Security Adoption in Heterogeneous Networks: the Influence of Cyber-insurance Market

Zichao Yang and John C.S. Lui
Computer Science & Engineering Department
The Chinese University of Hong Kong

Abstract. Hosts (or nodes) in the Internet often face epidemic risks such as virus and worms attack. Despite the awareness of these risks and the availability of anti-virus software, investment in security protection is still scare, hence, epidemic risk is still prevalent. Deciding whether to invest in security protection is an *inter-dependent process*: security investment decision made by one node can affect the security risk of others, and therefore affect their decisions also. The first contribution of this paper is to provide a fundamental understanding on how "*network externality*" and "*nodes heterogeneity*" may affect security adoption. Nodes make decisions on security investment by evaluating the epidemic risk and the expected loss. We characterize it as a *Bayesian network game* in which nodes only have the local information, e.g., number of neighbors, as well as minimum common information, e.g., degree distribution of the network. Our second contribution is in analyzing a new form of risk management called *cyber-insurance*. We investigate how the presence of competitive insurance market can affect the security adoption and show that if the insurance provider can observe the protection level of nodes, the insurance market is a positive incentive for security adoption provided that the protection quality is not high. We also find that cyber-insurance is more likely to be a good incentive for nodes with higher degree. This work provides the fundamental understanding on the economics aspect of security adoption, and sheds light on a new Internet security service which can be economically viable and sustainable.

Keywords: network science, security adoption, heterogeneous network, cyber-insurance, Bayesian network game

1 Introduction

Network security is a major problem in communication networks. One of its most common manifestations is in form of virus, worms and bonnet spreading, which we call the *epidemic risk*. In these epidemic risks, hosts (or nodes) which are infected become the sources of new infections, and adversaries can use these compromised nodes to generate new attacks. Epidemic risk is highly damaging, e.g., the Code Red worm [22] has infected thousands of computers and induced huge financial loss. To counter this risk, there have been great efforts in both the research and industrial fronts to come up with techniques and tools (i.e., anti-virus software, intrusion detection systems, firewalls etc) to detect virus/worms. Despite the sophistication of these tools, only a small percentage of hosts adopts some form of security protection, hence, epidemic risk is still prevalent.

A node's decision on adopting some security measures is not a simple individual and independent process, but rather, *depends* on the decisions of many other nodes in the network. Nodes which decide not to invest in security protection, also put other nodes at security risk. This *network externality effect* caused by the spreading of epidemic influences the degree of adoption of security measure. *Our first contribution in this paper* *is to provide a theoretical understanding on how the two factors: network externality and node heterogeneity, may influence security adoption in a network of interconnected nodes (i.e., the Internet) where virus/worms can propagate.*

Modeling such decision and security problem requires the combination of epidemic theory and game theory. While extensive studies in traditional literatures have been dedicated to epidemic theory [23], few works have addressed the problems of strategic behavior of security investment. In realistic situation, nodes which make decision in security investment usually do not have complete information like the network topology or knowledge of other nodes. So it is difficult for them to accurately evaluate the epidemic risk and other nodes' influence on itself. In this paper, we model the security investment as a *Bayesian network game* where nodes only have the local information of their degree, as well as a minimum common information of network's degree distribution. In contrast to graphical game [24], in which complete topology is given and analysis is complicated, we show that using Bayesian network game, one can elegantly tradeoff using partial topology information while making the analysis tractable. *We show how heterogeneous nodes, characterized by their degree, can estimate their epidemic risk and make decisions on security investment with incomplete information*. We show that nodes with higher degree are more likely to be infected by epidemic, making the secure measure less effective for them in terms of the reduction in infection probability. Moreover, nodes with higher degrees are more sensitive to externality effect, i.e., they are more likely to be affected by others' decision. The final fraction of nodes which adopt the security measure strongly depends on node degrees and their relative loss from epidemic.

While protection measures may limit the spread of virus/worms, another way to manage the epidemic risk is to transfer the risk to a third-party, which is called *cyber-insurance* [13]: nodes pay certain premium to insurance companies in return for compensation in the virus outbreaks. One main challenge in cyber-insurance is *moral hazard* [11, 13]. The combination of self-protection and insurance raises the problem of moral hazard, in which nodes covered by insurance may take fewer secure measures, or even falsify their loss. Moral hazard happens when the insurance provider cannot observe the protection level of nodes. In this paper, we investigate the effect of cyber-insurance on security adoption under competitive insurance market without moral hazard[1]. *Our second contribution is to show the conditions under which cyber-insurance is an incentive*. We find that cyber-insurance without moral hazard is an incentive for security adoption if the initial secure condition is poor and the quality of secure measure is not very high. Moreover, cyber-insurance is more likely to be an incentive for high degree nodes.

This is the outline of our paper. In Section 2, we present the epidemic and security investment models. In Section 3, we show how heterogeneous nodes can determine their infection probability and decide on proper security investment. In Section 4, we investigate the effect of insurance market without moral hazard on security adoption. Validations and performance evaluations are presented in Section 5. Section 6 gives related work and Section 7 concludes.

[1] We refer the readers to [28] for the analysis of the case with moral hazard.

2 Mathematical Models

Let us first present the mathematical models on how nodes make decision on security investment. Our models include: (a) *epidemic model*: to characterize the spread of virus or malware in a network, (b) *investment model*: to characterize node's decision in security investment, and (c) *Bayesian network game*: how nodes make decision under the incomplete information setting.

Epidemic Model: The interaction relation of N nodes is denoted by the undirected graph $G = (V, E)$ with the vertex set V, $|V| = N$ and the edge set E. For $i, j \in V$, if $(i, j) \in E$, then nodes i and j are neighbors and we use $i \sim j$ to denote this relationship. Let $S = \{\textit{healthy}, \textit{infected}\}$ represents the set of states each node can be in. If node i is infected (healthy), then $S_i = 1$ ($S_i = 0$). Each infected node can contaminate its neighbors independently with probability q. Note that this is similar to the *bond percolation process* [23] in which every edge is occupied with probability q. Each node has an *initial state* of being infected or not. This can represent whether the node has been attacked by the adversary. Let us denote it by s_i where $s_i = 1$ if node i is initially infected and $s_i = 0$ otherwise. Hence, at the steady state, a node is infected either because it is initially infected, or it contracts virus from its infected neighboring nodes. The final state of node i can be expressed in the following recursive equation:

$$1 - S_i = (1 - s_i) \prod\nolimits_{j \sim i} (1 - \theta_{ji} S_j) \quad \forall i \in V, \tag{1}$$

where θ_{ji} is a random variable indicating whether the edge (i, j) is occupied or not. Based on the above discussion, θ_{ji} is a Bernoulli random variable with $\Pr(\theta_{ji} = 1) = q$. Now, given the network topology G and the probabilities of infection, every node can evaluate the probability that it will eventually be infected. Since the infection will incur some financial loss, a node needs to decide whether to invest in self-protection to reduce the potential financial loss. In the following, we present a model such that a node can use it to make the decision.

Investment Model: Let's say that node i has an initial wealth $w_i \in \mathbb{R}_+$. A node's utility $u_i(w)$ is a function of wealth $w \in \mathbb{R}_+$. We consider nodes are *risk averse*, i.e., the utility function is strictly increasing and concave in w, i.e., $u'_i(w) > 0$ and $u''_i(w) < 0$. In this paper, we consider the *constant relative risk averse* utility function commonly used in the economic literature [4]: $u(w) = w^{1-\sigma}/(1-\sigma)$, for $0 < \sigma < 1$, where σ is a parameter for the degree of risk aversion. The condition $0 < \sigma < 1$ is added to eliminate the case of $\sigma = 1$ and also for tractability of analysis later on. For node i, the utility function is given by the above utility function with parameter σ_i. If node i is infected, then it will incur a financial loss of $l_i \in \mathbb{R}_+$.

To reduce the potential financial loss, a node can consider adopting some self-protection measures or purchasing insurance. In the first part of this paper, we consider the case of self-protection. In the second part of this paper, we consider both cases and study the influence of insurance market on security protection. A node's investment in self-protection can reduce the probability of being infected initially. For the amount of investment x, the probability of being infected initially is denoted as $p(x)$, which is a continuous differentiable decreasing function of x. In particular, we assumed the effort of security investment is separable with the wealth, which is a common standard in the literature [27]. We have

$$p_i u_i(w_i - l_i) + (1 - p_i) u_i(w_i) - x_i, \tag{2}$$

where p_i is the *final* probability that node i will be infected. p_i contains two parts: the probability of being infected initially, given by $p(x_i)$ and the probability of getting infected from neighbor nodes. For simplicity of analysis, we assume that the choice of node i regarding security self-protection is a binary decision: either the node invests unit amount with a cost of c_i, or it does not invest at all. We use the action set $A = \{\mathcal{S}, \mathcal{N}\}$ to denote the behavior, where $\mathcal{S}$ denotes taking secure measure while $\mathcal{N}$ denotes not taking secure measure. If it decides to invest, the node can still be infected with probability p^-. Otherwise, it will be infected with probability p^+. Obviously we have $0<p^-<p^+<1$. Let $a = (a_1, ..., a_i, ..., a_N)= (a_i, a_{-i})$ be an *action profile*. Given the action profile a_{-i} of other nodes, node i makes the decision by maximizing its expected utility. If node i takes action $\mathcal{N}$, the expected utility is: $p_i(\mathcal{N}, a_{-i})u_i(w_i-l_i)+(1-p_i(\mathcal{N}, a_{-i}))u_i(w_i)$, where $p_i(\mathcal{N}, a_{-i})$ is the final probability of node i being infected when it initially did not adopt security protection. On the other hand, the expected utility of a node which initially subscribed to security protection (or action $\mathcal{S}$) is: $p_i(\mathcal{S}, a_{-i})u_i(w_i - l_i) + (1 - p_i(\mathcal{S}, a_{-i}))u_i(w_i) - c_i$ where $p_i(\mathcal{S}, a_{-i})$ is the final probability of a node being infected when it initially subscribed to some self-protection measures with cost of c_i.

Each node needs to consider whether it should subscribe to some self-protection measures. The decision is based on the cost of investing in security measure, as well as the risk loss of being infected. The decision is non-trivial because one has to consider the *network externality effect*. In particular, node i will choose to invest in security protection if and only if

$$c_i < (p_i(\mathcal{N}, a_{-i}) - p_i(\mathcal{S}, a_{-i}))(u_i(w_i) - u_i(w_i - l_i)). \tag{3}$$

Note that the inequality is a function of a_{-i}, and this shows that a node's decision is based on the action of other nodes.

Bayesian Network Game: According to Inequality (3), each node needs to have the complete information of the network topology G so to make the proper decision. However, it is almost impossible in practice for each node to have the complete information of G. Instead, each node can only have some *local information* on G, i.e., a node may only know its neighbors, and some cases, only knows the number of neighbors it is to interact with. Secondly, it is impossible to know the exact loss of other nodes in a large network.

In here, we assume that nodes only have a *minimum common information*, that is, the knowledge of the degree distribution of G, as well as the distribution of financial loss of nodes caused by virus. Assume that the degree distribution of the graph is $\{p_k\}_{\underline{K}}^{\overline{K}}$, where $\overline{K}$ is the maximum degree and $\underline{K}$ is the minimum degree. In this paper, we consider the *asymptotic case* that N, the number of nodes, tends to infinity and the degree distribution converges to the fixed probability distribution $\{p_k\}_{\underline{K}}^{\overline{K}}$. For nodes with degree k, the loss distribution is given by the CDF $F_k(l)$. We assume that the cost of secure measure is the same for all nodes which have the same degree, we denoted this as c_k. Furthermore, these nodes have the same utility function u_k and the same initial wealth w_k. Nodes make decision on security investment based on the information of degree and loss. According to the discussion in the investment model, a node should know the probability of getting infected before deciding on security investment. Since nodes do not have the complete information, they should estimate these probabilities based on the limited common information. Next, we derive this infected probability using the *local mean field technique* [1,18].

3 Analysis for Strategic Security Self-protection

Let us show how nodes make decisions on security investment and how to determine the final security protection level. Determining the final infection probability is a difficult problem because of the complex network structure. In this work, we assume that a node only knows the degree distribution and consider the network topology as a *random graph* [23] with a given degree distribution $\{p_k\}_{\underline{K}}^{\overline{K}}$. Thus, nodes do not need to know the full network topology G to determine the final infection probability. Although real networks are not random graphs [23] and they have some characteristics, e.g., high clustering coefficient, community structure etc, that are not possessed by random graph. Recent study [20] has shown that random graph is very often accurate for real networks. Thus, it is reasonable to assume that the network topology is a random graph, especially here we consider incomplete information. Note that in [18], we also consider extension of random graphs with high clustering coefficient.

Estimating the Probability: A node can calculate its final infection probability by constructing a *local mean field tree* [1,18]. Fig. 1 illustrates the local mean field of node i which has degree k. For the ease of presentation, let's say that none of these nodes will take secure measure, i.e., the initial infection probability is p^+ for all nodes in this subsection. We will show how to relax this in later section.

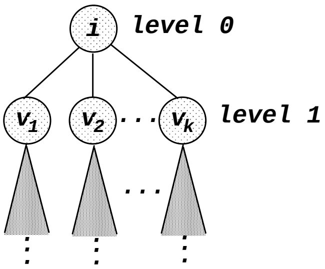


Fig. 1: local mean field tree for node i with degree k

The children of node i in the local mean field tree are denoted as $v_c, c \in [1,k]$. The triangle under each child node v_c denotes another tree structure. Based on the results in [1,18], for any node i, the local topology of a large random graph G can be modeled as a tree rooted at node i with high probability. In other words, we transform G to a tree rooted at node i (or local mean field of node i). Node i can be independently influenced by each subtree rooted at v_c. For every subtree rooted at v_c, it consists of its subtrees. Using this *recursive* structure, one can derive the total infection probability that other nodes in G can impose on node i.

Let us illustrate the derivation. First, we divide nodes into levels. The root node i is at the zero level. The neighbors of node i is at the first level and so on. Let Y_j be the final state of node $j, j \neq i$, *conditioned on its parent in the tree structure is not infected*, and y_j be the initial state of node j. For the root node i, we use S_i to denote its final state and s_i to denote its initial state, then we have

$$1 - S_i = (1 - s_i) \prod\nolimits_{j \sim i} (1 - \theta_{ji} Y_j). \tag{4}$$

The above equation indicates the root node i is either initially infected, or it can be infected by its neighbors. The state of its neighbors conditioned on that the root node i is not infected is also determined by the state of the children of the neighbors in the tree structure, or one can express it recursively as:

$$1 - Y_j = (1 - y_j) \prod\nolimits_{l \to j} (1 - \theta_{lj} Y_l) \quad j \neq i, \tag{5}$$

where $l \to j$ denotes that l is a child of j in the tree structure. To solve Eq. (5), we only need to know the degree distribution of a child node. This degree distribution can be expressed as: $\tilde{p}_k = \frac{kp_k}{\sum_{k=\underline{K}}^{\overline{K}} kp_k} = \frac{kp_k}{\bar{d}}$, where $\bar{d}$ is the average degree of nodes in G. The number of edges of a child other than the edge connecting to its parents is called the *excess degree* [23]. Let $\underline{K}' = \max\{0, \underline{K}{-}1\}$ and $\overline{K}' = \max\{0, \overline{K}{-}1\}$. The excess degree distribution of a child is $q_k = \tilde{p}_{k+1} = (k+1)p_{k+1}/\bar{d}$, for $k \in [\underline{K}', \overline{K}']$.

As in [1, 18], if nodes are at the same level of the tree structure, their states are independent of each other. Let $\rho_n, n \geq 1$ be the probability that a node at the n^{th} level is infected conditioned on its parent is not infected. By Eq. (5), we have

$$1 - \rho_n = (1 - p^+) \sum\nolimits_{k=\underline{K}'}^{\overline{K}'} q_k (1 - q\rho_{n+1})^k.$$

When we scale up the network (or let $n \to \infty$), we get ρ, the average probability that a child node of the root node i will be infected conditioned on the root node is not infected, which can be determined by the solution of the fixed point equation

$$1 - \rho = (1 - p^+) \sum\nolimits_{k=\underline{K}'}^{\overline{K}'} q_k (1 - q\rho)^k.$$

By Eq. (4), for a node with degree k, the infection probability under the condition of incomplete information is

$$\phi_k = 1 - (1 - p^+)(1 - q\rho)^k. \tag{6}$$

Security Adoption: In the previous subsection, we show how a node can compute the infection probability with incomplete information. The calculation is based on the assumption that none of the nodes take secure adoption, so that the initially infected probability is p^+. In here, we show how to use this infection probability for strategy selection. Let λ_k be the fraction of nodes with degree k which take action $\mathcal{S}$. Then by applying the method shown above, we have

Proposition 1. *If λ_k fraction of the nodes with degree k will take secure measure, let ρ be the probability a child node of the root node will get infected conditioned that its parent is not infected, is given by the unique solution of the fixed point equation in* $[0, 1]$*:*

$$\rho = 1 - \sum\nolimits_{k=\underline{K}'}^{\overline{K}'} q_k (1 - p^+ + \lambda_{k+1}(p^+ - p^-))(1 - q\rho)^k. \tag{7}$$

For a node with degree k, if it decides to take secure measure, then by Eq. (6), the infection probability is $\phi_k(\mathcal{S}, \lambda_{\underline{K}}, ..., \lambda_{\overline{K}}) = 1 - (1 - p^-)(1 - q\rho)^k$. If it does not invest in protection measure, the probability for this node to get infected is $\phi_k(\mathcal{N}, \lambda_{\underline{K}}, ..., \lambda_{\overline{K}}) = 1 - (1 - p^+)(1 - q\rho)^k$. The infection probability reduction for a node with degree k is

$$\phi_k(\mathcal{N}) - \phi_k(\mathcal{S}) = (p^+ - p^-)(1 - q\rho)^k. \tag{8}$$

Note that infection probability reduction decreases as degree increases. This implies that higher degree nodes have *less incentive* to invest in protection measure.

Lemma 1. *ρ, which is given by the solution of fixed point Eq. (7), has an unique solution in $[0,1]$, and $\rho(\lambda_{\underline{K}}, ..., \lambda_{\overline{K}})$ is a decreasing function of λ_k, $\forall k \in [\underline{K}, \overline{K}]$.*

Due to the limitation of space, we refer the readers to [28] for the proofs of propositions and lemmas in this paper.

Remark: Combining Lemma 1 with Eq. (8), we see that the reduction in infection probability by taking security measure increases as other nodes takes on security measure. This implies the impact of the *network externality effect*, i.e., the value of security measure increases as more nodes invest in self-protection.

Sensitivity Analysis: Nodes with different degree have different sensitivity to the externality effect. Define $\widetilde{\phi}_k = \phi_k(\mathcal{N}) - \phi_k(\mathcal{S}) = (p^+ - p^-)(1-q\rho)^k$. Assume ρ decreases by a small amount $\varDelta\rho$, then $\varDelta\widetilde{\phi}_k = (p^+ - p^-)(1-q\rho)^{k-1}kq\varDelta\rho$, the relative change is given by $\frac{\varDelta\widetilde{\phi}_k}{\widetilde{\phi}_k} = \frac{kq\varDelta\rho}{(1-q\rho)}$, indicating that nodes with degree k are k times sensitive to the network externality effect than nodes with degree 1.

A node with degree k will invest if and only if the utility with secure measure is higher than that of without secure measure, or

$$\begin{aligned} c_k <& (\phi_k(\mathcal{N}) - \phi_k(\mathcal{S}))(u_k(w_k) - u_k(w_k - l)) \\ =& (p^+ - p^-)(1-q\rho)^k(u_k(w_k) - u_k(w_k - l)) \end{aligned}$$

Note that the loss distribution of nodes with degree k is $F_k(l)$. Since the infection probability is varying with the fraction of security adopters, we consider the *self-fulfilling expectations equilibrium* [6] in analyzing the final adoption extent. Nodes form a shared expectation that the fraction of the nodes has adopted security measure and if each of them makes decision based on this expectation, then the final fraction is indeed the initial expectation. Let l_k^* be the minimum value that satisfies the above inequality in the equilibrium, then λ_k^*, the fraction of node of degree k will take the secure measure, is given by the equation $\lambda_k^* = 1 - F_k(l_k^*)$. Summarizing the previous analysis, we have the following proposition.

Proposition 2. *Nodes with degree k will take the secure measure if their loss is greater than l_k^*. The final fraction of nodes with degree k that will invest in self-protection is λ_k^*. l_k^* and λ_k^* are solutions of the following fixed point equations:*

$$\lambda_k^* = 1 - F_k(l_k^*), \tag{9}$$

$$c_k = (p^+ - p^-)(1-q\rho^*)^k(u_k(w_k) - u_k(w_k - l_k^*)), \tag{10}$$

where ρ^* is given by the solution of the following equation

$$\rho^* = 1 - \sum\nolimits_{k=\underline{K}'}^{\overline{K}'} q_k(1 - p^+ + \lambda_{k+1}^*(p^+ - p^-))(1-q\rho^*)^k. \tag{11}$$

Lemma 2. *Fixed point equations (9)–(11) has at least one solution.*

Lemma 3. *The equilibrium points given by fixed point equations (9)–(11) are monotone, i.e., If $\mathbf{\Lambda}^{*1} = (\lambda_{\underline{K}}^{*1}, ..., \lambda_k^{*1}, ..., \lambda_{\overline{K}}^{*1})$ and $\mathbf{\Lambda}^{*2} = (\lambda_{\underline{K}}^{*2}, ..., \lambda_k^{*2}, ..., \lambda_{\overline{K}}^{*2})$ are two equilibrium points, then we have either $\mathbf{\Lambda}^{*1} \geq \mathbf{\Lambda}^{*2}$ or $\mathbf{\Lambda}^{*1} \leq \mathbf{\Lambda}^{*2}$ and there exists at least one $k \in [\underline{K}, \overline{K}]$ such that $\lambda_k^{*1} \neq \lambda_k^{*2}$.*

Remark: The above lemmas prove the existence and monotonicity of equilibrium points. In our technical report [28], we analyzed the multiplicity of the equilibrium points by considering a special case. Readers can refer to the report to get more insights.

4 Analysis for Cyber-insurance Market

In here, we consider *cyber-insurance* and analyze its impact on security adoption.
A. **Supply of Insurance**

Let's say the insurance provider offers insurance at the price of $\pi < 1$. Nodes which buy insurance at the premium of πX from the insurance provider will be compensated X for the loss incurred if they are infected. Given the price π, node will choose to buy the amount of insurance that maximizes its utility. Define $\phi_k(\mathcal{S})(\phi_k(\mathcal{N}))$ as the probability that a node with degree k will be infected if it subscribes (does not subscribe) to a secure measure. In this paper, we consider cyber-insurance without adverse selection, in which the insurance provider can observe the degree of a node, hence the risk type of a node (high degree indicates high risk level). Thus, in the following, we drop the subscript k where the meaning is clear for general presentation. A node will choose the amount of insurance that maximizes

$$U(\pi, X) = \phi u(w - l + (1 - \pi)X) + (1 - \phi)u(w - \pi X) - x, \tag{12}$$

where x is the wealth spent on security protection. When a node chooses $\mathcal{N}$, ϕ becomes $\phi(\mathcal{N})$, $x = 0$. When a node chooses $\mathcal{S}$, ϕ becomes $\phi(\mathcal{S})$, $x = c$. Assume the insurance provider is risk neutral, so they only care about the expected wealth. If a node buys X amount of insurance, then the profit of the insurance is $(\pi - \phi)X$. In here, we consider a competitive market so the insurance provider has to offer the insurance at the price $\pi = \phi$, or the *actuarially fair price* [7].

Lemma 4. *When the insurance is offered at the actuarially fair price, the optimal insurance coverage is a full insurance coverage, i.e., a node will buy insurance amount equal to the loss l. The maximal expected utility is $u(w - \phi l) - x$, i.e., when a node chooses $\mathcal{N}$, the maximal expected utility is $u(w - \phi(\mathcal{N})l)$, when a node chooses $\mathcal{S}$, the maximal expected utility is $u(w - \phi(\mathcal{S})l) - c$.*

Lemma 5. *When the insurance is offered at price $\pi > \phi$, the optimal insurance coverage is partial insurance coverage, i.e., a node will buy insurance coverage less than l. The maximal expected utility is $u(w – \phi l – \delta(\phi, \pi)) – x$, where $\delta(\phi, \pi) > 0$.*

Remark: Lemma 4 shows that the expected utility without insurance market is $u(w - \phi l - r(\phi, x)) - x < u(w - \phi l) - x$. The utility of a node is *improved* by the insurance market with the fair price. But if the contract is at an unfair price, the utility improvement is smaller according to Lemma 5.

One problem with the combination of insurance and self-protection is *moral hazard*, which happens when the insurance provider cannot observe the protection level of a node. Insurance coverage may discourage the node to take self-protection measure to prevent the losses from happening, or even to encourage nodes to cause the loss and make insurance claims. In here, we examine the effect of the insurance market on the self-protection level. In this paper, we consider the case without moral hazard, where the insurance provider can observe the protection level of a node. We refer the reader to [28] for the analysis of the case with moral hazard, where insurance provider does not have any information about the protection level of a node. Without the moral hazard, the insurance provider can discriminate against the nodes with protection measure and those without protection measure. We investigate whether the insurance market will help to incentivize nodes to take secure measure.

B. Cyber-insurance Without Moral Hazard

Security Adoption with Insurance Market: Because the insurance provider can observe the protection level of a node, the insurance provider will offer insurance price of $\phi(\mathcal{S})$ (or $\phi(\mathcal{N})$) for those nodes with (or without) security protection. According to Lemma 4, nodes will buy the full insurance regardless of its protection level. As a result, the expected utility for nodes without protection is $u(w-\phi(\mathcal{N})l)$ and the expected utility for nodes with protection is $u(w-\phi(\mathcal{S})l)-c$. Thus, with insurance market, a node will invest in security protection if and only if

$$c < g(l,\rho) \triangleq u(w-\phi(\mathcal{S})l) - u(w-\phi(\mathcal{N})l).$$

Here $g(l,\rho)$ is a function of ρ because $\phi(\mathcal{S})$ and $\phi(\mathcal{N})$ can be expressed in ρ.

Lemma 6. *The function $g(l,\rho) \triangleq u(w-\phi(\mathcal{S})l) - u(w-\phi(\mathcal{N})l)$ is increasing with respect to the loss l.*

Lemma 7. *$g(l,\rho) \triangleq u(w-\phi(\mathcal{S})l) - u(w-\phi(\mathcal{N})l) = u(w-(1-(1-p^-)(1-q\rho)^k)l) - u(w-(1-(1-p^+)(1-q\rho)^k)l)$ is a decreasing function with ρ.*

Lemma 6 indicates that nodes with higher loss are more likely to invest in security. Lemma 7 shows that positive network externality still exists even in the presence of insurance market. Similar to the analysis in Sec. 3, we can arrive in the following proposition regarding the adoption fraction with insurance market:

Proposition 3. *With insurance market, nodes with degree k will take the secure measure if their loss is greater than $l_k^{*\mathtt{I}}$. The final fraction of nodes with degree k that will invest in self-protection is $\lambda_k^{*\mathtt{I}}$. $l_k^{*\mathtt{I}}$ and $\lambda_k^{*\mathtt{I}}$ are solutions of the following fixed point equations:*

$$\lambda_k^{*\mathtt{I}} = 1 - F_k(l_k^{*\mathtt{I}}), \tag{13}$$

$$c_k = u_k(w_k - \phi(\mathcal{S})l_k^{*\mathtt{I}}) - u_k(w_k - \phi(\mathcal{N})l_k^{*\mathtt{I}}), \tag{14}$$

where $\rho^{*\mathtt{I}}$ is given by the solution of the following equation

$$\rho^{*\mathtt{I}} = 1 - \sum_{k=\underline{K}'}^{\overline{K}'} q_k(1-p^+ + \lambda_{k+1}^{*\mathtt{I}}(p^+ - p^-))(1-q\rho^{*\mathtt{I}})^k. \tag{15}$$

Previous lemmas on the existence and monotonicity of equilibrium points also holds here. Comparing Proposition 3 with Proposition 2, we can recognize the only difference lies in Eq. (14) and Eq. (10). Buying insurance improves node's utility, hence changes their decision on security protection as well. In the following, we examine the effect of insurance market on security adoption. An overall and detailed analysis needs calculating out all the equilibrium points and comparing the equilibrium points specified by the two propositions. which is quite complicated. Instead, we examine the effect from the local point of view, but still provide enough insight.

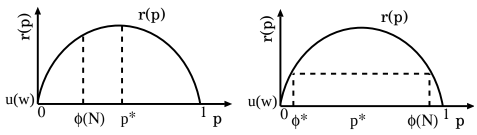
Incentive Analysis: According to previous analysis, a node will take secure measure if $c < c_{NI} \triangleq (\phi(\mathcal{N}) - \phi(\mathcal{S}))(u(w) - u(w-l))$, here c_{NI} is the threshold without insurance market. With insurance market, nodes will take secure measure if and only if $c < c_I \triangleq u(w - \phi(\mathcal{S})l) - u(w - \phi(\mathcal{N})l)$, where c_I denotes the threshold with insurance market. In order for insurance market to be a good incentive for self-protection, we should have $c_{NI} < c_I$, i.e.,

$$\begin{aligned} c_I - c_{NI} =& u(w - \phi(\mathcal{S})l) + \phi(\mathcal{S})(u(w) - u(w - l)) \\ & - [u(w - \phi(\mathcal{N})l) + \phi(\mathcal{N})(u(w) - u(w - l))] > 0. \end{aligned}$$

Define $r(p) \triangleq u(w - pl) + p(u(w) - u(w - l))$, the above condition becomes $r(\phi(\mathcal{S})) > r(\phi(\mathcal{N}))$. Next we investigate under what condition the above inequality will hold. Consider the function $r(p)$, we have the following lemma.

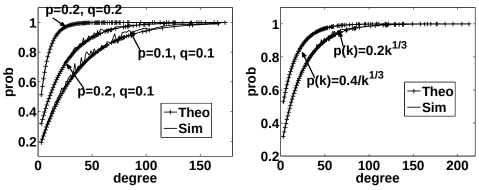
Lemma 8. *$r(p)$ is a concave function of p and has an unique max value at p^*.*

Proposition 4. *If the infection probability without secure measure $\phi(\mathcal{N})$ is greater than p^* and the quality of self-protection is not too high, i.e., $\phi(\mathcal{N}) - \phi(\mathcal{S})$ is bounded, insurance will be a good incentive for self-protection.*



(a) non-incentive (b) incentive

Fig. 2: Thresholds of $\phi(S)$



(a) homogeneous (b)heterogeneous

Fig. 3: Verifying local mean field

Fig. 2a shows the case where $\phi(\mathcal{N})$ is smaller than p^*. Cyber-insurance is not an incentive for security. In Fig. 2b, $\phi(\mathcal{N})$ is greater than the p^*, if $\phi(\mathcal{S})$ is with in the region $[\phi^*, \phi(\mathcal{N})]$, then cyber-insurance is a good incentive. From the figure, we can see that insurance will be more likely to be incentive with big $\phi(\mathcal{N})$ and small $\phi(\mathcal{N}) - \phi(\mathcal{S})$. Hence, if the initial secure situation is bad and the protection quality of secure measure is not too high, then insurance market is a positive incentive for self-protection; otherwise, insurance market is a negative incentive, i.e., if a node adopts secure measure without insurance, it may decide not to adopt secure measure with insurance market.

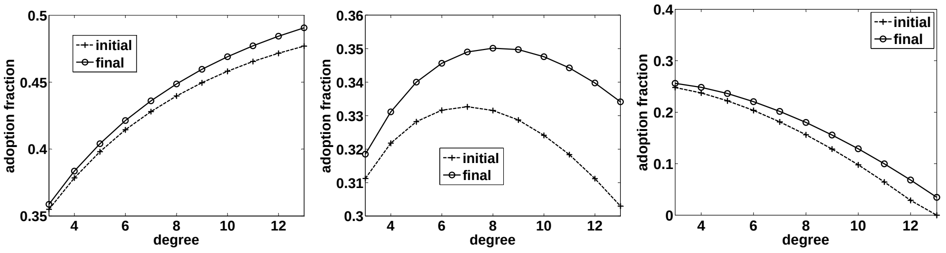
We can study the effect of cyber-insurance on nodes with *different degree* based on above analysis. For $k_1 < k_2$, we have $\phi_{k_1}(\mathcal{S}) < \phi_{k_2}(\mathcal{S})$, $\phi_{k_1}(\mathcal{N}) < \phi_{k_2}(\mathcal{N})$ and $\phi_{k_1}(\mathcal{N}) - \phi_{k_1}(\mathcal{S}) > \phi_{k_2}(\mathcal{N}) - \phi_{k_2}(\mathcal{S})$. In other words, nodes with higher degree have higher infection probability $\phi(\mathcal{N})$ and the protection measure will be less effective to nodes with higher degree. As a result, insurance market will be more likely to be incentive for nodes with higher degree. (A quantitative conclusion needs to examine the influence of wealth and loss difference of nodes.)

Whether insurance will be an incentive greatly depends on the parameters. Generally speaking, cyber-insurance can be positive insurance for all nodes, negative insurance for all nodes and negative incentive for low degree nodes, but positive incentive for high degree nodes. We provide extensive numerical results in the Section 5 to demonstrate the above cases.

5 Simulation and Numerical Results

Validate Final Infection Probability: We consider a large graph with power-law distribution [8]. We want to verify the accuracy of using the mean field on these power law graphs. We use the popular *Generalized Linear Preference (GLP)* method to generate power law graphs [5]. Parameters were selected so that the power law exponent $\gamma = -3$. We generate graphs with $10,000$ nodes and approximately $30,000$ edges. The minimum degree is 3 and the maximum degree is approximately 200. First, we verify the case when all the nodes have the same probability of being infected initially. The result is shown in Fig. 3a. Initially, every node is infected with the same probability p and every edge is occupied with probability q. We calculate the probability that nodes with certain degree is infected. Fig. 3a shows the simulation verifies the theoretical results.

Next, Fig. 3b shows the infection probability of nodes with different degree under different initial infection probability. For both curves, q is set to 0.1. For the curve above, we set the initial probability for nodes with degree k to be $p(k) = 0.4/k^{\frac{1}{3}}$. The probability decreases with degree. For the curve below, we set the initial infection probability to be $p(k) = 0.2k^{\frac{1}{3}}$. The probability increases with degree. From the figure, we see that the local mean field technique is very accurate and the theoretical results accurately match with simulation results.

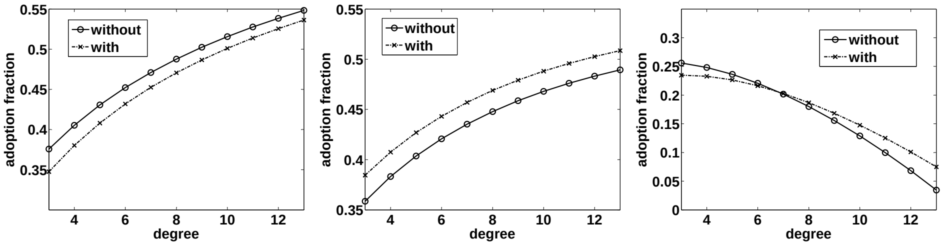


(a) $q = 0.05$ (b) $q = 0.10$ (c) $q = 0.15$

Fig. 4: adoption fraction of nodes with different degree

Security Adoption: Let us investigate how different parameters can influence the fraction of adopters with different degree. We consider a graph G with power law distribution with $\gamma = -3$, minimum and maximum degree are 3 and 13. Here maximum degree is set small for the convenience of selecting other parameters without affecting the result in general case. For example, with very large maximum degree, even a small q will make the final infection probability in Eq. 6 very big because of the power relationship. We set $\sigma = 0.5$ for all nodes. The initial wealth of nodes with degree k is $w_k = 10 * k + 50$. The loss follows uniform distribution from 0 to half of the initial wealth. The cost of secure measure of all nodes is $c = 0.3$. Initially, all nodes without (with) secure measure are infected initially with probability $p^+ = 0.3$ ($p^- = 0.2$). Having fixed the above parameters, we choose to change q to calculate the fraction of adopters with different degree because nodes with different degree are mainly differentiated via the term $(1 - q\rho)^k$. We want to examine the effect of heterogeneity by setting different q.

We show the initial fraction and final fraction of adoption in Fig. 4. Here the initial fraction means that every node assumes that other nodes will not adoption secure measure and makes its decision on this assumption. Final fraction means the fraction given by the minimum equilibrium point in Proposition 2. Due to the positive externality effect, final fraction is greater than initial fraction. We plot them to examine the externality effect. From Fig. 4a to Fig. 4c, we set q to be $0.05, 0.10$ and 0.15 respectively. In Fig. 4a, the adoption fraction increases with degree, in Fig. 4b, the adoption fraction first increases with degree, then decreases with degree, while in Fig. 4c , the adoption fraction decreases with degree. Comparing these three figures, we see that there is no general rule regarding the adoption of nodes as a function of the degree. It greatly depends on the parameters. However, we can see in all figures that the gap between the final adoption fraction and the initial adopt fraction increases with degree, indicating nodes with higher degree will be incentivized better than nodes with lower degree. This agrees with our previous result that higher degree nodes are more sensitive to the externality effect.



(a) negative (b) positive (c) negative & positive

Fig. 5: adoption fraction of nodes with different degree

Influence of cyber-insurance: We claim in previous section that insurance can be negative incentive for all nodes, positive incentive for all nodes and negative incentive for low degree nodes but positive incentive for high degree nodes. We demonstrate these cases through numerical results. In Fig. 5a, we set the parameters $p^+ = 0.3$, $p^- = 0.2$ and $q = 0.02$. We see that the fraction of nodes which adopt the secure measure without insurance market is greater than that with insurance market. This is because the infection probability $\phi(\mathcal{N})$ is low. In Fig. 5b, we set the parameters $p^+ = 0.8$, $p^- = 0.7$ and $q = 0.02$. As the figure shows, insurance market is positive incentive. In this case, the infection probability $\phi(\mathcal{N})$ is high and the protection quality is low. In Fig. 5c, we set the parameters $p^+ = 0.3$, $p^- = 0.2$ and $q = 0.15$. In contrast to Fig. 5a, the q is greater, making the infection probability $\phi(\mathcal{N})$ for low degree nodes small while for high degree nodes big. Thus insurance is negative incentive for low degree nodes, but positive incentive for high degree nodes.

6 Related Work

Recently there has been growing research in the economic of information security [2,3]. Some models consider the security investment game without incorporating the effect of network topology, i.e., [9, 10, 14]. Others assume that the graph topology is given

[12, 21, 25]. [15, 16] are the closely related to our work. The network topology is modeled as a Poisson random graph while real networks are with power law distribution. They assume that all nodes are the same to examine the average effect and do not consider the interaction among those nodes. In contrast, we consider the interaction of nodes by studying a Bayesian network game. Our modeling result provides significant insight on the influence of heterogeneity. [29] is our previous extended abstract in considering network heterogeneity, which is defined by dividing the nodes into classes by setting degree thresholds. Insurance was studied in the economic literature long time ago [7] [26]. But these literatures lack to consider many characteristics specific to computer network, such as the interdependence of security, heterogeneity considered in this work. Cyber-insurance was proposed to manage security risk [19] but is only modeled recently [13, 17, 27]. A key concern is that whether cyber-insurance is an incentive for security adoption. In [17], the authors do not consider the heterogeneity in modeling cyber-insurance. [27] assume the effort on security protection is continuous and did not consider the network topology.

7 Conclusion

Modeling strategic behavior in security adoption helps us to understand what are the factors that could result in under investment. In this paper, we show, via a Bayesian network game formulation, how “network externality” and “node heterogeneity” can affect security adoption in a large communication network. We also investigate the effect of cyber-insurance on protection level. We establish the conditions under which cyber-insurance is a positive incentive for security adoption. This work provides the fundamental understanding on the economics aspect of security adoption, and sheds light on a new Internet service which is economically viable.

Acknowledgement: This research is supported in part by the RGC 415310.

References

1. D. Aldous, A. Bandyopadhyay. Survey of max-type recursive distributional equations. *The Annals of Applied Prob.*, 15(2):1047–1110, ’05.
2. R. Anderson. Why information security is hard-an economic perspective. In *IEEE Computer Security Applications Conference’01*, pages 358–365.
3. R. Anderson and T. Moore. The economics of information security. *Science*, 314(5799):610, 2006.
4. R. Böhme and G. Schwartz. Modeling cyber-insurance: Towards a unifying framework. In *Workshop on the Economics of Information Security, Harvard University, Cambridge (June 2010).*
5. T. Bu and D. Towsley. On distinguishing between internet power law topology generators. In *INFOCOM*, pages 638–647. IEEE, 2002.
6. D. Easley and J. Kleinberg. *Networks, crowds, and markets: Reasoning about a highly connected world.* Cambridge Univ Pr, 2010.
7. I. Ehrlich and G. Becker. Market insurance, self-insurance, and self-protection. *The Journal of Political Economy*, 80(4):623–648, 1972.
8. M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *ACM SIGCOMM*, pages 251–262, 1999.

9. J. Grossklags, N. Christin, and J. Chuang. Secure or insecure? a game-theoretic analysis of information security games. In *WWW'08*.
10. G. Heal and H. Kunreuther. The vaccination game. *Center for Risk Management and Decision Process Working Paper*, 2005.
11. B. Hillier. *The economics of asymmetric information*. Palgrave Macmillan, 1997.
12. L. Jiang, V. Anantharam, and J. Walrand. Efficiency of selfish investments in network security. In *Proc. of the 3rd international workshop on Economics of networked systems*, pages 31–36. ACM, 2008.
13. J. Kesan, R. Majuca, and W. Yurcik. Cyberinsurance as a market-based solution to the problem of cybersecurity: a case study. In *Proc. WEIS*. Citeseer, 2005.
14. H. Kunreuther and G. Heal. Interdependent security. *Journal of Risk and Uncertainty*, 26(2):231–249, 2003.
15. M. Lelarge and J. Bolot. A local mean field analysis of security investments in networks. In *Proc. of the 3rd international workshop on Economics of networked systems*, pages 25–30. ACM, 2008.
16. M. Lelarge and J. Bolot. Network externalities and the deployment of security features and protocols in the internet. In *ACM SIGMETRICS*, 2008.
17. M. Lelarge, J. Bolot. Economic incentives to increase security in the internet: The case for insurance. In *INFOCOM*, pages 1494–1502, 2009.
18. Y. Li, B. Q. Zhao, and J. C. S. Lui. On modeling product advertisement in large scale online social networks. *Accepted for publication, IEEE/ACM Transactions on Networking*, 2011.
19. G. Medvinsky, C. Lai, and B. Neuman. Endorsements, licensing, and insurance for distributed system services. In *Proceedings of the 2nd ACM Conference on Computer and Communications Security*, pages 170–175. ACM, 1994.
20. S. Melnik, A. Hackett, M. Porter, P. Mucha, and J. Gleeson. The unreasonable effectiveness of tree-based theory for networks with clustering. *Physical Review E*, 83(3):036112, 2011.
21. R. Miura-Ko, B. Yolken, N. Bambos, and J. Mitchell. Security investment games of interdependent organizations. In *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, pages 252–260. IEEE, 2008.
22. D. Moore, C. Shannon, et al. Code-red: a case study on the spread and victims of an internet worm. In *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*, pages 273–284. ACM, 2002.
23. M. Newman. *Networks: an introduction*. Oxford Univ Pr, 2010.
24. N. Nisan. *Algorithmic game theory*. Cambridge Univ Pr, 2007.
25. J. Omic, A. Orda, and P. Van Mieghem. Protecting against network infections: A game theoretic perspective. In *INFOCOM 2009, IEEE*.
26. S. Shavell. On moral hazard and insurance. *The Quarterly Journal of Economics*, 93(4):541, 1979.
27. N. Shetty, G. Schwartz, M. Felegyhazi, and J. Walrand. Competitive cyber-insurance and internet security. *Economics of Information Security and Privacy*, pages 229–247, 2010.
28. Z. Yang and J. Lui. Security adoption in heterogeneous networks: the influence of cyber-insurance market. `http://www.cse.cuhk.edu.hk/%7ecslui/TR1.pdf`, 2011.
29. Z. Yang and J. Lui. Investigating the effect of node heterogeneity and network externality on security adoption. In *Thirteenth ACM Sigmetrics Workshop on MAthematical performance Modeling and Analysis (MAMA)*, Jun, 2011.