

An Extended Stochastic Model for Quantitative Security Analysis of Networked Systems

Maochao Xu and Shouhuai Xu

Abstract. Quantitative security analysis of networked computer systems has been an open problem in computer security for decades. Recently, a promising approach was proposed in [Li et al. 11], which, however, made some strong assumptions including the exponential distribution of, and the independence among, the relevant random variables. In this paper, we substantially weaken these assumptions while offering, in addition to the same types of analytical results as in [Li et al. 11], methods for obtaining the desired security quantities in practice. Moreover, we investigate the problem from a higher-level abstraction, which also leads to both analytical results and practical methods for obtaining the desired security quantities. These should represent a significant step toward ultimately solving the problem of quantitative security analysis of networked computer systems.

1. Introduction

Quantitative security analysis of networked computer systems from a whole-system, rather than from an individual-component, perspective is one of the most important open problems in cyber security. Recently, a promising stochastic

process approach to tackling this problem was proposed in [Li et al. 11]. In this approach, a networked computer system is modeled as a complex network, or graph, $G = (V, E)$, where V is the vertex, or node, set and E is the edge set over which direct attacks can be carried out. Then a stochastic abstraction of the interactions between the attacker and the defender taking place in G is considered. The main research task is to derive the following two security measures:

- $q(v)$: the probability that a node $v \in V$ is compromised (or attacked) in the steady state.
- q : the probability that a uniformly chosen node in G is compromised (or attacked) in the steady state.

The quantity q is particularly important for decision-making because it captures or reflects the global security level of a networked system in a useful way. For example, given q , the defender can design or deploy appropriate cryptographic or security schemes to mitigate the damage of the attacks. (Some extensions to q will be discussed in Section 7.)

1.1. Our Contribution

In this paper, we make substantial improvements upon [Li et al. 11], which assumed that the relevant random variables follow the exponential distribution and are independent of each other. Specifically, we consider two cases: (I) the case that the lower-level attack–defense process is observed, which is the same as in [Li et al. 11] but with much weaker assumptions; (II) the case that the higher-level alternating renewal process, but not the lower-level attack–defense process, is observed. The latter case was not investigated in [Li et al. 11] and may be more realistic.

In the case that the lower-level attack–defense process is observed (Section 4), we present new analytical results for deriving $q(v)$ under weaker assumptions, where $q(v)$ is the probability that an arbitrary node $v \in V$ is compromised in the steady state. On the one hand, these new analytical results lead to useful methods for obtaining the global security measure q in practice:

- (i) In the case that G is a regular graph with moderate node degree and the dependence between the nodes can be ignored, we derive a closed-form solution for q .

- (ii) In the case that G is an arbitrary graph (including the case of a regular graph but with large node degree) and the dependence between the nodes can be ignored, we give an approximation solution for q .
- (iii) In the case that G is an arbitrary graph and the dependence between the nodes cannot be ignored, we describe a statistical solution for q .

On the other hand, the new analytical results lead to bounds of q both in the case that G is regular graph and in the case that G is an arbitrary graph. Such results are interesting and useful because they demand much less information. The upper bounds are especially useful in decision-making because they capture a sort of worst-case scenario from the defender's perspective.

In the case that the lower-level attack–defense process is not observed, but the higher-level alternating renewal process is observed (Section 5), we present new analytical results for estimating $q(v)$. These analytical results lead to statistical methods for obtaining the global security measure q in practice. Note that this case was not investigated in [Li et al. 11], and is harder to deal with analytically due to the lack of sufficient lower-level information.

Note that the results in [Li et al. 11] and in the present paper are applicable to both undirected and directed graphs. To simplify notation, we will focus on the case of undirected graphs, while noting that when we adapt the results to the case of directed graphs, we need to replace “node degree” with “node in-degree” (a node v 's in-degree is the number of nodes that point to v).

1.2. Outline of the Paper

The rest of the paper is organized as follows. In Sections 2 and 3, we briefly review the model and results in [Li et al. 11] and some probabilistic notions, respectively. In Section 4, we present new results for the case in which the lower-level attack–defense process is observed. In Section 5, we present new results for the case in which the higher-level alternating renewal process is observed. In Section 6, we review related previous studies. In Section 7, we discuss the utility of the stochastic process approach and future research directions toward the ultimate goal. In Section 8, we conclude the paper.

1.3. Notation

The principal notation used throughout the paper is summarized as follows.

$G = (V, E)$: This is a graph that abstracts a networked system with $n = |V|$ nodes of average node degree μ .

$\text{deg}(v), q(v), K(v)$: In the steady state, a node $v \in V$ is compromised with probability $q(v)$, and among the v 's $\text{deg}(v)$ neighbors, $K(v)$ neighbors are compromised.

q, K : In the steady state, the probability that a uniformly chosen node is compromised is denoted by q , and the number of a uniformly chosen node's compromised neighbors is denoted by K , where $q = \sum_{v \in V} q(v)/n$ and $K = \sum_{v \in V} K(v)/n$.

$\{(S_j(v), C_j(v)); j \geq 1\}$: This is the sequence of node v 's cycles corresponding to the alternating renewal process, where each cycle consists of the time interval $S_j(v)$ during which v is secure and the time interval $C_j(v)$ during which v is compromised.

2. A Brief Review of the Model and Results in [Li et al. 11]

We start with a brief review of [Li et al. 11]. As mentioned above, a networked system is represented as a complex network $G = (V, E)$, where V is the set of n nodes or vertices, namely $|V| = n$, that abstract (for example) computers, and E is the set of edges that abstract the internode communications or interactions that can be exploited to carry out direct attacks. At any point in time, a node $v \in V$ is either secure or compromised. These abstractions are not new. The novelty of [Li et al. 11] is the attack–defense process over G , which is specified by the following random variables:

- X_1 : The time a secure node becomes compromised directly by the attacker outside of G . This models “drive-by download”-like attacks [Provos et al. 07].
- $X_{2,i}$: The time a secure node becomes compromised because of its i th compromised neighbor, $1 \leq i \leq \text{deg}(v)$. This models attacks such as standard malware spreading within a network.
- Y_1 : The time a compromised node becomes secure again because the compromise has been detected and cured. This models reactive defense.
- Y_2 : The time a compromised node becomes secure again for any other reason (e.g., patching or reinstalling the software system even if the compromise was not detected). This models proactive defense.

As mentioned above, the main security measure of interest is q , the probability that a uniformly chosen node is compromised in the steady state. This immediately leads to the more intuitive security measure $q \cdot n$, namely the expected

number of compromised nodes in the steady state. Since in the steady state an arbitrary node $v \in V$ is compromised with probability $q(v)$, we have

$$q = \frac{\sum_{v \in V} q(v)}{n}.$$

The above is perhaps the most important conceptual contribution of [Li et al. 11], which is very simple but insightful because q captures the global property of $G = (V, E)$ in the presence of attacks and defenses. The main research task is to derive an expression for q , which turns out to be a difficult problem because the states of the nodes are random variables that are dependent on each other according to the complex structure of G . As a first step, [Li et al. 11] focused on the special case in which G is regular graph, namely all nodes have the same degree and the $q(v)$'s are all the same (i.e., independent of v). Based on the following assumption, two main results were offered in [Li et al. 11].

Assumption 2.1. [Li et al. 11]

1. Assumption on the distributions of the random variables:
 - X_1 follows the exponential distribution with rate α .
 - The $X_{2,i}$'s follow the exponential distribution with rate γ .
 - Y_1 follows the exponential distribution with rate β .
 - Y_2 follows the exponential distribution with rate η .
2. Assumption on the independence between the random variables: X_1 , Y_1 , Y_2 , and the $X_{2,i}$'s are independent of each other.

Theorem 2.2. [Li et al. 11] *Suppose G is regular graph. Under Assumption 2.1, we have*

$$q = \frac{1}{1+m}, \quad \text{where } m = \mathbb{E} \left[\frac{\beta + \eta}{\alpha + \gamma K} \right],$$

and the random variable K indicates the number of compromised neighbors of a node.

Theorem 2.3. [Li et al. 11] *For a regular graph G of node degree μ , under Assumption 2.1, we have*

$$\frac{\alpha}{\alpha + \beta + \eta} \leq q \leq \frac{\alpha + \gamma\mu}{\alpha + \beta + \eta + \gamma\mu}.$$

Simulation studies in [Li et al. 11] showed that the upper bound is reasonably tight not only for regular graphs, but also for Erdős–Rényi random graphs and power-law graphs (in the latter two cases, μ is naturally the average node degree). This means that we can plausibly use the upper bound in decision-making when there is a lack of information.

3. A Brief Review of Some Probabilistic Notions Used in This Paper

3.1. Some Probabilistic Distributions

A random variable Z is said to have an exponential distribution with rate $\lambda > 0$ if the distribution function is

$$F(z) = 1 - \exp\{-\lambda z\}.$$

This distribution has the memoryless property, which played an essential role in [Li et al. 11]. The exponential distribution is a special case of the Weibull distribution with distribution function

$$F(z) = 1 - \exp\{-(\gamma z)^\alpha\},$$

where $\gamma > 0$ and $\alpha > 0$ are scale and shape parameters, respectively.

The Lomax (or Pareto II) distribution has distribution function

$$F(z) = 1 - \left(1 + \frac{z}{\beta}\right)^{-\eta},$$

where $\beta > 0$ is the scale parameter and $\eta > 0$ is the shape parameter [Johnson et al. 94]. This distribution is interesting because it can model heavy-tailed phenomena.

3.2. Some Notions of Nonindependence between Random Variables

Two random variables Z_1 and Z_2 are said to be positively quadrant dependent (PQD) if for all $z_1, z_2 \in \mathbb{R}$,

$$P(Z_1 > z_1, Z_2 > z_2) \geq P(Z_1 > z_1)P(Z_2 > z_2).$$

In general, a family $\{Z_1, \dots, Z_n\}$ of random variables is said to be positively upper orthant dependent (PUOD) if for all $z_i \in \mathbb{R}$, $i = 1, \dots, n$,

$$P(Z_1 > z_1, \dots, Z_n > z_n) \geq \prod_{i=1}^n P(Z_i > z_i).$$

A family $\{Z_1, \dots, Z_n\}$ of random variables is said to be positively associated (PA) if

$$\text{Cov}(f(Z_1, \dots, Z_n), g(Z_1, \dots, Z_n)) \geq 0$$

for all coordinatewise nondecreasing functions f, g on \mathbb{R}^n as long as the covariance exists. Finally, we recall the multivariate Marshall–Olkin model, which offers a good deal of flexibility in modeling dependence structures [Marshall and Olkin 67, Li 08].

Let $\{E_B, B \subseteq \{1, 2, \dots, b\}\}$ be a sequence of independent, exponentially distributed random variables, where E_B has mean $1/\lambda_B$. Let

$$T_j = \min\{E_B : j \in B\}, \quad j = 1, \dots, b.$$

The joint distribution of $T = (T_1, \dots, T_b)$ is called the Marshall–Olkin exponential distribution with parameters $\{\lambda_B, B \subseteq \{1, 2, \dots, b\}\}$.

It is known [Barlow and Proschan 81, Theorem 3.2, Chapter 2] that $\text{PA} \implies \text{PUOD(PQD)}$, meaning that PA is (not necessarily strictly) stronger than PUOD (PQD). It is also known that the Marshall–Olkin exponential distribution is PA, and hence PUOD [Mueller and Stoyan 02].

4. New Results for the Case in Which the Attack–Defense Process Is Observed

In this section, we present new analytical results for deriving $q(v)$, methods for deriving q in practice, and bounds of q in the absence of sufficient information.

4.1. New Analytical Results for Deriving $q(v)$ under Weaker Assumptions

We seek to substantially weaken Assumption 2.1 while considering more general graph structures (i.e., not just regular graphs). We start with the following assumption.

Assumption 4.1. (A substantially weaker assumption.)

1. Rather than assuming that X_1 , the $X_{2,i}$'s, Y_1 , and Y_2 are exponential random variables, we assume the following:
 - X_1 follows an arbitrary continuous distribution F_1 with finite mean.
 - $X_{2,i}$ follows an arbitrary continuous distribution $F_{2,i}$ with finite mean, where the distributions may be different for different i 's.

- Y_1 follows an arbitrary continuous distribution G_1 with finite mean.
 - Y_2 follows an arbitrary continuous distribution G_2 with finite mean.
2. Rather than assuming that X_1 , the $X_{2,i}$'s, Y_1 , and Y_2 are independent of each other, we assume that X_1 , the $X_{2,i}$'s, Y_1 , and Y_2 can be mutually dependent.

In practice, Assumption 4.1 can be tested using some nonparametric methods (cf. [Hollander and Douglas 99, Chapter 8]).

Theorem 4.2. Denote the joint survival function for $\{X_1, X_{2,1}, \dots, X_{2,n}\}$ by

$$\bar{H}_{n+1}(x_1, \dots, x_{n+1}) = P(X_1 > x_1, X_{2,1} > x_2, \dots, X_{2,n} > x_{n+1}),$$

and the joint survival function for Y_1 and Y_2 by

$$\bar{G}(y_1, y_2) = P(Y_1 > y_1, Y_2 > y_2).$$

Under Assumption 4.1, the probability $q(v)$ that an arbitrary node $v \in \mathbf{V}$ is compromised in the steady state is

$$q(v) = \frac{1}{1+m}, \tag{4.1}$$

where

$$m = E \left[\int_0^\infty \bar{H}_{K(v)+1}(x, \dots, x) dx \right] / \int_0^\infty \bar{G}(x, x) dx,$$

and $K(v)$ is the number of compromised neighbors of node v .

Proof. Observe that in general, the states of node $v \in \mathbf{V}$ can be modeled as an alternating renewal process, with each cycle composed of two states: secure with random time $S_j(v)$ and compromised with random time $C_j(v)$ for $j = 1, 2, \dots$. Under Assumption 4.1, for each cycle j we can express $(S_j(v), C_j(v))$ as

$$S_j(v) = S(v) = \min \{X_1, \{X_{2,i}; 1 \leq i \leq K(v)\}\}, \quad C_j(v) = C(v) = \min \{Y_1, Y_2\},$$

where $K(v)$ is the number of v 's compromised neighbors. Hence,

$$\begin{aligned} E[S(v)] &= E[\min\{X_1, X_{2,1}, \dots, X_{2,K(v)}\}] \\ &= E[E[\min\{X_1, X_{2,1}, \dots, X_{2,K(v)}\} \mid K(v)]] \\ &= E\left[\int_0^\infty \bar{H}_{K(v)+1}(x, \dots, x) dx\right]. \end{aligned}$$

Similarly, the time that a node is compromised in a cycle is

$$E[C(v)] = E[\min\{Y_1, Y_2\}] = \int_0^\infty \bar{G}(x, x) dx.$$

Note also that Assumption 4.1 implies that

$$E[S(v) + C(v)] < \infty.$$

This and the fact that the distribution of $S(v) + C(v)$ is nonlattice guarantee (cf. [Ross 96, Theorem 3.4.4]) that the process is steady, and moreover,

$$q(v) = \frac{E[C(v)]}{E[S(v)] + E[C(v)]} = \frac{\int_0^\infty \bar{G}(x, x) dx}{E \left[\int_0^\infty \bar{H}_{K(v)+1}(x, \dots, x) dx \right] + \int_0^\infty \bar{G}(x, x) dx}.$$

The desired result follows. \square

The following corollary relates Theorem 2.2 to Theorem 4.2.

Corollary 4.3. *Theorem 2.2 is a special case of Theorem 4.2.*

Proof. By replacing Assumption 4.1 with Assumption 2.1 and letting G be a regular graph in Theorem 4.2, we obtain

$$\begin{aligned} m &= E \left[\int_0^\infty \bar{H}_{K(v)+1}(x, \dots, x) dx \right] / \int_0^\infty \bar{G}(x, x) dx \\ &= E \left[\int_0^\infty \bar{F}_1(x) \prod_{i=1}^{K(v)} \bar{F}_{2,i}(x) dx \right] / \int_0^\infty \bar{G}_1(x) \bar{G}_2(x) dx \\ &= E \left[\int_0^\infty \exp\{-\alpha x\} \exp\{-K(v) \cdot \gamma\} dx \right] / \int_0^\infty \exp\{-\beta x\} \exp\{-\eta x\} dx \\ &= E \left[\frac{\beta + \eta}{\alpha + \gamma K(v)} \right]. \end{aligned}$$

Therefore, the resulting q , namely $q(v)$ in the case of regular graphs, is the same as in Theorem 2.2. \square

Now we consider the case that the random variables are independent of each other but follow the Weibull distribution (rather than the exponential distribution as required in [Li et al. 11]).

Assumption 4.4. (Weaker than Assumption 2.1 but stronger than Assumption 4.1.)

1. Assumption on the distributions of the random variables:

- X_1 follows the Weibull distribution with survival function $\bar{F}_1(t) = e^{-(\lambda_1 t)^\alpha}$.
- The $X_{2,i}$'s follow the Weibull distribution with survival function $\bar{F}_2(t) = e^{-(\lambda_2 t)^\alpha}$.

- Y_1 follows the Weibull distribution with survival function $\bar{G}_1(t) = e^{-(\gamma_1 t)^\beta}$.
 - Y_2 follows the Weibull distribution with survival function $\bar{G}_2(t) = e^{-(\gamma_2 t)^\beta}$.
2. X_1 , the $X_{2,i}$'s, Y_1 , and Y_2 are independent of each other (as in Assumption 2.1).

In practice, the Weibull distribution in Assumption 4.4 can be tested using standard statistical techniques such as QQ-plot and the Kolmogorov–Smirnov test (cf. [Lehmann and Romano 05, Chapter 14]). The following result is more general than Theorem 2.2 but less general than Theorem 4.2.

Theorem 4.5. *Replacing Assumption 4.1 in Theorem 4.2 with Assumption 4.4, we have*

$$m = \frac{\Gamma\left(1 + \frac{1}{\alpha}\right)}{\Gamma\left(1 + \frac{1}{\beta}\right)} \mathbb{E} \left[\frac{\left(\gamma_1^\beta + \gamma_2^\beta\right)^{1/\beta}}{\left(\lambda_1^\alpha + K(v) \cdot \lambda_2^\alpha\right)^{1/\alpha}} \right].$$

Proof. Note that

$$\begin{aligned} \int_0^\infty \bar{G}_1(x)\bar{G}_2(x) dx &= \int_0^\infty \exp\{-[(\gamma_1 x)^\beta + (\gamma_2 x)^\beta]\} dx \\ &= \int_0^\infty \exp\{-[\gamma_1^\beta + \gamma_2^\beta] x^\beta\} dx \\ &= \frac{1}{\left(\gamma_1^\beta + \gamma_2^\beta\right)^{1/\beta}} \Gamma\left(1 + \frac{1}{\beta}\right), \end{aligned}$$

where $\Gamma(\cdot)$ is the gamma function. Note also that

$$\begin{aligned} \mathbb{E} \left[\int_0^\infty \bar{F}_1(x)\bar{F}_2^{K(v)}(x) dx \right] &= \mathbb{E} \left[\int_0^\infty \exp\{-[(\lambda_1 x)^\alpha + K(v) \cdot (\lambda_2 x)^\alpha]\} dx \right] \\ &= \mathbb{E} \left[\int_0^\infty \exp\{-[(\lambda_1^\alpha + K(v) \cdot \lambda_2^\alpha) x^\alpha]\} dx \right] \\ &= \Gamma\left(1 + \frac{1}{\alpha}\right) \mathbb{E} \left[\frac{1}{\lambda_1^\alpha + K(v) \cdot \lambda_2^\alpha} \right]^{1/\alpha}. \end{aligned}$$

Using Theorem 4.2, we have the desired result. □

4.2. Applying the Analytical Results to Obtain q in Practice

Now we show how to use the above analytical results to derive methods for obtaining q . While allowing X_1 , the $X_{2,i}$'s, Y_1 , and Y_2 to be possibly dependent on each other, we further consider the (in)dependence between the states of a node's neighbors. This leads to three cases:

- (i) In the case that G is a regular graph of node degree μ and the dependence between neighbors can be ignored, we offer closed-form results in Section 4.2.1.
- (ii) In the case that G is an arbitrary graph and the dependence between neighbors can be ignored, we offer approximation results in Section 4.2.2.
- (iii) In the case that G is an arbitrary graph and the dependence between the neighbors cannot be ignored, we offer statistical results in Section 4.2.3.

4.2.1. Regular Graphs with Ignorable Dependence between Nodes. In case (i), every node has μ neighbors, denoted by $1, 2, \dots, \mu$, which are **compromised** in the steady state with probability $q_1 = q_2 = \dots = q_\mu = q$. Let I_i be a Bernoulli random variable with parameter q , which is the probability that the i th neighbor is **compromised**. Then $K(v) = \sum_{i=1}^{\mu} I_i$ is a binomial distribution with parameter q , and

$$P(K(v) = k) = \binom{\mu}{k} q^k (1 - q)^{\mu - k}, \quad k = 0, \dots, \mu.$$

According to Theorem 4.2, which requires Assumption 4.1, we have

$$\begin{aligned} m &= \mathbb{E} \left[\int_0^\infty \bar{H}_{K(v)+1}(x, \dots, x) dx \right] / \int_0^\infty \bar{G}(x, x) dx \\ &= \sum_{k=0}^{\mu} \left[\binom{\mu}{k} q^k (1 - q)^{\mu - k} \int_0^\infty \bar{H}_{k+1}(x, \dots, x) dx \right] / \int_0^\infty \bar{G}(x, x) dx. \end{aligned}$$

Hence, we have

$$q + \sum_{k=0}^{\mu} \left[\binom{\mu}{k} q^{k+1} (1 - q)^{\mu - k} \int_0^\infty \bar{H}_{k+1}(x, \dots, x) dx \right] / \int_0^\infty \bar{G}(x, x) dx = 1. \tag{4.2}$$

In order to solve (4.2), we need to know the specific $\bar{H}_{k+1}(x, \dots, x)$ and $\bar{G}(x, x)$ that can be obtained in scenarios such as the following:

- Closed-form result under Assumption 2.1, which is stronger than Assumption 4.1: In this special case, we have Theorem 2.2 and thus (4.2) becomes

$$q + \sum_{k=0}^{\mu} \frac{\beta + \eta}{\alpha + \gamma k} \binom{\mu}{k} q^{k+1} (1 - q)^{\mu - k} = 1. \tag{4.3}$$

- Closed-form result under Assumption 4.4, which is stronger than Assumption 4.1. In this special case, we have Theorem 4.5, and thus (4.2) becomes

$$q + \sum_{k=0}^{\mu} \binom{\mu}{k} q^{k+1} (1 - q)^{\mu - k} \frac{\Gamma(1 + \frac{1}{\alpha})}{\Gamma(1 + \frac{1}{\beta})} \left[\frac{(\gamma_1^\beta + \gamma_2^\beta)^{1/\beta}}{(\lambda_1^\alpha + k\lambda_2^\alpha)^{1/\alpha}} \right] = 1. \tag{4.4}$$

When μ is not large, the solutions for q in (4.3) and (4.4) can be derived using an appropriate iterative algorithm, such as the Newton–Raphson method [Stoer and Bulirsch 02]. When μ is large, it might be infeasible to derive explicit solutions for q , but we can instead use the approximation result in Section 4.2.2.

4.2.2. Arbitrary Graphs with Ignorable Dependence between the Nodes. For an arbitrary node $v \in V$, let I_i be a Bernoulli random variable with parameter q_i , which is the probability that the i th neighbor of node v is compromised. Note that $K(v) = \sum_{i=1}^{\text{deg}(v)} I_i$. Suppose the mean and variance of $K(v)$ exist and we can estimate them (e.g., based on sufficiently many observations on v 's neighbors). Set

$$\bar{k} = E[K(v)], \quad \sigma^2 = \text{Var}(K(v)).$$

In this case, we are unable to give closed-form solutions, and instead propose using two approximations:

- (a) Normal approximation: If $\sigma^2 \rightarrow \infty$ as $\text{deg}(v) \rightarrow \infty$, by Lindeberg's theorem (cf. [Billingsley 95, p. 359]),

$$P(K(v) \leq x) \approx \int_0^x \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(y - \bar{k})^2}{2\sigma^2}\right\} dy.$$

According to Theorem 4.2, which requires Assumption 4.1, we have

$$\begin{aligned} m &= E \left[\int_0^\infty \bar{H}_{K(v)+1}(x, \dots, x) dx \right] / \int_0^\infty \bar{G}(x, x) dx \\ &\approx \int_0^\infty \int_0^{\text{deg}(v)} \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{(k - \bar{k})^2}{2\sigma^2}\right\} \\ &\quad \times \bar{H}_{k+1}(x, \dots, x) dk dx / \int_0^\infty \bar{G}(x, x) dx. \end{aligned} \tag{4.5}$$

As a result, we can obtain more specific approximation solutions as follows:

- Approximation result under Assumption 2.1, which is stronger than Assumption 4.1. In this special case we have Theorem 2.2, and thus (4.5) becomes

$$m \approx \frac{\beta + \eta}{\sqrt{2\pi}\sigma} \int_0^{\deg(v)} \frac{1}{\alpha + \gamma x} \exp\left\{-\frac{(k - \bar{k})^2}{2\sigma^2}\right\} dk.$$

- Approximation result under Assumption 4.4, which is stronger than Assumption 4.1. In this special case, we have Theorem 4.5, and thus (4.5) becomes

$$m \approx \frac{\Gamma\left(1 + \frac{1}{\alpha}\right) (\gamma_1^\beta + \gamma_2^\beta)^{1/\beta}}{\sqrt{2\pi}\sigma\Gamma\left(1 + \frac{1}{\beta}\right)} \times \int_0^{\deg(v)} \frac{1}{(\lambda_1^\alpha + k\lambda_2^\alpha)^{1/\alpha}} \exp\left\{-\frac{(k - \bar{k})^2}{2\sigma^2}\right\} dk.$$

- (b) Poisson approximation. Note that $\bar{k} \geq 0$. If $\deg(v) \rightarrow \infty$ and

$$\max\{q_1, \dots, q_{\deg(v)}\} \rightarrow 0,$$

then [Billingsley 95, Theorem 23.2] says that

$$P(K(v) = i) \approx \exp\{-\bar{k}\} \frac{(\bar{k})^i}{i!}, \quad i = 0, 1, \dots, \deg(v).$$

According to Theorem 4.2, which requires Assumption 4.1, we have

$$m = \mathbb{E} \left[\int_0^\infty \bar{H}_{K(v)+1}(x, \dots, x) dx \right] / \int_0^\infty \bar{G}(x, x) dx \\ \approx \sum_{i=0}^{\deg(v)} \exp\{-\bar{k}\} \frac{(\bar{k})^i}{i!} \int_0^\infty \bar{H}_{i+1}(x, \dots, x) dx / \int_0^\infty \bar{G}(x, x) dx. \quad (4.6)$$

As a result, we can have more specific approximation solutions as follows:

- Approximation result under Assumption 2.1, which is stronger than Assumption 4.1. In this special case, we have Theorem 2.2, and thus (4.6) becomes

$$m \approx (\beta + \eta) \exp\{-\bar{k}\} \sum_{i=0}^{\deg(v)} \frac{1}{\alpha + \gamma i} \frac{(\bar{k})^i}{i!}.$$

- Approximation result under Assumption 4.4, which is stronger than Assumption 4.1. In this special case, we have Theorem 4.5, and thus

(4.6) becomes

$$m \approx \frac{\Gamma(1 + \frac{1}{\alpha}) (\gamma_1^\beta + \gamma_2^\beta)^{1/\beta}}{\Gamma(1 + \frac{1}{\beta})} \sum_{i=0}^{\deg(v)} \frac{(\bar{k})^i \exp\{-\bar{k}\}}{(\lambda_1^\alpha + i \cdot \lambda_2^\alpha)^{1/\alpha} i!}.$$

In both normal approximation and Poisson approximation, having obtained m allows one to derive $q(v)$ according to (4.1) in Theorem 4.2. Note that this method is sound because both Assumption 2.1 and Assumption 4.4 are stronger than Assumption 4.1, which underlies Theorem 4.2. From a sufficiently large sample $V' \subset V$ with each $v \in V'$ being uniformly chosen from V , we can derive $q = \sum_{v \in V'} q(v)/|V'|$.

4.2.3. Arbitrary Graphs with Nonignorable Dependence between the Nodes. If the dependence between neighbors cannot be ignored, the probability mass function of $K(v)$ may be estimated from data as follows. Suppose we have observations (k_1, k_2, \dots, k_b) on $K(v)$, where b is the number of observations on the number of v 's compromised neighbors in the steady state (i.e., at b different points of observation time). Then the probability mass function of $K(v)$ can be estimated as

$$p_k = P(K(v) = k) = \frac{\sum_{i=1}^b \mathbf{I}(k_i = k)}{b}, \quad b \rightarrow \infty,$$

where $\mathbf{I}(\cdot)$ is the indicator function. Thus, according to Theorem 4.2, we have

$$\begin{aligned} m &= \mathbb{E} \left[\int_0^\infty \bar{H}_{K(v)+1}(x, \dots, x) dx \right] / \int_0^\infty \bar{G}(x, x) dx \\ &= \sum_{k=0}^{\deg(v)} p_k \int_0^\infty \bar{H}_{k+1}(x, \dots, x) dx / \int_0^\infty \bar{G}(x, x) dx. \end{aligned} \tag{4.7}$$

To be specific, let us consider a concrete example under the following assumption:

Assumption 4.6. (Weaker than Assumption 2.1 but stronger than Assumption 4.1.)

1. Assumption on the distributions of the random variables:
 - X_1 follows the exponential distribution with parameter λ .
 - The $X_{2,i}$'s follow the multivariate Marshall–Olkin exponential model.
 - Y_1 and Y_2 follow the bivariate Marshall–Olkin exponential model.
2. The $X_{2,i}$'s can be dependent on each other; Y_1 and Y_2 can be dependent on each other.

Under Assumption 4.6, the joint survival function of $(X_{2,1}, \dots, X_{2,k})$ is

$$\begin{aligned} \bar{H}_k(x_1, \dots, x_k) &= P(X_{2,1} > x_1, \dots, X_{2,k} > x_k) \\ &= \exp \left\{ - \sum_{i=1}^k \lambda_i x_i - \sum_{i < j} \lambda_{ij} \max\{x_i, x_j\} - \lambda_{12\dots k} \max\{x_1, \dots, x_k\} \right\}. \end{aligned}$$

Therefore,

$$\begin{aligned} \int_0^\infty \bar{H}_{k+1}(x, \dots, x) dx &= \int_0^\infty \exp\{-\lambda t\} \exp\{-\phi(\boldsymbol{\lambda}, kt)\} dt \\ &= \frac{1}{\lambda + \phi(\boldsymbol{\lambda}, k)}, \end{aligned}$$

where

$$\phi(\boldsymbol{\lambda}, k) = \sum_{i=1}^k \lambda_i + \sum_{i < j} \lambda_{ij} + \dots + \lambda_{12\dots k}.$$

Similarly, the joint survival function of (Y_1, Y_2) is

$$\begin{aligned} \bar{G}(y_1, y_2) &= P(Y_1 > y_1, Y_2 > y_2) \\ &= \exp \{ -\gamma_1 y_1 - \gamma_2 y_2 - \gamma_{12} \max\{y_1, y_2\} \}. \end{aligned}$$

Then

$$\int_0^\infty \bar{G}(x, x) dx = \frac{1}{\gamma_1 + \gamma_2 + \gamma_{12}}.$$

Using (4.7), we have

$$m = (\gamma_1 + \gamma_2 + \gamma_{12}) \sum_{k=0}^{\deg(v)} \frac{p_k}{\lambda + \phi(\boldsymbol{\lambda}, k)}.$$

Having obtained m , one can derive $q(v)$ according to (4.1) in Theorem 4.2. Note that this method is sound because Assumption 4.6 is stronger than Assumption 4.1, which underlies Theorem 4.2. From a sufficiently large sample $\mathbf{V}' \subset \mathbf{V}$ with each $v \in \mathbf{V}'$ being uniformly chosen from \mathbf{V} , we can derive $q = \sum_{v \in \mathbf{V}'} q(v) / |\mathbf{V}'|$.

4.3. New Bounds on q

Above, we discussed how to obtain q in various settings. When one cannot obtain q (for example, due to the lack of sufficient information), one can instead use the bounds of q in decision-making. This is plausible as long as the bounds are reasonably tight, and reasonable because it requires much less information to compute the bounds. In particular, one could use the upper bound of q in decision-making because it can be seen as a sort of worst-case scenario. In what

follows we will differentiate the case of regular graphs from the case of arbitrary graphs under the following three assumptions:

Assumption 4.7. (Assumption underlying our general bounds of q .)

1. Assumption on the distributions of the random variables:
 - X_1 follows an arbitrary continuous distribution with survival function \bar{F}_1 and finite mean.
 - The $X_{2,i}$'s are PUOD with the same marginal continuous survival function \bar{F}_2 and finite mean.
 - Y_1 follows an arbitrary distribution with continuous survival function \bar{G}_1 and finite mean.
 - Y_2 follows an arbitrary distribution with continuous survival function \bar{G}_2 and finite mean.
2. X_1 and the $X_{2,i}$'s are independent, and Y_1 and Y_2 may be dependent on each other.

Assumption 4.8. (Assumption underlying one specific bound of q .)

1. Assumption on the distributions of the random variables:
 - X_1 follows the exponential distribution with rate α .
 - The $X_{2,i}$'s are PUOD with the same marginal distribution function F_2 of the exponential distribution with rate γ .
 - Y_1 follows the exponential distribution with rate β .
 - Y_2 follows the exponential distribution with rate η .
2. X_1 and the $X_{2,i}$'s are independent, and Y_1 and Y_2 are independent of each other.

Assumption 4.9. (Assumption underlying another specific bound of q .)

1. Assumption on the distributions of the random variables:
 - X_1 follows the Lomax distribution with scale parameter λ and shape parameter $\alpha_1 > 1$.
 - The $X_{2,i}$'s are PUOD with the same marginal distribution function F_2 of the Lomax distribution with scale parameter λ and shape parameter $\alpha_2 > 1$.

- Y_1 follows the Lomax distribution with scale parameter γ and shape parameter $\beta_1 > 1$.
 - Y_2 follows the Lomax distribution with scale parameter γ and shape parameter $\beta_2 > 1$.
2. X_1 and the $X_{2,i}$'s are independent, and Y_1 and Y_2 are independent of each other.

In practice, the PUOD dependence assumption can be tested by the method suggested in [Scaillet 05], while the exponential distribution assumption and the Lomax distribution assumption can be tested using QQ-plot or the Kolmogorov–Smirnov test (see [Lehmann and Romano 05]).

4.3.1. Bounds on q in the Case of Regular Graphs. We have the following theorem.

Theorem 4.10. *Suppose G is regular graph of node degree μ . Under Assumption 4.7, in the steady state we have*

$$\int_0^\infty \bar{G}(x, x) dx \Big/ \int_0^\infty [\bar{F}_1(x) + \bar{G}(x, x)] dx \\ \leq q \leq \int_0^\infty \bar{G}(x, x) dx \Big/ \int_0^\infty [\bar{F}_1(x) \bar{F}_2^{\bar{k}}(x) + \bar{G}(x, x)] dx,$$

where \bar{k} is the expected number of compromised neighbors for an arbitrary node.

Proof. Denote by $1, 2, \dots, \mu$ the neighbors of an arbitrary node $v \in V$. Let I_i be the indicator function that $I_i = 1$ if and only if the i th neighbor of v is **compromised**, and let q_i be the probability that the i th neighbor is **compromised**. Under Assumption 4.7, the system will enter the steady state (see [Ross 96, Theorem 3.4.4]). Since G is regular graph, we have $q_1 = \dots = q_\mu = q$. In the steady state, we have

$$\bar{k} = E[K(v)] = E \left[\sum_{i=1}^{\mu} I_i \right] = q\mu.$$

Assume that the $X_{2,i}$'s are PUOD with the same marginal survival function \bar{F}_2 . Then

$$\begin{aligned}
 m &= \mathbb{E} \left[\int_0^\infty \bar{F}_1(x) \bar{H}_{K(v)}(x, \dots, x) dx \right] / \int_0^\infty \bar{G}(x, x) dx \\
 &\geq \mathbb{E} \left[\int_0^\infty \bar{F}_1(x) \bar{F}_2^{K(v)}(x) dx \right] / \int_0^\infty \bar{G}(x, x) dx \\
 &\geq \left[\int_0^\infty \bar{F}_1(x) \bar{F}_2^{\mathbb{E}[K(v)]}(x) dx \right] / \int_0^\infty \bar{G}(x, x) dx \\
 &= \left[\int_0^\infty \bar{F}_1(x) \bar{F}_2^{\bar{k}}(x) dx \right] / \int_0^\infty \bar{G}(x, x) dx, \tag{4.8}
 \end{aligned}$$

where the last inequality follows from Jensen's inequality. Using (4.1) in Theorem 4.2, one can obtain the upper bound of q .

On the other hand, we observe that for every $x \geq 0$, we have

$$\bar{H}_{K(v)}(x, \dots, x) dx \leq 1.$$

The upper bound of m can be derived as

$$m \leq \int_0^\infty \bar{F}_1(x) dx / \int_0^\infty \bar{G}(x, x) dx. \tag{4.9}$$

Using (4.1) in Theorem 4.2, one can obtain the lower bound of q . □

Theorem 4.11. *Suppose G is a regular graph of node degree μ . Under Assumption 4.8, in the steady state we have*

$$\frac{\alpha}{\alpha + \beta + \gamma} \leq q \leq \frac{-(\beta + \eta + \alpha - \mu\gamma) + \sqrt{(\beta + \eta + \alpha - \mu\gamma)^2 + 4\mu\gamma\alpha}}{2\mu\gamma}.$$

Proof. From (4.8), it follows that

$$m \geq \frac{\beta + \eta}{\alpha + q\mu\gamma}.$$

Using (4.1) in Theorem 4.2, we have

$$q = \frac{1}{m + 1} \leq \frac{1}{\frac{\beta + \eta}{\alpha + q\mu\gamma} + 1}.$$

This leads to

$$q^2 \mu\gamma + q(\beta + \eta + \alpha - \mu\gamma) - \alpha \leq 0.$$

Hence

$$q \leq \frac{-(\beta + \eta + \alpha - \mu\gamma) + \sqrt{(\beta + \eta + \alpha - \mu\gamma)^2 + 4\mu\gamma\alpha}}{2\mu\gamma}.$$

On the other hand, according to (4.9) and (4.1), one can derive

$$q \geq \frac{\alpha}{\alpha + \beta + \gamma}.$$

□

Theorem 4.12. *Suppose G is a regular graph of node degree μ . Under Assumption 4.9, in the steady state we have*

$$\frac{\gamma(\alpha_1 - 1)}{\lambda(\beta_1 + \beta_2 - 1) + \gamma(\alpha_1 - 1)} \leq q \leq \frac{-B + \sqrt{B^2 + 4\gamma^2\alpha_2(\alpha_1 - 1)\mu}}{2\alpha_2\gamma\mu},$$

where $B = \lambda(\beta_1 + \beta_2 - 1) + \gamma(\alpha_1 - 1) - \gamma\alpha_2\mu$.

Proof. Note that F_i has Lomax survival function

$$\bar{F}_i(x) = \left(1 + \frac{x}{\lambda}\right)^{-\alpha_i}, \quad \alpha_i \geq 1, \quad i = 1, 2.$$

Similarly, G_i has Lomax survival function

$$\bar{G}_i(x) = \left(1 + \frac{x}{\gamma}\right)^{-\beta_i}, \quad \beta_i \geq 1, \quad i = 1, 2.$$

From (4.8), it follows that

$$m \geq \frac{\lambda}{\alpha_1 + \alpha_2\mu q - 1} \bigg/ \frac{\gamma}{\beta_1 + \beta_2 - 1} = \frac{\lambda(\beta_1 + \beta_2 - 1)}{\gamma(\alpha_1 + \alpha_2\mu q - 1)}.$$

Using (4.1) in Theorem 4.2, we have

$$q \leq \frac{\gamma(\alpha_1 + \alpha_2\mu q - 1)}{\lambda(\beta_1 + \beta_2 - 1) + \gamma(\alpha_1 + \alpha_2\mu q - 1)}.$$

This leads to

$$q^2\alpha_2\gamma\mu + qB - \gamma(\alpha_1 - 1) \leq 0,$$

where

$$B = \lambda(\beta_1 + \beta_2 - 1) + \gamma(\alpha_1 - 1) - \gamma\alpha_2\mu.$$

Hence, the upper bound of q is

$$q \leq \frac{-B + \sqrt{B^2 + 4\gamma^2\alpha_2(\alpha_1 - 1)\mu}}{2\alpha_2\gamma\mu}.$$

On the other hand, according to (4.9) and (4.1), we have

$$m \leq \frac{\lambda(\beta_1 + \beta_2 - 1)}{\gamma(\alpha_1 - 1)}.$$

So the lower bound of q is

$$q \geq \frac{\gamma(\alpha_1 - 1)}{\lambda(\beta_1 + \beta_2 - 1) + \gamma(\alpha_1 - 1)}.$$

□

4.3.2. Bounds of q in the Case of Arbitrary Graphs. For an arbitrary node $v \in \mathbf{V}$, it follows that

$$K(v) = \sum_{i=1}^{\deg(v)} I_i,$$

where I_i is a Bernoulli random variable with parameter q_i , which is the probability that the i th neighbor of v is compromised. Note that the I_i 's may be dependent on each other and the q_i 's may be different. Note also that in this case, we cannot obtain a counterpart of Theorem 4.10, because the concavity needed to apply Jensen's inequality may not hold. As such, we are able to give only the following specific results in this more general case.

Theorem 4.13. *Suppose \mathbf{G} is an arbitrary graph with average node degree μ . Under Assumption 4.8, in the steady state we have*

$$\frac{\alpha}{\alpha + \beta + \gamma} \leq q \leq \frac{\alpha + \gamma\mu}{\alpha + \beta + \eta + \gamma\mu}.$$

Proof. In a fashion similar to the proof of Theorem 4.10, we can show that for an arbitrary node v ,

$$\frac{\alpha}{\alpha + \beta + \gamma} \leq q(v) \leq \frac{\alpha + \gamma \mathbf{E}[K(v)]}{\alpha + \beta + \eta + \gamma \mathbf{E}[K(v)]}.$$

Since

$$q = \frac{\sum_{v \in \mathbf{V}} q(v)}{n},$$

it follows that

$$q \geq \frac{\alpha}{\alpha + \beta + \gamma}.$$

On the other hand, because

$$\mathbf{E}[K(v)] = \mathbf{E} \left[\sum_{i=1}^{\deg(v)} I_i \right] = \sum_{i=1}^{\deg(v)} q_i \leq \deg(v),$$

we have

$$\frac{\alpha + \gamma \mathbb{E}[K(v)]}{\alpha + \beta + \eta + \gamma \mathbb{E}[K(v)]} \leq \frac{\alpha + \gamma \deg(v)}{\alpha + \beta + \eta + \gamma \deg(v)}.$$

Hence

$$q(v) \leq \frac{\alpha + \gamma \deg(v)}{\alpha + \beta + \eta + \gamma \deg(v)}.$$

Therefore,

$$q \leq \frac{1}{n} \sum_{v \in \mathcal{V}} \frac{\alpha + \gamma \deg(v)}{\alpha + \beta + \eta + \gamma \deg(v)}. \quad (4.10)$$

Because

$$h(x) = \frac{\alpha + \gamma x}{\alpha + \beta + \eta + \gamma x}, \quad x \geq 0,$$

is a concave function, Jensen's inequality yields

$$\frac{1}{n} \sum_{v \in \mathcal{V}} \frac{\alpha + \gamma \deg(v)}{\alpha + \beta + \eta + \gamma \deg(v)} \leq \frac{\alpha + \gamma \frac{1}{n} \sum_{v \in \mathcal{V}} \deg(v)}{\alpha + \beta + \eta + \gamma \frac{1}{n} \sum_{v \in \mathcal{V}} \deg(v)}.$$

Hence, by (4.10), we have

$$q \leq \frac{\alpha + \gamma \mu}{\alpha + \beta + \eta + \gamma \mu}.$$

This completes the proof. \square

In a similar fashion, we can prove the following result.

Theorem 4.14. *Suppose G is an arbitrary graph with average degree μ . Under Assumption 4.9, in the steady state we have*

$$\frac{\gamma(\alpha_1 - 1)}{\lambda(\beta_1 + \beta_2 - 1) + \gamma(\alpha_1 - 1)} \leq q \leq \frac{\gamma(\alpha_1 + \alpha_2 \mu - 1)}{\lambda(\beta_1 + \beta_2 - 1) + \gamma(\alpha_1 + \alpha_2 \mu - 1)}.$$

5. New Results for the Case in Which Only the Alternating Renewal Process Is Observed

The results obtained thus far are applicable when the lower-level attack–defense process is observed, or more specifically the distributions F_1 , the $F_{2,i}$'s, G_1 , and G_2 and their dependence structure (if applicable) are observed. What if they are not observed? This is a very reasonable question, because obtaining all detailed information can be very costly, and it may be much simpler to observe

the higher-level alternating renewal process $\{(S_j(v), C_j(v)); j \geq 1\}$ mentioned above. In this case, neither the results in [Li et al. 11] nor the results described above are applicable, which motivates us to present the following results for this possibly more realistic scenario.

5.1. Analytical Results for Estimating $q(v)$

Let $p_v(t)$ be the probability that node v is secure at time t , and let the limiting average probability be

$$\bar{p}(v) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t p_v(s) ds.$$

For node $v \in \mathbf{V}$ with $\{(S_j(v), C_j(v)) : 1 \leq j \leq b\}$, define

$$W_b(v) = S_1(v) + S_2(v) + \dots + S_b(v), \quad D_b(v) = C_1(v) + C_2(v) + \dots + C_b(v).$$

Lemma 5.1. [Marlow and Tortorella 95, Theorem 1] *If*

$$\lim_{b \rightarrow \infty} \frac{W_b(v)}{b} = w(v) \text{ a.s.} \quad \text{and} \quad \lim_{b \rightarrow \infty} \frac{D_b(v)}{b} = d(v) \text{ a.s.,}$$

where “a.s.” means “almost surely” and $0 < w(v), d(v) < \infty$, then

$$\bar{p}(v) = \frac{w(v)}{w(v) + d(v)}.$$

Assumption 5.2. Assume that $\{S_j(v) : j \geq 1\}$ and $\{C_j(v) : j \geq 1\}$ are sequences of pairwise PQD random variables with finite variances, and

- (a) $\sum_{j=1}^{\infty} j^{-2} \text{Cov}(S_j(v), W_j(v)) < \infty$;
- (b) $\sum_{j=1}^{\infty} j^{-2} \text{Cov}(C_j(v), D_j(v)) < \infty$;
- (c) $\sup_{j \geq 1} \mathbf{E} |S_j(v) - \mathbf{E}[S_j(v)]| < \infty$;
- (d) $\sup_{j \geq 1} \mathbf{E} |C_j(v) - \mathbf{E}[C_j(v)]| < \infty$.

In practice, the above assumption can be tested as follows: PQD dependence can be tested using the statistical methods described in [Denuit and Scaillet 04]. To test for finite variance, we may use the methods described in [Resnick 07] for heavy-tailed data. If the data are not extremely heavy-tailed, we may assume that the data have finite variance.

Lemma 5.3. [Birkel 89, Theorem 1] *Let $\{S_j(v) : j \geq 1\}$ be a sequence of pairwise PQD random variables with finite variance. Assume*

$$(a) \sum_{j=1}^{\infty} j^{-2} \text{Cov}(S_j(v), W_j(v)) < \infty.$$

$$(b) \sup_{j \geq 1} \mathbb{E} |S_j(v) - \mathbb{E}[S_j(v)]| < \infty.$$

Then as $b \rightarrow \infty$, we have $b^{-1}(W_b(v) - \mathbb{E}[W_b(v)]) \rightarrow 0$ almost surely.

Theorem 5.4. *Under Assumption 5.2 and given a sufficiently large number b of observations on the states of node v , the probability $q(v)$, namely the probability that node v is **compromised**, if it exists (i.e., the alternating renewal process is steady), can be expressed as*

$$q(v) = \frac{d^*(v)}{w^*(v) + d^*(v)},$$

where $w^(v) = \lim_{b \rightarrow \infty} \mathbb{E}[W_b(v)]/b$ and $d^*(v) = \lim_{b \rightarrow \infty} \mathbb{E}[D_b(v)]/b$.*

Proof. According to Lemma 5.3, conditions (a) and (c) imply

$$\lim_{b \rightarrow \infty} \frac{W_b(v)}{b} = \lim_{b \rightarrow \infty} \frac{\mathbb{E}[W_b(v)]}{b} = w^*(v), \quad a.s.,$$

and similarly, (b) and (d) imply

$$\lim_{b \rightarrow \infty} \frac{D_b(v)}{b} = \lim_{b \rightarrow \infty} \frac{\mathbb{E}[D_b(v)]}{b} = d^*(v), \quad a.s.$$

According to Lemma 5.1,

$$\bar{p}(v) = \frac{w^*(v)}{w^*(v) + d^*(v)}.$$

The limiting probability is

$$p(v) = \lim_{t \rightarrow \infty} p_v(t),$$

which means that there exists a sufficiently large t_0 such that for every $\epsilon > 0$,

$$|p(v) - p_v(t)| < \epsilon, \quad t \geq t_0.$$

Now for $t > t_0$, we have

$$\frac{1}{t} \int_0^t p_v(s) ds \leq \frac{1}{t} \int_0^{t_0} p_v(s) ds + \frac{1}{t} \int_{t_0}^t p_v(s) ds.$$

Hence,

$$\bar{p}(v) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t p_v(s) ds \leq p(v) + \epsilon.$$

Similarly,

$$\bar{p}(v) \geq p(v) - \epsilon.$$

So we conclude that

$$p(v) = \bar{p}(v) = \frac{w^*(v)}{w^*(v) + d^*(v)},$$

and

$$q(v) = 1 - p(v) = \frac{d^*(v)}{w^*(v) + d^*(v)}.$$

□

In what follows, we present two variants of Theorem 5.4. In the first variant, we use the stronger PA assumption to replace the PQD assumption in Theorem 5.4 (or more precisely, in the underlying Assumption 5.2), which yields that conditions (c) and (d) can be dropped. To see this, we need the following lemma.

Lemma 5.5. [Birkel 89, Theorem 2] *Let $\{S_j(v) : j \geq 1\}$ be a sequence of pairwise PA random variables with finite variance. Assume that*

$$\sum_{j=1}^{\infty} j^{-2} \text{Cov}(S_j(v), W_j(v)) < \infty.$$

Then as $b \rightarrow \infty$, we have $b^{-1}(W_b(v) - E[W_b(v)]) \rightarrow 0$ almost surely.

Assumption 5.6. Assume that $\{S_j(v) : j \geq 1\}$ and $\{C_j(v) : j \geq 1\}$ are sequences of pairwise PA random variables with finite variances, and that

- (a) $\sum_{j=1}^{\infty} j^{-2} \text{Cov}(S_j(v), W_j(v)) < \infty$;
- (b) $\sum_{j=1}^{\infty} j^{-2} \text{Cov}(C_j(v), D_j(v)) < \infty$.

Theorem 5.7. (One variant of Theorem 5.4.) *Under Assumption 5.6 and given a sufficiently large number b of observations on node v 's states, the probability $q(v)$, if it exists (i.e., the alternating renewal process is steady), can be expressed as*

$$q(v) = \frac{d^*(v)}{w^*(v) + d^*(v)}.$$

Proof. The proof follows from Lemma 5.5 and the proof of Theorem 5.4. □

In the following we present another useful variant of Theorem 5.4 under another stronger assumption.

Assumption 5.8. Assume that $\{S_j(v) : j \geq 1\}$ and $\{C_j(v) : j \geq 1\}$ are independent sequences with finite variances.

In practice, we can test the independence for sequences $\{S_j : j \geq 1\}$ and $\{C_j : j \geq 1\}$ using the methods described in [Hollander and Douglas 99].

Theorem 5.9. (Another variant of Theorem 5.4.) *Under Assumption 5.8 and given a sufficiently large number b of observations on node v 's states, the probability $q(v)$, if it exists (i.e., the alternating renewal process is steady), can be expressed as*

$$q(v) = \frac{d^*(v)}{w^*(v) + d^*(v)}.$$

Proof. The theorem follows from Kolmogorov's strong law of large numbers and the proof of Theorem 5.4. \square

The following corollary expresses the connection between Theorem 5.9, which corresponds to the case in which the higher-level alternating renewal process is observed, and Theorem 2.2, which corresponds to the case in which the lower-level attack–defense process is observed.

Corollary 5.10. *Theorem 2.2 is a special case of Theorem 5.9.*

Proof. Under Assumption 2.1 used in Theorem 2.2, we have

$$\begin{aligned} \text{Var}(S(v)) &< 2 \mathbb{E} \left[\frac{1}{\alpha + K(v) \cdot \gamma} \right]^2 \leq 2 \left(\frac{1}{\alpha} \right)^2 < \infty, \\ \text{Var}(C(v)) &= \left(\frac{1}{\beta + \eta} \right)^2 < \infty. \end{aligned}$$

Hence, the conditions in Theorem 5.9 are fulfilled. Therefore, we have

$$w^*(v) = \mathbb{E}[S(v)] = \mathbb{E} \left[\frac{1}{\alpha + \gamma K(v)} \right], \quad d^*(v) = \mathbb{E}[C(v)] = \frac{1}{\beta + \eta}.$$

As a result,

$$q(v) = \mathbb{E} \left[\frac{\beta + \eta}{\alpha + \gamma K(v)} \right],$$

which is the same as in Theorem 2.2. \square

5.2. Applying the Analytical Results to Estimating q in Practice

Theorems 5.4, 5.7, and 5.9 require observations only on $\{(S_l(v), C_l(v)); l \geq 1\}$. In order to show how the above analytical results can guide the computation of q in practice, we use Theorem 5.9 as an example. Suppose we select m nodes, denoted by v_1, \dots, v_m , uniformly at random from V to observe their respective $\{(S_l(v), C_l(v)); l \geq 1\}$'s. For each $v_i, 1 \leq i \leq m$, we estimate $q(v_i)$ as follows:

1. Observe the node for a sufficiently long time,¹ and get j_i observations $(S_1(v_i), C_1(v_i)), \dots, (S_{j_i}(v_i), C_{j_i}(v_i))$.
2. Test whether $\{S_l(v_i) : j_i \geq l \geq 1\}$ and $\{C_l(v_i); j_i \geq l \geq 1\}$ are independent sequences (e.g., using methods described in [Hollander and Douglas 99]). If they are independent, the algorithm continues executing the next steps; otherwise, the algorithm aborts.
3. Test whether $\{S_l(v_i) : j \geq l \geq 1\}$ and $\{C_l(v_i); j \geq l \geq 1\}$ have finite variances (e.g., using the methods described in [Resnick 07]). If so, the algorithm continues executing the next step; otherwise, the algorithm aborts.
4. Compute $W_{i,j} = S_1(v_i) + \dots + S_j(v_i)$ and $D_{i,j} = C_1(v_i) + \dots + C_j(v_i)$.
5. Compute

$$q(v_i) = \frac{D_{i,j}}{W_{i,j} + D_{i,j}}. \tag{5.1}$$

Having obtained $q(v_1), \dots, q(v_m)$, we can compute

$$\bar{q} = \frac{\sum_{i=1}^m q(v_i)}{m},$$

which, while intuitive, is not unconditionally applicable, as our algorithm demonstrates.

6. Related Work

As extensively discussed in [Li et al. 11], there have been some (promising) attempts at quantitative security analysis of networked systems. In what follows

¹In practice, we can first observe the time interval $[0, t_1]$ and use (5.1) to compute $q(v_i)$. Then we can observe a longer interval, say $[0, t_2]$, and compute $q(v_i)$ again. Assume that we collect many points $(q(v_i), t_z), z = 1, 2, \dots$. Then we may plot them to look for a horizontal line. If a horizontal line pattern appears after t_z for some z , we may say that t_z is a sufficiently long time. This method can also be used to determine whether the process is steady in practice.

we discuss three main approaches, which are complementary to the “stochastic process” approach of [Li et al. 11] and the present paper.

An extensively studied approach is the “attack-graph” approach, which was initiated in [Phillips and Swiler 98]. Basically, an attack graph is derived from a networked system and its known vulnerabilities. The nodes in an attack graph are the states of the networked system, and the edges reflect the attack steps (i.e., certain computers having been compromised can cause the compromise of more computers). Along this line, substantial progress has been made in the last decade [Jha and Wing 01, Sheyner et al. 02, Ammann et al. 02, Noel et al. 03, Ingols et al. 06, Ou et al. 06, Wang et al. 07, Sawilla and Ou 08, Wang et al. 08a, Wang et al. 08b, Xie et al. 10, Huang et al. 11]. A particular application of attack graphs is to (optimally) harden specific assets by identifying the relevant attack paths. Our approach is complementary to the attack-graph approach because of the following. First, the attack-graph approach mainly focuses on studying known (but unpatched) vulnerabilities. The only exception we are aware of is the recent extension to considering the number of unknown vulnerabilities that are needed in order to compromise some specific assets [Wang et al. 10]. In contrast, our approach aims to embed unknown vulnerabilities into the model inputs (rather than as an outcome of the model). Second, the attack-graph approach is algorithmic or combinatorial in nature (e.g., computing the number of attack paths or hardening security of a given set of assets with minimal effort). In contrast, our approach aims to model the evolution of node states, which could allow us to derive basic laws or principles for the security of networked systems.

An approach that is closely related to ours is what we call the “dynamical system” approach, in which dynamical system models play an essential role. This approach was rooted in epidemic models [McKendrick 26, Kermack and McKendrick 27], which were first introduced to computer security for studying computer viruses in [Kephart and White 91]. This approach has been coupled with complex network structures since [Wang et al. 03, Ganesh et al. 05, Chakrabarti et al. 08]. The state-of-the-art result in this line of research is [Xu et al. 12], which also presented perhaps the first numerical result for estimating the global state that is comparable to our concept of q . While the results in [Xu et al. 12] can accommodate arbitrary network structures, they are based on an independence assumption that is comparable to that the $X_{2,i}$'s in our model are independent of each other. As shown above, we aim to get rid of this independence and other assumptions as much as we can.

The third approach is what we call the “statistical physics” approach (cf. [Albert and Barabási 02] for a large body of literature and its numerous follow-ons). This approach mainly aims to characterize the robustness of network connectivity (i.e., robustness of giant components in the presence of node and/or edge

deletions). While relevant, there is a fundamental difference between network connectivity and security. Specifically, as noted in [Li et al. 11], a malicious attacker may aim to compromise as many computers as possible. However, the ultimate goal of the attacker may be to steal sensitive information rather than to disrupt the connectivity of networks. In other words, an attacker may compromise sensitive information while likely not disrupting the communications of legitimate users.

7. Discussion and Research Directions

In this section we discuss the utility of the stochastic process approach, and suggest future research directions.

7.1. On the Utility of the Stochastic Process Approach

The stochastic process approach, introduced in [Li et al. 11], can be seen as a new way of thinking about quantitative security. The present paper is a first step toward eliminating exponential-distribution assumptions and the independence assumptions in [Li et al. 11]. The approach is centered on stochastic processes that model the interaction between attack and defense. The approach is in its infant stage and is currently based on first-principle modeling (due to the lack of real-life data). Nevertheless, the approach has important real-life implications, as we elaborate below.

First, since stochastic processes capture the evolution of node states, the results are pertinent to the evolution of system states. This will help deepen our understanding of the problem and likely will allow us to draw insights or laws and principles that govern the outcome of the interaction between attack and defense. Note that the characterization is not fundamentally based on knowledge of the parameters, because the parameters exist in any case. In other words, the resulting insights or laws and principles are valid whether we know the specific values of the model parameters or not. Such general (or even universal) laws and principles are of paramount importance.

Second, the analytical results could lead to guidelines (as shown in [Li et al. 11]) for adjusting the defense in a cost-effective, if not optimal, fashion. As mentioned above, the upper bounds can be utilized in practice to capture the worst-case scenario, and can be adopted for decision-making. For example, if we know that q is bounded from above, then some appropriate proactive threshold cryptosystems [Herzberg et al. 97] may be deployed to tolerate the bounded

compromise. Moreover, the study can suggest quantitative defense adjustments in order to contain q below a desired threshold (as shown in [Li et al. 11]).

Third, we have offered here statistical methods for obtaining security measures in practice. In Section 4, we presented numerical methods for deriving the global security measure q in various practical settings. In Section 5, we presented statistical methods for obtaining the global security measure q when only high-level information about the attack–defense process can be obtained.

7.2. Extending the Stochastic Process Models

The models in the present paper can be extended in multiple directions. The first direction is to further weaken the assumptions as much as possible. One question is this: how should we accommodate node and edge heterogeneity (other than the degree or topology heterogeneity), meaning that different nodes (edges) exhibit different parameter distribution characteristics? Note that our analytical results on $q(v)$, the probability that node v is compromised in the steady state, already accommodate node and edge heterogeneity. However, our analytical results on the global security measure q , especially its bounds, are based on the fact that the nodes (edges) exhibit the same parameter distribution characteristics.

Another question is this: how can we bridge the gap between what can be observed in practice (i.e., what can be offered by real-life data) and the weakest possible assumptions that we have to make in order to derive analytical results? This includes the design of new statistical methods for testing assumptions that currently cannot be tested.

The second direction is to accommodate multiple vulnerability classes but from a different perspective. Our models are based on aggregating the effects of vulnerabilities at individual computers. Another perspective is to consider individual vulnerability classes that are “compatible” with each other, where “compatibility” is an intuitive concept that needs to be formalized. The intuition is that the exploitation of some vulnerability in one computer can cause the exploitation of another compatible (i.e., not necessarily the same) vulnerability in another computer. This would lead to multiple stochastic attack–defense processes that might be dependent on each other. This quickly becomes very complicated, but certainly worthy of study, because it may lead to deeper insights. Moreover, it is certainly interesting and important to compare the two perspectives.

The third direction is to accommodate the consequence heterogeneity between nodes. This is important because the damage that is caused by the compromise of one node (e.g., a server) may be greater than the damage caused by the compromise of another (e.g., a desktop). Our models can partially accommodate this consequence heterogeneity via the $q(v)$'s, where $q(v)$ is the probability that

node v is compromised in the steady state. However, our results on, including the bounds of, the global security measure q do not consider the asset on node v , which may be denoted by $\text{asset}(v)$. One simple extension would be to introduce a derivative security measure $q(v) \times \text{asset}(v)$, which can lead to the naturally extended global security measure $\sum_{v \in V} q(v) \times \text{asset}(v) / |V|$. Nevertheless, other extensions are possible.

The fourth direction is to characterize the distribution of the random variable that indicates the number of compromised nodes in the steady state. Note that $q \times |V|$ corresponds to the expected number of compromised nodes in the steady state. Knowledge about the distribution of the random variable will allow us to conduct better decision-making.

7.3. Unifying the Approaches

As mentioned in the introduction, the problem of quantitative security analysis of networked computer systems has been outstanding for decades. In the above section on related work, we highlighted three main approaches. Both the attack-graph approach and the statistical-physics approach have been extensively studied, but both the dynamical system approach and the stochastic process approach are in their infant stages. Nevertheless, it is imperative to unify these approaches into a comprehensive and systematic framework.

8. Conclusion

We extended the quantitative security model and analysis presented in [Li et al. 11] by substantially weakening its assumptions regarding the distributions of, and the dependence between, the relevant random variables. In particular, our extensions lead to practical methods for obtaining the desired security measures for both the case that the lower-level attack–defense process is observed and the case that the higher-level alternating renewal process is observed. We discussed future research directions toward ultimately solving the problem of quantitative security analysis of networked systems.

Acknowledgements. We are very grateful to an anonymous reviewer whose comments helped us improve the paper.

Shouhuai Xu's work was supported in part by grants sponsored by AFOSR, AFOSR MURI, and NSF. The views and conclusions contained in the article are those of the authors and should not be interpreted as in any sense the official policies or endorsements of the United States government or any of its agencies.

References

- [Albert and Barabási 02] R. Albert and A. Barabási. “Statistical Mechanics of Complex Networks.” *Reviews of Modern Physics* 74 (2002), 47–97.
- [Ammann et al. 02] P. Ammann, D. Wijesekera, and S. Kaushik. “Scalable, Graph-Based Network Vulnerability Analysis.” In *Proc. ACM Conference Computer and Communications Security (CCS '02)*, pp. 217–224, 2002.
- [Barlow and Proschan 81] R. Barlow and F. Proschan. *Statistical Theory of Reliability and Life Testing*. Silver Spring, MD: To Begin With, 1981.
- [Billingsley 95] P. Billingsley. *Probability and Measure*, 3rd edition. New York: Wiley, 1995.
- [Birkel 89] T. Birkel. “A Note on the Strong Law of Large Numbers for Positively Dependent Random Variables.” *Statistics and Probability Letters* 7 (1989), 17–20.
- [Chakrabarti et al. 08] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos. “Epidemic Thresholds in Real Networks.” *ACM Trans. Inf. Syst. Secur.* 10:4 (2008), 1–26.
- [Denuit and Scaillet 04] M. Denuit and Q. Scaillet. “Nonparametric Tests for Positive Quadrant Dependence.” *Journal of Financial Econometrics* 2 (2004), 422–450.
- [Ganesh et al. 05] A. Ganesh, L. Massoulie, and D. Towsley. “The Effect of Network Topology on the Spread of Epidemics.” In *Proc. INFOCOM'05*, pp. 1455–1466, 2005.
- [Herzberg et al. 97] A. Herzberg, M. Jakobsson, S. Jarecki, H. Krawczyk, and M. Yung. “Proactive Public Key and Signature Systems.” In *ACM Conference on Computer and Communications Security 1997*, pp. 100–110, 1997.
- [Hollander and Douglas 99] M. Hollander and A. Douglas. *Nonparametric Statistical Methods*, 2nd edition. Wiley, 1999.
- [Huang et al. 11] H. Huang, S. Zhang, X. Ou, A. Prakash, and K. Sakallah. “Distilling Critical Attack Graph Surface Iteratively through Minimum-Cost SAT Solving.” In *Proc. ACSAC'11*, pp. 31–40, 2011.
- [Ingols et al. 06] K. Ingols, R. Lippmann, and K. Piwowarski. “Practical Attack Graph Generation for Network Defense.” In *Proc. of the 22nd Annual Conf. on Computer Security Applications*, pp. 121–130, 2006.
- [Jha and Wing 01] S. Jha and J. Wing. “Survivability Analysis of Networked Systems.” In *Proc. 23rd Int'l Conf. Software Eng. (ICSE '01)*, pp. 307–317, 2001.
- [Johnson et al. 94] N. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions*, vol. 1, 2nd edition. New York: Wiley, 1994.
- [Kephart and White 91] J. Kephart and S. White. “Directed-Graph Epidemiological Models of Computer Viruses.” In *IEEE Symposium on Security and Privacy*, pp. 343–361, 1991.
- [Kermack and McKendrick 27] W. Kermack and A. McKendrick. “A Contribution to the Mathematical Theory of Epidemics.” *Proc. Roy. Soc. Lond. A* 115 (1927), 700–721.

- [Lehmann and Romano 05] E. Lehmann and J. Romano. *Testing Statistical Hypotheses*, 3rd edition. Springer, 2005.
- [Li 08] H. Li. “Tail Dependence Comparison of the Survival Marshall–Olkin Copulas.” *Methodology and Computing in Applied Probability* 10 (2008), 39–54.
- [Li et al. 11] X. Li, T. Parker, and S. Xu. “A Stochastic Model for Quantitative Security Analysis of Networked Systems.” *IEEE Transactions on Dependable and Secure Computing* 8:1 (2011), 28–43.
- [Marlow and Tortorella 95] N. Marlow and M. Tortorella. “Some General Characteristics of Two-State Reliability Models for Maintained Systems.” *Journal of Applied Probability* 32 (1995) 805–820.
- [Marshall and Olkin 67] A. Marshall and I. Olkin. “A Multivariate Exponential Distribution.” *Journal of American Statistical Association* 2 (1967), 84–98.
- [McKendrick 26] A. McKendrick. “Applications of Mathematics to Medical Problems.” *Proc. Edin. Math. Society* 14 (1926), 98–130.
- [Mueller and Stoyan 02] A. Mueller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks*. Chichester: Wiley and Sons, 2002.
- [Noel et al. 03] S. Noel, S. Jajodia, B. O’Berry, and M. Jacobs. “Efficient Minimum Cost Network Hardening via Exploit Dependency Graphs.” In *Proc. the 19th Annual Conf. on Computer Security Applications*, pp. 86–95, 2003.
- [Ou et al. 06] X. Ou, W. Boyer, and M. McQueen. “A Scalable Approach to Attack Graph Generation.” In *Proc. ACM Conference on Computer and Communications Security 2006*, pp. 336–345, 2006.
- [Phillips and Swiler 98] C. Phillips and L. Swiler. “A Graph-Based System for Network-Vulnerability Analysis.” In *Proc. Workshop New Security Paradigms (NSPW ’98)*, pp. 71–79, 1998.
- [Provos et al. 07] N. Provos, D. McNamee, P. Mavrommatis, K. Wang, and N. Modadugu. “The Ghost in the Browser: Analysis of Web-Based Malware.” In *Proceedings of the First Workshop on Hot Topics in Understanding Botnets (HotBots ’07)*, pp. 4–13, 2007.
- [Resnick 07] S. Resnick. *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, 2007.
- [Ross 96] S. Ross. *Stochastic Processes*. Wiley and Sons, 1996.
- [Sawilla and Ou 08] R. Sawilla and X. Ou. “Identifying Critical Attack Assets in Dependency Attack Graphs.” In *Proc. ESORICS’08*, pp. 18–34, 2008.
- [Scaillet 05] Q. Scaillet. “A Kolmogorov–Smirnov Type Test for Positive Quadrant Dependence.” *Canadian Journal of Statistics* 33 (2005), 415–427.
- [Sheyner et al. 02] O. Sheyner, J. Haines, S. Jha, R. Lippmann, and J. Wing. “Automated Generation and Analysis of Attack Graphs.” In *Proceedings of the IEEE Symposium on Security and Privacy*, pp. 273–284, 2002.
- [Stoer and Bulirsch 02] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*, 3rd edition. Springer, 2002.

- [Wang et al. 03] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos. “Epidemic Spreading in Real Networks: An Eigenvalue Viewpoint.” In *Proc. of the 22nd IEEE Symposium on Reliable Distributed Systems (SRDS’03)*, pp. 25–34, 2003.
- [Wang et al. 07] L. Wang, A. Singhal, and S. Jajodia. “Measuring the Overall Security of Network Configurations Using Attack Graphs.” In *Prof. DBSec’07*, pp. 98–112, 2007.
- [Wang et al. 08a] L. Wang, T. Islam, T. Long, A. Singhal, S. Jajodia. “An Attack Graph-Based Probabilistic Security Metric.” In *Proc. DBSec’08*, pp. 283–296, 2008.
- [Wang et al. 08b] L. Wang, C. Yao, A. Singhal and S. Jajodia. “Implementing Interactive Analysis of Attack Graphs Using Relational Databases.” *Journal of Computer Security* 16:4 (2008), 419–437.
- [Wang et al. 10] L. Wang, S. Jajodia, A. Singhal, and S. Noel. “ k -Zero Day Safety: Measuring the Security Risk of Networks against Unknown Attacks.” In *Proc. ES-ORICS’10*, pp. 573–587, 2010.
- [Xie et al. 10] P. Xie, J. Li, X. Ou, P. Liu, and R. Levy. “Using Bayesian Networks for Cyber Security Analysis.” In *Proc. DSN’10.*, pp. 211–220, 2010.
- [Xu et al. 12] S. Xu, W. Lu, and L. Xu. “Push- and Pull-Based Epidemic Spreading in Networks: Thresholds and Deeper Insights.” To appear in *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 2012.

Maochao Xu, Department of Mathematics, Illinois State University, Normal, IL 61790-4520, USA (mxu2@ilstu.edu)

Shouhuai Xu, Department of Computer Science, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA (shxu@cs.utsa.edu)