (2.5), the Kendall's Tau coefficients for the Clayton copula are

$$(\tau_1, \tau_2) = \left(\frac{1}{2}, \frac{10}{11}\right). \tag{3.2}$$

By substituting Eq. (3.2) to Eq. (2.6) and Eq. (2.7), the Gumbel and Frank copula parameters are obtained as follows.

$$(\epsilon_1; \epsilon_2; \delta_1; \delta_2) = (2; 11; 5.7363; 42.2883).$$

This Monte Carlo simulation is run $\xi = 3000$ times with a contract period of $k = 365$ days (1 year). When the contract starts, all nodes are in a secure state. Simulations are carried out to estimate the premium in the first year. For the following year, premium calculation is based on historical data in the previous year based on risk selection, the amount of claims submitted by the insured company, and the frequency of claims occurring in one period.

Table 2 shows the mean and standard deviation of the number of infections that occur for the Kendall's Tau coefficient $\tau_1 = \frac{1}{2}$. The copula parameters are $\theta_1 = 2$ for the Clayton copula, $\epsilon_1 = 2$ for the Gumbel copula, and $\delta_1 = 5.7363$ for the Frank copula. In Fig. 1, node 3 is the node with most neighbors, whereas node 4 is the node with least neighbors. Therefore, the expected number of infections that occur at node 3 is the most compared to other nodes, while the expected number of infections that occur at node 4 is the least compared to other nodes. The number of infections in the other nodes is between the number of infections of these two particular nodes. As the Clayton copula has strong lower tail dependence, nodes with fewer neighbors might show higher infections because the infections at their few neighboring nodes are more likely to be correlated, especially in the early stages. Meanwhile, the Gumbel copula has strong upper tail dependence. At first, the infection spreads slowly across the network as it originates from an external source. However, its spread picks up pace as it reaches nodes with more neighbors, which are considered high-risk nodes. The infection then accelerates as it continues to affect other high-risk nodes. On the other hand, due to the Frank copula not prioritizing the lower or upper tails, the infection spreads more evenly across the network.

Table 2. Table of the number of infections for the three copulas with $\tau_1 = \frac{1}{2}$.

Node	Clayton ($\theta_1 = 2$)		Gumbel ($\epsilon_1 = 2$)		Frank ($\delta_1 = 5.7363$)	
	$\mathbb{E}[N]$	$\sqrt{\mathrm{Var}(N)}$	$\mathbb{E}[N]$	$\sqrt{\mathrm{Var}(N)}$	$\mathbb{E}[N]$	$\sqrt{\mathrm{Var}(N)}$
1	76.44	6.79	76.34	6.57	76.32	6.68
2	76.37	6.70	76.44	6.64	76.33	6.63
3	76.59	6.67	76.52	6.76	76.56	6.89
4	76.16	6.73	76.25	6.80	76.21	6.82
5	76.29	6.58	76.44	6.74	76.55	6.74
6	76.57	6.69	76.40	6.72	76.38	6.66
7	76.39	6.80	76.46	6.82	76.34	6.70
8	76.56	6.70	76.34	6.86	76.38	6.83
9	76.36	6.56	76.41	6.74	76.37	6.78
10	76.19	6.69	76.34	6.91	76.33	6.60

(Number of Infections N)

From Table 2, $\mathbb{E}[N]$ for node 3 is 76.59 for simulation using Clayton copula, 76.52 using Gumbel copula, and 76.56 uses Frank copula. This number means that from the 3000 simulations, the expected number of infections that occurred at node 3 is about 77 times for the three copulas. Table 3 shows the number of infections, N, for the Kendall's Tau coefficient $\tau_2 = \frac{10}{11}$. The copula parameters are $\theta_2 = 20$ for the Clayton copula, $\epsilon_2 = 11$ for the Gumbel copula, and $\delta_2 = 42.2883$ for the Frank copula. From Table 3, as in Table 2, the expected number of infections that occur at node 3 is the most compared to other nodes, and vice versa for node 4. In addition, when compared with Table 2, the change in the Kendall's Tau coefficient affects the num-

Table 3. Table of the number of infections for the three copulas with $\tau_2 = \frac{10}{11}$.

Node	Number of Infections N					
	Clayton ($\theta_2 = 20$)		Gumbel ($\epsilon_2 = 11$)		Frank ($\delta_2 = 42.2883$)	
	$\mathbb{E}[N]$	$\sqrt{\mathrm{Var}(N)}$	$\mathbb{E}[N]$	$\sqrt{\mathrm{Var}(N)}$	$\mathbb{E}[N]$	$\sqrt{\mathrm{Var}(N)}$
1	76.36	6.71	76.44	6.72	76.54	6.73
2	76.28	6.66	76.48	6.67	76.43	6.80
3	76.54	6.93	76.54	6.74	76.63	6.80
4	76.02	6.76	76.24	6.68	76.20	6.73
5	76.05	6.70	76.34	6.79	76.33	6.74
6	76.41	6.77	76.50	6.77	76.37	6.80
7	76.32	6.72	76.41	6.83	76.43	6.65
8	76.41	6.87	76.50	6.90	76.36	6.68
9	76.50	6.66	76.39	6.78	76.27	6.68
10	76.28	6.79	76.20	6.83	76.36	6.72

ber of infections, although insignificantly. The higher Kendall's Tau coefficient causes the times-to-infections for expected number of infections for $\tau_2 = \frac{10}{11}$ is less than $\tau_1 = \frac{1}{2}$ for most nodes.

3.2 Predicted losses

Next, the mean and standard deviation of the company's losses for the Kendall's Tau coefficient $\tau_1 = \frac{1}{2}$ is shown at Table 4. The expected loss for node 3 is

Table 4. Table of the company's losses for the three copulas with $\tau_1 = \frac{1}{2}$.

Node	Company's Losses S (USD)					
	Clayton ($\theta_1 = 2$)		Gumbel ($\epsilon_1 = 2$)		Frank ($\delta_1 = 5.7363$)	
	$\mathbb{E}[S]$	$\sqrt{\mathrm{Var}(S)}$	$\mathbb{E}[S]$	$\sqrt{\mathrm{Var}(S)}$	$\mathbb{E}[S]$	$\sqrt{\mathrm{Var}(S)}$
1	25.57	2.75	25.49	2.70	25.44	2.72
2	25.52	2.72	25.51	2.70	25.46	2.69
3	25.65	2.68	25.64	2.69	25.59	2.80
4	25.22	2.76	25.28	2.77	25.40	2.74
5	25.46	2.67	25.49	2.74	25.59	2.76
6	25.52	2.72	25.47	2.70	25.51	2.69
7	25.50	2.74	25.57	2.73	25.44	2.73
8	25.60	2.74	25.50	2.78	25.48	2.79
9	25.48	2.66	25.52	2.70	25.48	2.79
10	25.43	2.72	25.49	2.85	25.43	2.65
System	254.93	6.27	254.94	6.34	254.82	6.40

higher than for other nodes, and vice versa for node 4. From this, it can be concluded that the dependency effect also affects the cumulative loss. Table 5 shows the company's losses (S) for the Kendall's Tau coefficient $\tau_2 = \frac{10}{11}$. As in Table 4, the expected loss for node 3 is higher than for other nodes, and vice versa for node 4. In addition, when compared with Table 4, the higher Kendall's Tau coefficient causes an insignificant decrease in expected loss for most nodes. However, this is not the case

Table 5. Table of the company's losses for the three copulas with $\tau_2 = \frac{10}{11}$.

Node	Company's Losses S (USD)					
	Clayton ($\theta_2 = 20$)		Gumbel ($\epsilon_2 = 11$)		Frank ($\delta_2 = 42.2883$)	
	$\mathbb{E}[S]$	$\sqrt{\mathrm{Var}(S)}$	$\mathbb{E}[S]$	$\sqrt{\mathrm{Var}(S)}$	$\mathbb{E}[S]$	$\sqrt{\mathrm{Var}(S)}$
1	25.53	2.77	25.53	2.69	25.43	2.72
2	25.46	2.72	25.45	2.68	25.53	2.78
3	25.62	2.82	25.59	2.78	25.56	2.78
4	25.36	2.74	25.19	2.70	25.21	2.73
5	25.37	2.73	25.55	2.73	25.47	2.72
6	25.50	2.76	25.52	2.70	25.55	2.83
7	25.47	2.75	25.54	2.73	25.52	2.67
8	25.49	2.83	25.54	2.78	25.49	2.72
9	25.52	2.71	25.50	2.77	25.46	2.72
10	25.43	2.77	25.46	2.76	25.52	2.68
System	254.75	6.42	254.86	6.33	254.74	6.28

for node 4 as it is the only node with the least number of neighboring nodes.

By comparing Tables 4 and 5, it is observed that the connection between infection processes impacts how an infection spreads and the losses it causes. A higher Kendall's Tau coefficient means stronger positive dependence between infection processes between nodes [15]. This implies that the waiting time to infection increases, resulting in fewer nodes being infected within the contract period. Fewer nodes infected naturally leads to lower losses.

3.3 Estimation of premium charges

The premiums are calculated based on two principles: P_1 for the standard deviation principle and P_2 for the exponential utility principle for the Kendall's Tau coefficient $\tau_1 = \frac{1}{2}$. From Table 6, the esti-

Table 6. Table of premium charge for the three copulas with $\tau_1 = \frac{1}{2}$.

Node	Clayton ($\theta_1 = 2$)		Gumbel ($\epsilon_1 = 2$)		Frank ($\delta_1 = 5.7363$)	
	P_1(USD)	P_2(USD)	P_1(USD)	P_2(USD)	P_1(USD)	P_2(USD)
1	26.12	27.50	26.03	27.51	25.98	27.40
2	26.06	27.38	26.05	27.43	25.99	27.33
3	26.18	27.69	26.18	27.70	26.15	27.65
4	25.77	27.17	25.83	27.24	25.95	27.25
5	25.99	27.31	26.03	27.48	26.14	27.60
6	26.06	27.49	26.01	27.42	26.05	27.38
7	26.04	27.45	26.11	27.53	25.98	27.45
8	26.15	27.61	26.05	27.69	26.04	27.51
9	26.01	27.27	26.06	27.37	26.04	27.44
10	25.97	27.43	26.06	27.56	25.96	27.29
System	256.18	264.38	256.21	266.54	256.09	264.34

mated premiums for node 3 is more expensive than for other nodes, and vice versa for node 4. Therefore, it is concluded that the positive dependence also insignificantly affects the premium charge. In the next table, we compare estimated premiums P_1 and P_2 for the Kendall's Tau coefficient $\tau_2 = \frac{10}{11}$. From Table 7, the estimated premiums for node 3 are more expensive than for other nodes, and vice versa for node 4. In addition, when compared with Table 6, the higher Kendall's Tau coefficient causes an insignificant decrease in estimated premiums most nodes.

Table 7. Table of premium charge for the three copulas with $\tau_2 = \frac{10}{11}$.

Node	Clayton ($\theta_2 = 20$)		Gumbel ($\epsilon_2 = 11$)		Frank ($\delta_2 = 42.2883$)	
	P_1(USD)	P_2(USD)	P_1(USD)	P_2(USD)	P_1(USD)	P_2(USD)
1	26.08	27.42	26.06	27.35	25.97	27.43
2	26.01	27.39	25.98	27.39	26.08	27.53
3	26.18	27.67	26.14	27.66	26.11	27.57
4	25.90	27.11	25.73	27.22	25.76	27.32
5	25.92	27.30	26.09	27.43	26.01	27.55
6	26.05	27.51	26.06	27.42	26.11	27.49
7	26.02	27.49	26.09	27.52	26.05	27.33
8	26.05	27.64	26.10	27.51	26.04	27.38
9	26.07	27.41	26.06	27.65	26.01	27.40
10	25.99	27.61	26.02	27.47	26.06	27.42
System	256.03	264.29	256.13	265.60	255.99	264.29

Furthermore, in Tables 6 and 7, the infection times modeled using the Gumbel copula resulted in higher premium charge than the other two copulas, calculated by both principles, even though the differences are not material. Therefore, the infection times in the first year contract is modeled using the Gumbel copula. After a year, considering the actual number of infections and the actual loss, the insurance company gets to decide whether or not to lower the premium. If the insurance company chooses to lower the premium, the infection process is modeled using the Clayton or Frank copula for the following year. The selection of the Archimedean copula which tends to produce the highest premium charge is based on the possibility that the insured company is more interested in extending the policy contract with the same or cheaper value.

3.4 Analysis of changes in infection time between neighboring nodes

The effects of the changes in infection process between nodes within the internal system, denoted by parameter μ, are also analyzed. Initially, due to the assumption of infection processes follow the Log-normal distributions with the μ_1 and σ_1^2 parameters from Eq. (3.1), the expected time for node j to infect node v, denoted by $\mathbb{E}[Y_{v,j}]$ with $v, j \in \{1, 2, \ldots, 10\}; v \neq j$.

$$\mathbb{E}[Y_{v,j}] = \exp\left(\mu_1 + \frac{\sigma_1^2}{2}\right) = 5.$$

Therefore, the expected infection time is halved and doubled to examine its impact on the simulation. Specifically, in the first simulation, $\mathbb{E}[Y_{v,j}] = 2.5$, and in the second simulation, $\mathbb{E}[Y_{v,j}] = 10$. The resulting parameter sets, incorporating the new μ_1 values while keeping $\sigma_1^2, \mu_2, \sigma_2^2, \mu_v, \sigma_v^2$ constant, are

$$\left(\mu_1; \sigma_1^2; \mu_2; \sigma_2^2, \mu_v; \sigma_v^2\right) = (0.4163; 1; 0.1931; 1; -0.5; 1),$$

and

$$\left(\mu_1; \sigma_1^2; \mu_2; \sigma_2^2, \mu_v; \sigma_v^2\right) = (1.8026; 1; 0.1931; 1; -0.5; 1).$$

The following table shows the percentage of changes of the expected number of infections, $\mathbb{E}[N]$, by calculating:

$$\text{Change} = \text{Final value} - \text{Initial value},$$

$$\%\text{Change} = \frac{\text{Change}}{\text{Initial value}} \times 100\%,$$

using the initial values from Table 2 and the final values from the simulations based on the new set of parameters with Kendall's Tau coefficient of $\tau_1 = \frac{1}{2}$. From Table 8, a

Table 8. Table of percentage of changes of the expected number of infections.

Percentage of Changes of the Expected Number of Infections $\mathbb{E}[N]$ (%)						
Node	$\mu_1 = 0.4163$			$\mu_1 = 1.8026$		
	Clayton	Gumbel	Frank	Clayton	Gumbel	Frank
1	−0.61	−0.48	−0.38	−0.09	0.13	0.33
2	−1.68	−0.51	−0.39	0.29	−0.07	0.07
3	−0.37	−0.34	−0.54	−0.13	0.27	0.25
4	−0.08	−0.62	−0.18	0.38	0.20	0.30
5	−0.29	−0.58	−1.07	0.19	0.05	−0.09
6	−0.86	−0.32	−0.68	−0.02	0.31	0.22
7	0.71	−0.67	−0.75	0.06	−0.03	0.15
8	−0.62	−0.40	−0.58	−0.13	0.02	−0.05
9	−0.34	−0.53	−0.37	−0.11	0.17	0.07
10	−0.51	−0.58	−0.44	0.17	0.19	0.14

decrease in μ_1 causes a decrease in the expected number of infections which can be seen from the percentage of changes that are negative, while an increase in μ_1 causes an increase in the expected number of infections which can be seen from the percentage of changes that are positive. When μ_1 is decreased, the overall expected number of infections decreased by 0.46% for the Clayton copula, 0.50% for the Gumbel copula, and 0.54% for the Frank copula. On the other hand, when μ_1 is increased, the overall expected number of infections increased by 0.06% for the Clayton copula, 0.12% for the Gumbel copula, and 0.14% for the Frank copula. From the small percentage values, we conclude that the changes in μ_1 does not affect the expected number of infections. In the next table, we observe the changes in μ_1 also affects the expected loss.

From Table 9, a decrease in μ_1 also causes a decrease in the expected loss which can be seen from the percentage of changes that are negative, and vice versa for an increase in μ_1. When μ_1 is decreased, the network system's expected loss decreased by 0.37% for the Clayton copula, 17.75% for the Gumbel copula, and 0.47% for the Frank copula.

On the other hand, when μ_1 is increased, the network system's expected loss increased by 0.08% for the Clayton copula, 0.11% for the Gumbel copula, and 0.17% for the Frank copula. From the small percentage values, we conclude that the changes in μ_1 does not affect the expected loss. Next, we investigate whether the changes in μ_1 also affects the estimated premium using the standard deviation principle. From Table 10, a decrease in μ_1 also causes a decrease in the estimated premium using the standard deviation principle which can be seen from the percentage of changes that are negative, and vice versa for an increase in μ_1. When μ_1 is decreased, the premium charge decreased by 0.37% for the Clayton copula, 0.46% for the Gumbel copula, and 0.43% for the Frank copula. On the other hand, when μ_1 is increased, the premium charge increased by 0.09% for the Clayton copula, 0.12% for the Gumbel copula, and 0.17% for the Frank copula. From the small percentage values, the changes in μ_1 does not affect the estimated premium using the standard deviation prin-

Table 9. Table of percentage of changes of expected loss.

Percentage of Changes of Expected Loss $\mathbb{E}[S]$ (%)						
Node	$\mu_1 = 0.4163$			$\mu_1 = 1.8026$		
	Clayton	Gumbel	Frank	Clayton	Gumbel	Frank
1	−0.65	−0.30	0.04	−0.25	0.11	0.49
2	−0.26	−0.35	−0.31	0.09	−0.18	−0.02
3	−0.48	−0.33	−0.54	−0.07	0.24	0.07
4	−0.04	−0.61	−0.14	0.31	0.07	0.25
5	−0.15	−0.45	−1.05	0.23	0.25	−0.13
6	−0.63	−0.28	−0.72	0.23	0.61	0.24
7	−0.34	−0.85	−0.59	0.21	−0.06	0.19
8	−0.54	−0.49	−0.48	−0.18	−0.25	0.07
9	−0.26	−0.50	−0.09	0.13	0.09	0.30
10	−0.39	−0.51	−0.51	0.08	0.25	0.25
System	−0.37	−0.47	−0.44	0.08	0.11	0.17

Table 10. Table of percentage of changes of premium P_1.

Percentage of Changes of P_1 (%)						
Node	$\mu_1 = 0.4163$			$\mu_1 = 1.8026$		
	Clayton	Gumbel	Frank	Clayton	Gumbel	Frank
1	−0.66	−0.21	0.08	−0.24	0.13	0.49
2	−0.27	−0.30	−0.31	0.07	−0.17	0.03
3	−0.44	−0.28	−0.54	−0.04	0.26	0.00
4	−0.04	−0.62	−0.12	0.26	0.04	0.29
5	−0.07	−0.42	−1.02	0.27	0.23	−0.15
6	−0.60	−0.26	−0.70	0.24	0.60	0.27
7	−0.33	−0.83	−0.58	0.21	−0.02	0.14
8	−0.56	−0.52	−0.52	−0.14	−0.28	0.03
9	−0.19	−0.49	−0.11	0.17	0.15	0.24
10	−0.34	−0.58	−0.41	0.07	0.19	0.27
System	−0, 37	−0.46	−0.43	0.09	0.12	0.17

ciple. Lastly, we will see if the changes in μ_1 also affects the estimated premium using the exponential utility principle.

Table 11. Table of percentage of changes of premium P_2.

Node	Percentage of Changes of P_2 (%)					
	$\mu_1 = 0.4163$			$\mu_1 = 1.8026$		
	Clayton	Gumbel	Frank	Clayton	Gumbel	Frank
1	−0.68	−0.17	0.55	−0.12	0.22	0.24
2	−0.35	0.15	0.06	0.52	0.54	0.47
3	0.01	−0.02	−0.60	0.56	1.11	−0.36
4	0.11	−0.39	0.29	0.37	0.06	0.59
5	0.21	−0.13	−0.74	0.56	0.38	−0.38
6	−0.58	−0.64	−0.72	0.10	0.37	0.96
7	−0.09	−1.08	−0.95	0.40	0.35	−0.33
8	−0.70	−1.14	−0.63	−0.06	−0.66	−0.37
9	0.14	−0.36	0.12	0.70	0.36	0.08
10	0.27	−0.77	0.15	−0.32	−0.06	0.45
System	−0, 46	−0.72	−0.43	0.15	0.25	0.17

From Table 11, a decrease in μ_1 also causes a decrease in the estimated premium using the exponential utility principle which can be seen from the percentage of changes that are negative, and vice versa for an increase in μ_1. When μ_1 is decreased, the premium charge decreased by 0.46% for the Clayton copula, 0.72% for the Gumbel copula, dan 0.43% for the Frank copula. On contrary, when μ_1 is increased, the premium charge increased by 0.15% for the Clayton copula, 0.25% for the Gumbel copula, and 0.17% for the Frank copula. Due to the small percentage values, the changes in μ_1 does not affect the estimated premium using the exponential utility principle.

4. Conclusion

When taking dependencies into account in the cybersecurity insurance model, nodes (representing devices) which connected with a number of nodes are more likely to be infected than a node connected with fewer number of nodes. Also, an increase in Kendall's tau coefficient causes a decrease in number of infections, losses, and premium charge insignificantly. Modeling infection times between nodes using Gumbel copula function generates higher premiums than other copula functions (even though the differences are not material), so it is better to use Gumbel copula function for modeling cybersecurity insurance in the first contract period because insured companies will be more interested in extending the contract if the next year's premium charge less than or equal to previous year's premium. In addition, changes in parameter of times-to-infections from neighbors does not affect expected number of infections, expected losses, and premium charge.

For further research, we suggest to use copula types other than the Archimedean class to model the dependence between nodes. Besides that, we also suggest to analyze the sensitivity of the simulation results and charged premiums.

References

[1] Andrievskaya IK, Penikas HI. Copula-application to modelling Russian banking system capital adequacy according to Basel II IRB-approach. Model Assisted Statistics and Applications. 2012 Jan 1; 7(4):267-80.

[2] Gordeev VA, Knyazev AG, Shemyakin AE. Selection of copula model for inter-market dependence. Model Assisted Statistics and Applications. 2012 Jan 1; 7(4):315-25.

[3] Griscioli F, Pizzonia M. USBCaptchaIn: Preventing (un)conventional attacks from promiscuously used USB devices in industrial control systems. Journal of Computer Security. 2021 Jan 1; 29(1):51-76.

[4] Jethva B, Traoré I, Ghaleb A, Ganame K, Ahmed S. Multilayer ransomware detection using grouped registry key operations, file entropy and file signature monitoring. Journal of Computer Security. 2020 Jan 1; 28(3):337-73.

[5] Johnson NL, Kotz S, Balakrishnan N. Continuous univariate distributions, volume 2. John Wiley & sons; 1995 May 8.

[6] Kaas R, Goovaerts M, Dhaene J, Denuit M. Modern actuarial risk theory: using R. Springer Science & Business Media; 2008 Aug 17.

[7] Kim JM, Sungur EA, Choi T, Heo TY. Generalized bivariate copulas and their properties. Model Assisted Statistics and Applications. 2011 Jan 1; 6(2):127-36.

[8] Murali R, Shunmuga Velayutham C. A preliminary investigation into automatically evolving computer viruses using evolutionary algorithms. Journal of Intelligent & Fuzzy Systems. 2020 Jan 1; 38(5):6517-26.

[9] Nelsen RB. An introduction to copulas. Springer; 2006.

[10] Rosen KH. Discrete mathematics and its applications. 8th ed. New York: McGraw-Hill; 2018.

[11] Ross S. A first course in probability. 10th edition. New Jersey: Pearson; 2020.

[12] Schwartz GA, Sastry SS. Cyber-insurance framework for large scale interdependent networks. In Proceedings of the 3rd international conference on High confidence networked systems 2014 Apr 15 (pp. 145-154).

[13] Shimizu K, Edwin LC. History, genesis, and properties. InLognormal distributions 2018 May 2 (pp. 1-25). Routledge.

[14] Sibi Chakkaravarthy S, Sangeetha D, Venkata Rathnam M, Srinithi K, Vaidehi V. Futuristic cyber-attacks. International Journal of Knowledge-based and Intelligent Engineering Systems. 2018 Jan 1; 22(3):195-204.

[15] Xu M, Hua L. Cybersecurity insurance: Modeling and pricing. North American Actuarial Journal. 2019 Apr 3; 23(2):220-49.