Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 518213, 11 pages
http://dx.doi.org/10.1155/2013/518213

Research Article

Effective Proactive and Reactive Defense Strategies against Malicious Attacks in a Virtualized Honeynet

Frank Yeong-Sung Lin, Yu-Shun Wang, and Ming-Yang Huang

Department of Information Management, National Taiwan University, Taipei, Taiwan

Correspondence should be addressed to Yu-Shun Wang; yu.shun.wang.tw@gmail.com

Received 11 April 2013; Accepted 22 July 2013

Academic Editor: Anyi Chen

Copyright © 2013 Frank Yeong-Sung Lin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Virtualization plays an important role in the recent trend of cloud computing. It allows the administrator to manage and allocate hardware resources flexibly. However, it also causes some security issues. This is a critical problem for service providers, who simultaneously strive to defend against malicious attackers while providing legitimate users with high quality service. In this paper, the attack-defense scenario is formulated as a mathematical model where the defender applies both proactive and reactive defense mechanisms against attackers with different attack strategies. In order to simulate real-world conditions, the attackers are assumed to have incomplete information and imperfect knowledge of the target network. This raises the difficulty of solving the model greatly, by turning the problem nondeterministic. After examining the experiment results, effective proactive and reactive defense strategies are proposed. This paper finds that a proactive defense strategy is suitable for dealing with aggressive attackers under "winner takes all" circumstances, while a reactive defense strategy works better in defending against less aggressive attackers under "fight to win or die" circumstances.

1. Introduction

The vision for most service providers is to provide high-quality service and improve customer satisfaction, thus maximizing profit. From an infrastructure perspective, the evolution of computing architecture has shifted from mainframe, cluster computing, distributed computing, and grid computing to cloud computing. As a recent and increasingly noteworthy trend in information technology (IT), cloud computing describes a new service model based on the Internet. From a managerial perspective, one of the key success factors of cloud computing is its ability to promise and achieve high quality and availability of service.

Applying virtualization technology makes the job of providing enterprise security more difficult. As observed in the IBM X-Force Mid Year Trend and Risk Report conducted in August 2010 [1], attackers continue to take advantage of security flaws. The rate of vulnerability disclosures in 2010 is higher than any point from 2000 to 2009. The number of virtualization vulnerability disclosures from 1999 through the end of 2009 ascended rapidly, peaking in 2008 at 100. While this number fell by 12 percent to 88 in 2009, this drop indicates that virtualization vendors have recognized the threat these flaws pose and have increased their attention to security. Although the ratio of virtualization vulnerability disclosures increased by only 1 percent from 2007 through 2009, these vulnerabilities still represent a notable security threat. Therefore, it is increasingly important to understand the security implications of virtualization technology.

In addition to external malicious attackers, another weak component of networks is insiders. They insensibly assist people with bad intentions in their legal use of computer systems or networks. Such insider-assisted malicious attacks adopt social engineering as a method to exploit common human behavior. These attacks require a lower degree of technical ability than standard malicious attacks but can cause higher degrees of damage. As a result, organizations emphasize the importance of policy enforcement and increase employee education to mitigate the risks posed by social engineering.

In response to these problems, this paper considers both proactive defense resources, such as firewalls, IDS, IPS, and reactive defense techniques. All types of resources considered

Table 1: Given parameters.

Notation	Description
N	The index set of all nodes
C	The index set of all core nodes
L	The index set of all links
M	The index set of all levels of virtual machine monitors (VMMs)
H	The index set of all types of honeypots
P	The index set of candidate nodes equipped with false target function
Q	The index set of candidate nodes equipped with fake traffic generating function
R	The index set of candidate nodes equipped with false target and fake traffic generating function
S	The index set of all kinds of services
B	The defender's total budget
w	The cost of constructing one intermediate node
o	The cost of constructing one core node
P	The cost of each virtual machine (VM)
k_i	The maximum number of virtual machines on VMM level i, where $i \in M$
α_i	The weight of ith service, where $i \in S$
E	All possible defense configurations, including resources allocation and defending strategies
Z	All possible attack configurations, including attacker attributes, strategies, and transition rules
$\overrightarrow{A_{ij}}$	An attack configuration, including the attributes, strategies, and transition rules of the attacker launches jth attack on ith service, where $i \in S$, $1 \le j \le F_i$
F_i	The total attacking times on ith service for all attackers, where $i \in S$
γ	The cost of constructing a reconfiguration function to one node
λ	The minimum number of hops from core node attackers can start to compromise
c_i	The number of hops from core node category i attackers starts to compromise, where $i \in Z$

Table 2: Decision variables.

Notation	Description
$\overrightarrow{D_i}$	A defense configuration, including resource allocation and defending strategies on ith service, where $i \in S$
$T_{ij}(\overrightarrow{D_i}, \overrightarrow{A_{ij}})$	1 if the attacker achieves his goal successfully and 0 otherwise, where $i \in S$, $1 \le j \le F_i$
n_i	The proactive defense resource allocated to node i, where $i \in N$
l_i	The number of VMM level i purchased, where $i \in M$
δ_i	The number of services that honeypot i can simulate, where $i \in H$
ε_i	The interactive capability of false target honeypot i, where $i \in P$
θ_i	The maximum throughput of fake traffic of honeypot i, where $i \in Q$
$V(l_i)$	The cost of VMM level i with l_i VMMs, where $i \in M$
$h(\delta_i, \varepsilon_i)$	The cost of constructing a false target honeypot, where $i \in P$
$f(\delta_i, \theta_i)$	The cost of constructing a fake traffic generator honeypot, where $i \in Q$
$t(\delta_i, \varepsilon_i, \theta_i)$	The cost of constructing a honeypot equipped with false target and fake traffic functions, where $i \in R$
B_{NL}	The budget of constructing nodes and links
$B_{\text{proactive}}$	The budget of proactive defense resource
B_{special}	The budget of reactive defense resource
$B_{\text{virtualized}}$	The budget of virtualization
B_{honeypot}	The budget of honeypots
B_{reconfig}	The budget of reconfiguration functions
e	The total number of intermediate nodes
q_{ij}	The capacity of direct link between nodes i and j, where $i \in N$, $j \in N$
$g(q_{ij})$	The cost of constructing a link from node i to node j with capacity q_{ij}, where $i \in N$, $j \in N$
x_i	1 if node i is equipped with false target function and 0 otherwise, where $i \in N$
y_i	1 if node i is equipped with fake traffic generating function and 0 otherwise, where $i \in N$
z_i	1 if node i is equipped with reconfiguration function and 0 otherwise, where $i \in N$

in this work one is quantified not only by monetary, but also by time, labor, and other possible factors. As defenders seek out an appropriate defense resource allocation within budget limitations and observe the Quality of Service (QoS) requirement, it becomes increasingly important to best determine how to find a defense mechanism that can detect risky attack behavior early and mislead attackers to routes distant from the server before attackers are at the gates.

In this paper, we assume that organizations may encounter a great diversity of threats. Whether these threats are external or internal, they can bring about vast loss to finances and reputation. However, budgets for security and training are often inadequate. Hence, it is much more important for a system or network to enhance robustness in order to satisfy QoS requirements for service users, than to prevent all categories of malicious attacks. This symbiotic concept to security is called survivability, which is widely defined and applied in previous works [2–6].

Survivability is a typical metric that measures network performance under intentional attacks or other failures. Traditionally, network security status is divided into two discrete types: compromised and safe. Typically, when providing services, the evaluation criterion is the quality of service, but applying this dichotomy of compromise and safety ignores the intermediate status. For instance, while under an attack, the performance level of each service continually declines. Therefore, network status should be represented in a continuous form [2]. Hence, in this paper, survivability is chosen as the metric for describing network status. According to [2], survivability is defined as: "The capability of a system to fulfill its mission, in a timely manner, in the presence of

Table 3: Verbal notations.

Notation	Description
G_{core_i}	Loading of each residual core node i, where $i \in C$
U_{link_i}	Link utilization of each link i, where $i \in L$
K_{effect}	Negative effect caused by applying fake traffic adjustment
I_{effect}	Negative effect caused by applying dynamic topology reconfiguration
J_{effect}	Negative effect caused by applying local defense
O_{tocore}	The number of hops legitimate users experienced from one boundary node to core nodes
Y	The total compromise events
$W_{\text{threshold}}$	The predefined threshold about QoS
W_{final}	The QoS level at the end of attack
$W(\cdot)$	The value of QoS determined by G_{core_i}, U_{link_j}, K_{effect}, I_{effect}, J_{effect}, and O_{tocore}, where $i \in C$, $j \in L$
ρ_{defense_i}	The total defense resource of the shortest path from compromised nodes detected to core node i divided by total defense resource, where $i \in C$
τ_{hops_i}	The number of hops from compromised nodes detected to core node i divided by the number of hops from attacker's starting point, where $i \in C$
ω_{degree_i}	The linking number of core node i divided by the maximum number in the topology, where $i \in C$
$s^i_{\text{priority}_j}$	The priority of service j provided by core node i divided by the maximum service priority in the topology, where $i \in S$, $j \in C$
$\beta_{\text{threshold}}$	The risk threshold of core nodes
$\beta(\cdot)$	The risk status of each core node, which is the aggregation of defense resource, number of hops, link degree, and service priority

Table 4: Environment Parameters.

Testing platform	
Programming language	C
Compiler	GNU GCC 4.6.2
Total evaluation times	60,000
Distributions applied to describe attackers' attributes	Normal distribution

Table 5: Parameters for defender.

Parameters	Value
Topology type	Scale-free network
Number of nodes	49
Number of core nodes (each service)	6 (1, 2, 3)
Number of terminal nodes	5
Number of services	3
Weight of each service	1 : 2 : 3
Number of users	30
Defender total budget	1,700,000
Topology construction budget	700,000
Proactive defense budget	400,000
Reactive defense budget	600,000

Table 6: Parameters for attacker.

Parameters	Value
Total attack budget	A normal distribution with lower bound 300,000 and upper bound 1,500,000
Capability	A normal distribution with lower bound ε and upper bound 1
Aggressiveness	A normal distribution with lower bound 0.1 and upper bound 0.9
Attacker's objective	Service disruption or steal confidential information

attacks, failure, or accidents. . .including networks and large-scale systems of systems."

Along with the concept of survivability, the vulnerability of each node is determined by the Contest Success Function (CSF), which is also applied in [7–10]. CSF originates from the economic rent seeking problem found in Economic Theory. This method also applies a continuous approach to the problem. The form of the CSF is success probability $= T^m/(T^m + t^m)$, when applied to attack and defense scenarios, where T represents the resources invested by the attacker and t stands for resources deployed by the defender. Further, m is known as contest intensity, which illustrates the nature of the battle, while success probability is the probability of a node being compromised. When the value m is between 0 and 1, it represents "fight to win or die" circumstance [11], which means the effectiveness of resources is insignificant. For $m \geq 1$, the effectiveness of resources invested by both sides is exponentially increasing. If m closes to ∞, it stands for a "winner takes all" circumstance; significant advantage is granted to the stronger side, even if that side is stronger by only one invested resource [12].

There are many popular methodologies applied for solving survivability problems. In recent years, game theory is a widely used one. Nevertheless, in [3], the authors point out that this solution approach is limited for deterministic scenarios. Even the emerging branch of game theory, stochastic game, is still confined with this assumption that all values of probabilistic variables have to be determined before the attack and defense starts. This feature has a negative effect on creating cyber attack and defense scenario since, in real world, decisions made during attack and defense depend on current status. Given all values of variables makes the scenario far from reality.

For example, when choosing a victim from candidate nodes, an attacker should apply information like the loading of each node, traffic amount on each link, and/or number of users on each node to evaluate the importance of all candidates. Then, choosing the most appropriate one as the target, yet the restriction of game theory enforces the choosing probability which should be determined at the beginning of the cyber warfare. In other words, those variations happened during attack and defense, like traffic reroute, link status, node conditions, are ignored. Consequently, in this work, Monte Carlo simulation is applied to consider hopefully and cover every angle in the attack and defense scenario.

2. Problem Formulation

2.1. Problem Description. In order to improve system survivability, the defender deploys both proactive and reactive defense resources to confront different attacks. Proactive defense resources are deployed before an attack is launched. In this paper, proactive resources include a firewall system, antivirus software, detection techniques, such as Intrusion Detection System (IDS) and Intrusion Protection System (IPS).

Alternatively, reactive defense resources are activated during an attack as an immediate action for the defender. The mechanisms considered in this paper can strengthen defenses, provide deception, and provide resource concentration.

For strengthening defenses, since the scenario considered in this paper is constructed in a virtualized environment, a number of Virtual Machines (VMs) are governed by a Virtual Machine Monitor (VMM) that controls all information details for all VMs. When an attack event is detected, local defense functions for each VMM are activated automatically to raise the defense capability of the virtualized nodes belonging to the same VMM. However, this mechanism does not always create positive effects. Once an attacker determines that the target network is a virtualized environment and discovers the existence of a VMM, he can compromise the VMM through vulnerabilities in APIs [13]. If the VMM is compromised, all VMs belonging to this VMM are also compromised.

Deception mechanisms are widely applied in defense, and in this paper honeypots are considered to distract attackers. According to [14–17], honeypots not only serve as a passive decoy fooling attackers into believing they have achieved their goal and preemptively terminating the attack but also as an active lure that acts as a service-providing node to attract attackers. The former is known as a false target, and the latter can be implemented by a fake core node that spreads service-like traffic to attract attackers. When facing different attackers, the deceiving probability of each honeypot is different. This probability is jointly determined by attackers' capability and the interaction level of a honeypot.

As for resource concentration, by modifying the concept of "rotation" discussed in [18–20] and adapting it to our scenario, the defender can adopt dynamic responsive strategies to improve system survivability. Hence, while under an attack, the defender can apply dynamic topology reconfiguration to exchange the neighbor of one core node, which has the strongest defensive resources, with a node that is close to the attacker's current location.

However, since dynamic topology reconfiguration requires node rotation, it negatively impacts QoS. Therefore, unless the risk level of a core node exceeds a predefined threshold, the defender will not activate this mechanism. Furthermore, false positive and false negative situations for every defense mechanism are considered.

In addition to the defense mechanisms described above, the following attributes are also considered for creating a realistic attack-defense scenario.

2.1.1. Goal. Generally, attackers target a network to either disrupt services or to compromise servers and steal sensitive information. Therefore in this paper an attacker's goal may be service disruption or the theft of confidential information. Service disruption is achieved by ensuring that the minimal level of service quality is not fulfilled. In the case of information theft, attackers usually establish their target before launching an attack, and once core nodes with the desired information are compromised, the defenders lose.

2.1.2. Budget. Budget stands for the primary resources for an attack, including money, manpower, computing effort, time, and other important factors. Without sacrificing generality, an attackers' budget follows a general distribution. When determining the result of a compromising event, for example, one attacker invests a certain amount of attack resources on compromising the target node, and the CSF is applied to decide the compromised probability. In contrast with [21] and [22], if an attacker invests more resources than the defender on a given target node, it is not guaranteed that the attacker wins; it only raises the compromised probability.

2.1.3. Capability. This criterion depicts an attackers' proficiency and is also described by a general distribution. For highly proficient attackers, there is a high probability that they will see through reactive defense mechanisms, such as honeypots.

2.1.4. Attack Type. In general, a malicious attacker launches an attack from one of the boundary nodes, which is commonly considered an external attack. In contrast, other attacks can be launched by malicious insiders and cause more severe consequences [13]. Malicious insiders are able to choose an internal node as their starting position for compromising the network.

In the real world, external attackers may apply social engineering to escalate their access privileges. This mechanism allows the attacker to bypass some proactive defense facilitates, like a firewall. As a result, attackers have an edge on compromising the network.

In order to best mimic real-world conditions, all attack types discussed above are considered in this paper.

2.1.5. Aggressiveness. This metric describes the preferred compromised probability of an attacker when attacking a target node. This attribute is highly dependent on budget, since the compromised probability is the left-hand side of the CSF. In other words, the attack resources required for compromising a target node are calculated by the given defense resources, contest intensity (m), and attacker's aggressiveness. Highly aggressive attackers prefer a high success probability and like to spend a larger amount of resources to compromise the target node than less aggressive attackers. An attacker's aggressiveness is determined by a general distribution.

2.1.6. Next Hop Selection Criterion. In this paper, the attackers are assumed to have incomplete information and imperfect knowledge. Since they can only gather local information,

(1) Attackers only have incomplete information.
(2) The defender only has incomplete information regarding the network since there are unaware vulnerabilities.
(3) A service is provided by multiple core nodes.
(4) Each service has different weights.
(5) One virtual machine only provides one service.
(6) Only malicious nodal attacks are considered.
(7) The compromise probability of a target node is determined by the Contest Success Function (CSF).

Box 1: Problem assumptions.

such as the defense level of one hop neighbors, link traffic, or link utilization, attackers must use available information wisely to choose a proper strategy to select the next victim. Inspired by the seminal work of [23], possible attack strategies may be developed based on the quantity of proactive defense resources, link utilization, or a portion of the target service traffic of each candidate node. Additionally, this paper also considers a more irrational strategy of a blind attack.

By applying general distributions to describe attacker-related attributes, the total number of attacker categories is nearly infinite. This feature increases the generalizability of our model. The detailed assumptions are described in Box 1.

To describe the attack procedures in more detail, we develop the following idea to explain the interaction between the defender and attackers. The following figures are drawn from the attacker's view for clarity. In other words, all figures are a logical topology that has been already virtualized, since one physical machine may represent many VMs. The explanations of components are listed in Figure 1.

First, the defender deploys proactive defense resource on each node, and the attacker starts to compromise the network from the edge nodes (S). Before launching an attack, the attacker probes all the candidate nodes to gather sufficient information and determine the next hop selection criterion for this compromise event. By applying the next hop selection criterion, the victim is chosen. The result of this compromise event is determined by the CSF. If the target is compromised, it becomes a spring board for attackers to assault other uncompromised neighbors. Corresponding information is shown in Figure 2. The topology demonstrated in Figures 2 and 3 for presentation purposes; the topology structure implemented in the simulations is a scale-free network.

The risk level of each core node is evaluated by minimum defense resources, the number of hops, the link degree, and service priority. Minimum defense resource stands for the shortest path from compromised nodes to one core node, divided by total defense resources. The number of hops is determined by the minimum number of hops from compromised nodes to one core node, divided by the maximum number of hops from the attackers' starting point to one core node. Link degree is the linking number of each core node divided by the maximum linking number in the topology. Service priority is the weight of the service that is provided by the target core node divided by the highest weight of service in the topology. If any of the core nodes is in danger, the defender can activate defense mechanisms, such as a fake traffic generator and a dynamic topology reconfiguration under QoS limitations.

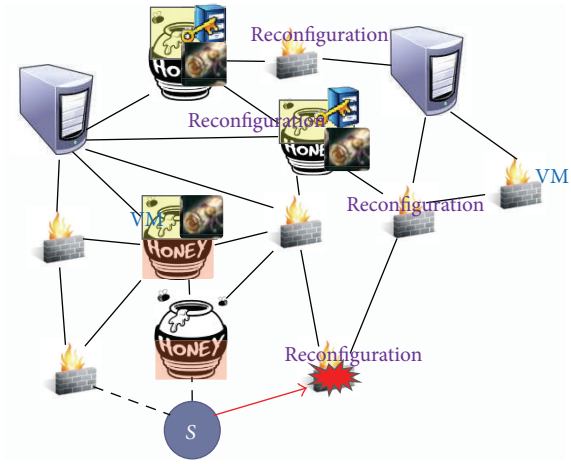

Figure 1: Explanations of Components.

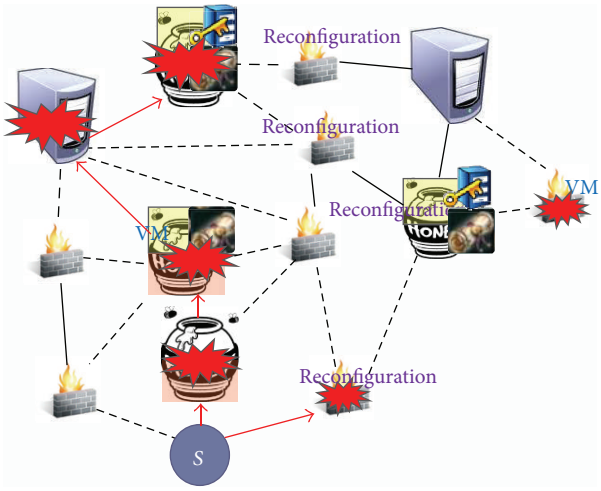

Figure 2: Sample network and resource allocation scheme.

When the defender activates a dynamic topology reconfiguration, only nodes implemented with the reconfiguration function are considered. First, the defender chooses the least proactively defended neighbor with the reconfiguration function of a risky core node. The defender then selects another node that is both not a neighbor of the core node and the most proactively defended nearby the node selected from the previous step. If there is a fake traffic generating honeypot, the defender is also able to activate the mechanism for influencing an attacker's next hop selecting criterion.

In the event that attackers attack the virtualized nodes and this malicious event is detected by the defender, the local

Given
(1) all possible defense configurations set, including defense resource allocation and defending strategies,
(2) all possible attack configurations set, including attacker attributes, strategies, and selection criterion,
(3) the total attack times on each service.
Objective
to minimize the service compromised probability of the target network.
Subject to
(1) budget constrain for both the defender and attackers,
(2) the minimum QoS requirement for legitimate users.
To determine
the effective defense strategies to allocate resources.

Box 2: Problem description.

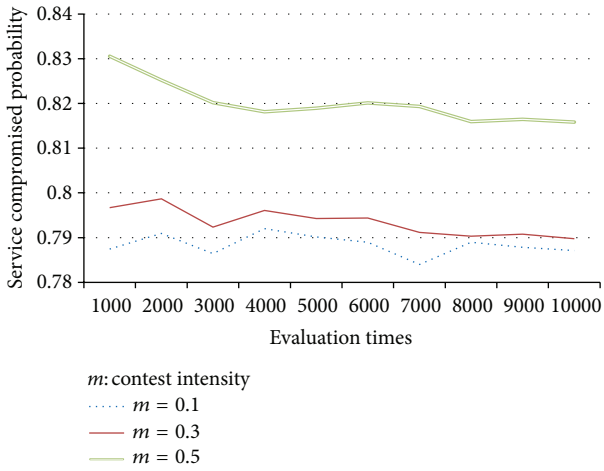

Figure 3: A possible result of an attack and defense scenario.

defense mechanism is activated automatically. Consequently, the defense level increases for the attacked node, the VMM, and the rest of this virtualized group. If attackers see that the node is under a virtualized environment, they may continue to compromise the VMM.

When a core node is compromised, the defender must evaluate whether the performance level still fulfills the minimum QoS requirement and update the degree of risk regarding each core node. If an attacker targets multiple services, once they compromise a false target honeypot and are deceived by it, they will change their target to the next service. Finally, if attackers exhaust their budget, the attack is terminated and the defender wins the battle. One possible result is shown in Figure 3. A structural description of the problem is presented in Box 2.

2.2. Mathematical Formulation. The scenario discussed previously is modeled as a minimization problem, given the parameters and decision variables shown in Tables 1 and 2, respectively. Furthermore, since the high amount of randomness involved renders the nature of this problem nondeterministic, it is quite difficult to formulate this problem purely through mathematics. To explain the nondeterministic nature, the following example is used: considering two adversaries with the same attack strategy, since the problem is nondeterministic, the consequences can be totally different. One may be distracted by a honeypot and the other may successfully achieve the goal. This feature dramatically expands the width in survivability studies. Further, the scenario considered in this work gives the defender the average survivability analysis of a network.

There are three major reasons that an average survivability analysis is more meaningful for the defender. First, in the real world, there are only few adversaries holding "complete" information regarding the target network. This kind of "worst case" rarely happen. Secondly, analyzing the worst case though gives a defender the lower bound of network survivability. Nevertheless, it overestimates the budget required for defending a network. An average survivability analysis makes conclusions closer to reality. Last but not the least, in average case analysis, if the defender wants to evaluate the survivability when facing a stronger adversary, it can be achieved by simply tuning the parameter. Average case analysis is flexible and adjustable according to the demand of a defender.

Thus, to well describe the problem, some verbal notations and constraints are included, which are shown in Table 3.

Objective function:

$$\min_{\overrightarrow{D_i}} \frac{\sum_{i\in S}\left[\alpha_i \times \sum_{j=1}^{F_i} T_{ij}\left(\overrightarrow{D_i}, \overrightarrow{A_{ij}}\right)\right]}{\sum_{i\in S}\left(\alpha_i \times F_i\right)}, \quad \text{(IP 1)}$$

subject to
mathematical constraints:

$$\overrightarrow{D_i} \in E \quad \forall i \in S, \quad \text{(IP 1.1)}$$

$$\overrightarrow{A_{ij}} \in Z \quad \forall i \in S,\ 1 \le j \le F_i, \quad \text{(IP 1.2)}$$

$$q_{ij} \ge 0 \quad \forall i, j \in N, \quad \text{(IP 1.3)}$$

$$x_i + y_j \ge 1 \quad \forall i, j \in H, \quad \text{(IP 1.4)}$$

$$c_i \ge \lambda \quad \forall i \in Z, \quad \text{(IP 1.5)}$$

$$B_{\text{NL}} + B_{\text{proactive}} + B_{\text{reactive}} \le B, \quad \text{(IP 1.6)}$$

$$B_{\text{virtualized}} + B_{\text{honeypot}} + B_{\text{reconfig}} \le B_{\text{reactive}}, \quad \text{(IP 1.7)}$$

$$w \times e + o \times \|C\| + \frac{\sum_{i \in N} \sum_{j \in N} g\left(q_{ij}\right)}{2} \leq B_{\mathrm{NL}}, \quad \text{(IP 1.8)}$$

$$\sum_{i \in N} n_i \leq B_{\mathrm{proactive}}, \quad \text{(IP 1.9)}$$

$$\sum_{i \in M} v\left(l_i\right) + p \times \sum_{i \in M} l_i \times k_i \leq B_{\mathrm{virtualized}}, \quad \text{(IP 1.10)}$$

$$\sum_{i \in P} x_i \times h\left(\delta_i, \varepsilon_i\right) + \sum_{j \in Q} y_j \times f\left(\delta_j, \theta_j\right) + \sum_{i \in N} \sum_{j \in N} x_i \times y_j \times t\left(\delta_i, \varepsilon_i, \theta_j\right) \leq B_{\mathrm{honeypot}}, \quad \text{(IP 1.11)}$$

$$\sum_{i \in N} z_i \times r \leq B_{\mathrm{reconfig}}, \quad \text{(IP 1.12)}$$

$$w \times e \geq 0, \quad \text{(IP 1.13)}$$

$$g\left(q_{ij}\right) \geq 0 \quad \forall i, j \in N, \quad \text{(IP 1.14)}$$

$$n_i \geq 0 \quad \forall i \in N, \quad \text{(IP 1.15)}$$

$$v\left(l_i\right) \geq 0 \quad \forall i \in M, \quad \text{(IP 1.16)}$$

$$h\left(\delta_i, \varepsilon_i\right) \geq 0 \quad \forall i \in P, \quad \text{(IP 1.17)}$$

$$f\left(\delta_j, \theta_j\right) \geq 0 \quad \forall j \in Q, \quad \text{(IP 1.18)}$$

$$t\left(\delta_i, \varepsilon_i, \theta_j\right) \geq 0 \quad \forall i \in N, \quad \text{(IP 1.19)}$$

$$x_i = 0 \text{ or } 1 \quad \forall i \in N, \quad \text{(IP 1.20)}$$

$$y_i = 0 \text{ or } 1 \quad \forall i \in N, \quad \text{(IP 1.21)}$$

$$z_i = 0 \text{ or } 1 \quad \forall i \in N, \quad \text{(IP 1.22)}$$

verbal constraints:

$$\frac{\int_{y=1}^{Y} \left[W\left(G_{\mathrm{core}_i}, U_{\mathrm{link}_j}, K_{\mathrm{effect}}, I_{\mathrm{effect}}, J_{\mathrm{effect}}, O_{\mathrm{tocore}}\right)\right] dy}{Y} \geq W_{\mathrm{threshold}}, \quad \text{where } i \in C,\ j \in L, \quad \text{(IP 1.23)}$$

$$W_{\mathrm{final}} \geq W_{\mathrm{threshold}}. \quad \text{(IP 1.24)}$$

For each core node, when

$$\beta\left(\rho_{\mathrm{defense}}, \tau_{\mathrm{hops}}, \omega_{\mathrm{degree}}, s_{\mathrm{priority}_i}\right) \geq \beta_{\mathrm{threshold}}, \quad \text{where } i \in S, \quad \text{(IP 1.25)}$$

the defender can activate the reactive defense mechanisms.

The goal of the objective function is to minimize the compromised service probability, which is modeled as the weighted total of successful attacks, divided by the total weighted number of attacks. The result of an attack is determined by $T_{ij}(\overrightarrow{D_i}, \overrightarrow{A_{ij}})$. For each service, the defender constructs a defense configuration, including defense resource allocation and defending strategies to oppose attackers targeting the service with different attack configurations, including attacker attributes, strategies and transition rules. Not only malicious attacks but also QoS issues are taken into consideration.

Equation (IP 1.1) stands for the feasibility of the defense configuration of each service. For the attacking side, (IP 1.2) denotes that the attack configuration should be feasible. Equation (IP 1.6) denotes that the summation of defense resources spent should not exceed total budget B. Equations (IP 1.7)~(IP 1.19) jointly restrain the budget of each type of defense resource individually. Equations (IP 1.20)~(IP 1.22) impose binary restrictions on decision variables. Finally, (IP 1.23)~(IP 1.25) jointly represent that the performance reduction caused by either malicious attacks or activating reactive defense mechanisms should not violate the QoS requirement.

3. Numerical Analysis

In this paper, a scale-free network is constructed for evaluation, since this structure is similar to real-world forms, such as the Internet. The implementation algorithm is referenced from [24].

3.1. Simulation Environment. Table 4 presents the system experiment parameters. Defender-related parameters are illustrated in Table 5. The unit of budget is dollar. For attackers, the parameters are shown in Table 6.

The value of each attacker's budget, capability and aggressiveness is governed by a normal distribution with different lower and upper bounds.

3.2. Numerical Result. As mentioned previously, the Contest Success Function is applied for quantifying vulnerability. The value of contest intensity is classified into two groups: scores greater than 1 and scores smaller than 1.

3.2.1. Convergence. The first issue to address before constructing meaningful simulations is convergence. In this paper, the convergence of data is considered as numerical stability. While the magnitude of data vibrations is within the acceptable interval, for example, 0.2%, the corresponding number of simulation times (M) is set as the number of evaluations for each attack and defense scenario.

For rigorousness, in convergence experiments, the effect of contest intensity is jointly considered. Three different magnitudes are simulated on a 49 node scale-free network. For the following experiments in this subsection, the horizontal axis represents the evaluation of the number of attacks, and the vertical axis stands for the network system compromised probability, which is the objective function of the proposed mathematical model.

Figures 4 and 5 show the result of 10,000 simulations with different contest intensity. Here, it is clear that the value of service compromised probability is unstable. In Figures 6 and 7, the vibration becomes alleviative but is still not convergent. When the number of simulations is raised to 100,000, as

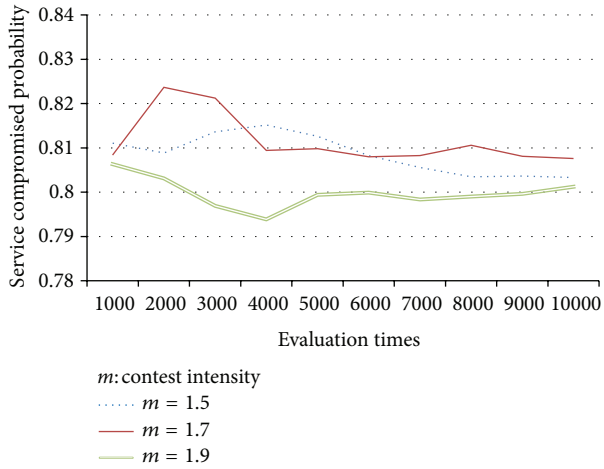

Figure 4: Convergence test on $m < 1$ group for 10,000 simulations.

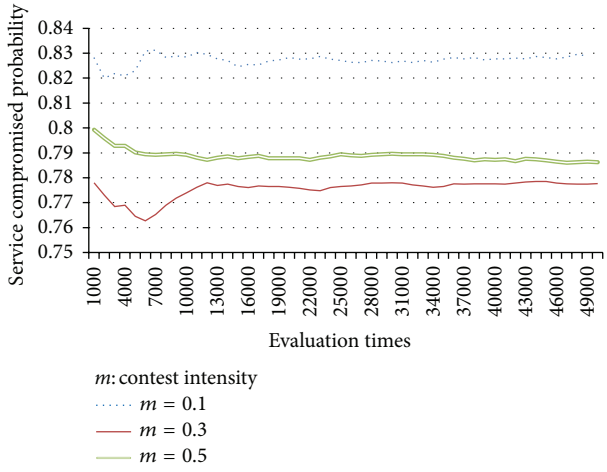

Figure 5: Convergence test on $m > 1$ group for 10,000 simulations.

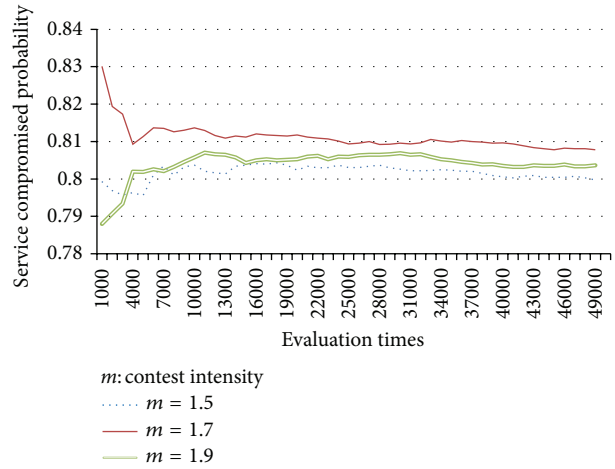

Figure 6: Convergence test on $m < 1$ group for 50,000 simulations.

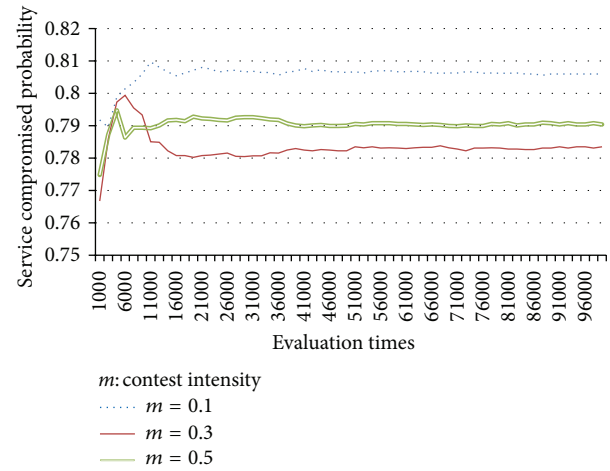

Figure 7: Convergence test on $m > 1$ group for 50,000 simulations.

shown in Figures 8 and 9, there is a stable trend after 60,000 among all different values of contest intensity.

From Figures 4 to 9, it is clear that the 60,000 simulations are a large enough number to give converging results among all values of contest intensity. Therefore, M is set to be 60,000.

3.2.2. Analysis of Key Factors Influencing Service Compromised Probability. Contest intensity has great influence on the nature of network attack and defense. However, as shown in Figure 10, there is no consistent tendency between contest intensity and service compromised probability. The same conclusions also appear in [8–10].

If the effects of contest intensity and attacker's aggressiveness are jointly considered, there are some meaningful and explainable results.

In Figure 11, three value intervals of aggressiveness including 0.1 to 0.9 (average), 0.1 to 0.5 (less aggressive), and 0.5 to 0.9 (aggressive) are simulated among different values of contest intensity. It is clear that when both the effect of contest intensity and an attack's aggressiveness are considered, there are consistent trends. For aggressive attackers, it is more advantageous to have higher values of contest intensity. Alternatively, less aggressive attackers have leverage when the value of contest intensity is small. However, for average attacks there is no clear tendency.

The value of contest intensity is separated into high-value and low-value groups. Here Figure 12 illustrates the variation of service compromised probability in low-level contest intensity groups when facing less aggressive attackers. While in Figure 12 the objective function value presents a linear decreasing form, Figure 13 demonstrates an exponentially decreasing from.

Figures 14 and 15 also present a similar phenomenon; when facing aggressive attackers, the variation of service compromised probability shows an exponentially increasing trend for contest intensity belonging to the low-level group. For the high-level group the tendency is more linear.

According to the previous results, the key factors influencing service compromised probability include contest intensity and attack aggressiveness. These results are summarized in Figure 16.

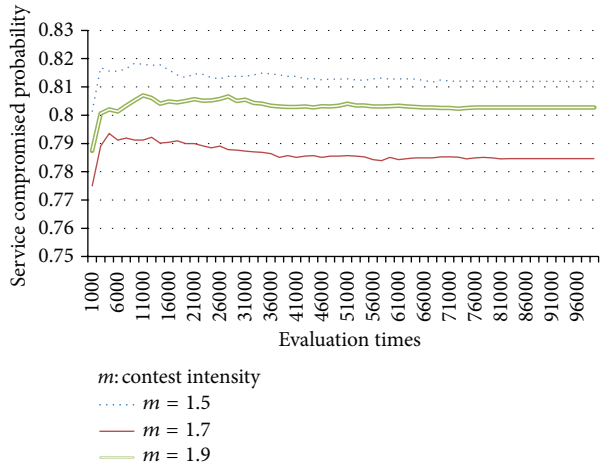

Figure 8: Convergence test on $m < 1$ group for 100,000 simulations.

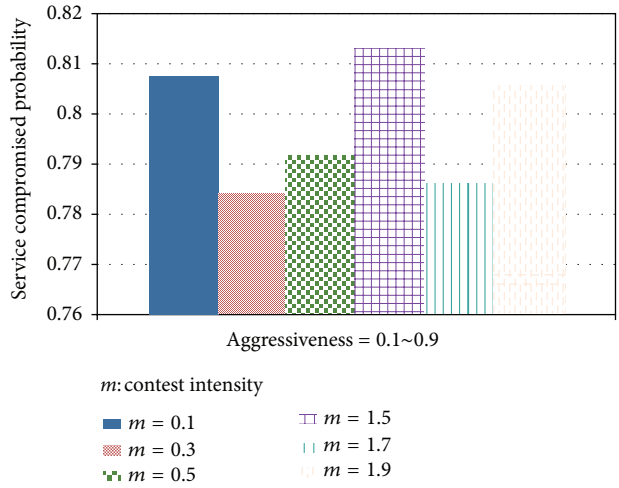

Figure 9: Convergence test on $m > 1$ group for 100,000 simulations.

4. Discussion of Results

Based on the simulation results, some interesting and meaningful arguments are presented in this section.

4.1. The Effectiveness of Resources Invested by the Defender and Attackers Is Highly Dependent on the Nature of Battle. The objective function value shown in Figure 12 presents a linear decreasing form when the contest intensity belongs to a low-level group and the defender is dealing with less aggressive attackers. The reason is that the effectiveness of resources invested by both players is insignificant. Thus, the service compromised probability is less sensitive with contest intensity under "fight to win or die" circumstances.

Furthermore, less aggressive attackers only invest few resources into compromising a target. With the same total budget, they are capable of launching more attacks than aggressive attackers. Therefore, the service compromised probability of less aggressive attackers is much higher than more aggressive attackers.

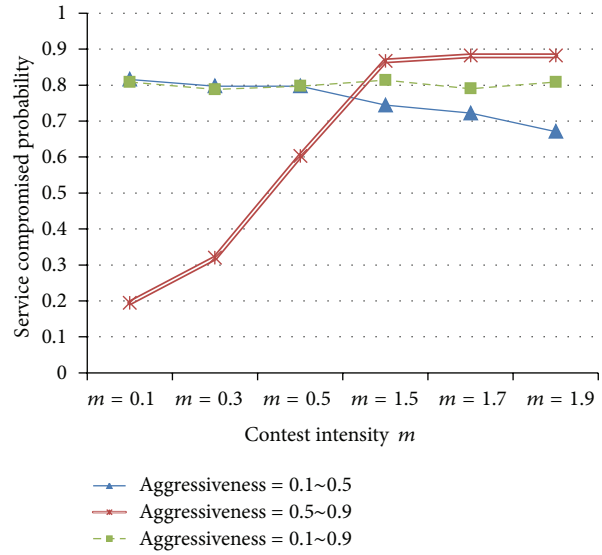

Figure 10: Service compromised probability under different contest intensity.

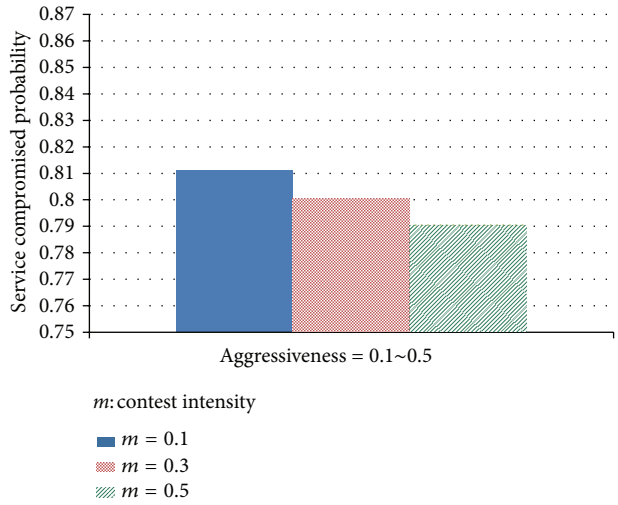

Figure 11: Effect of Aggressiveness among Different Contest Intensity.

In the case of "winner takes all," Figure 13 illustrates an exponentially decreasing trend among service compromised probability; this is because the effectiveness of resources invested is significant to the result of a battle. With a larger value of contest intensity, the vantage of a defender against less aggressive attackers is more obvious.

On the contrary, for aggressive attackers, there is an exponentially increasing tendency under "fight to win or die" circumstances. Since the contest intensity serves as the exponent of the Contest Success Function, when the value increases the effectiveness of resources invested grows exponentially. This is further shown in Figure 14.

Nevertheless, the exponential tendency does not appear through all values of contest intensity. For the "winner takes all" circumstance, in Figure 15, the increasing rate of service

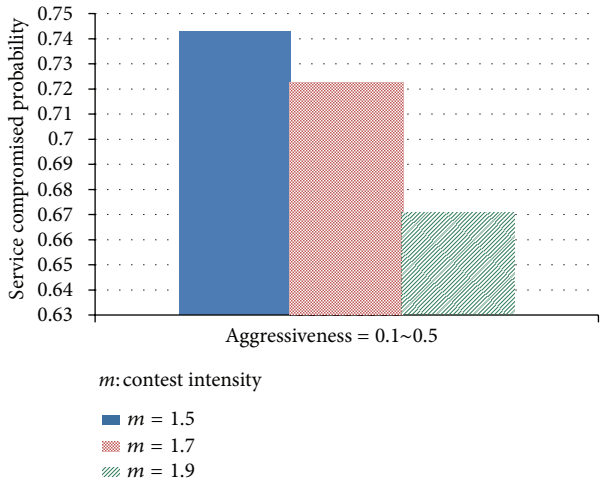

Figure 12: Service compromised probability of less aggressive attacker under $m < 1$ circumstance.

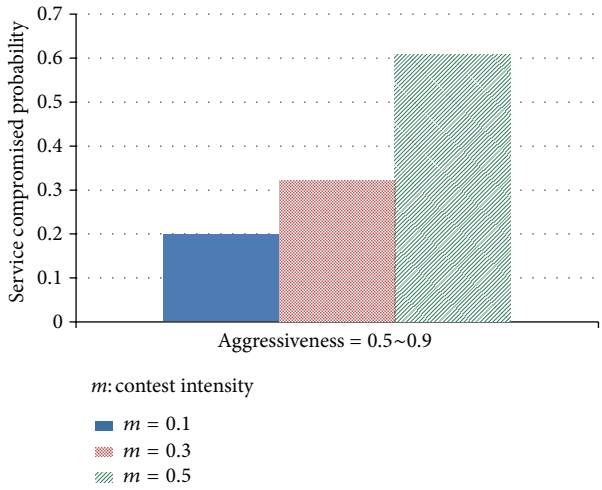

Figure 13: Service compromised probability of less aggressive attacker under $m > 1$ circumstance.

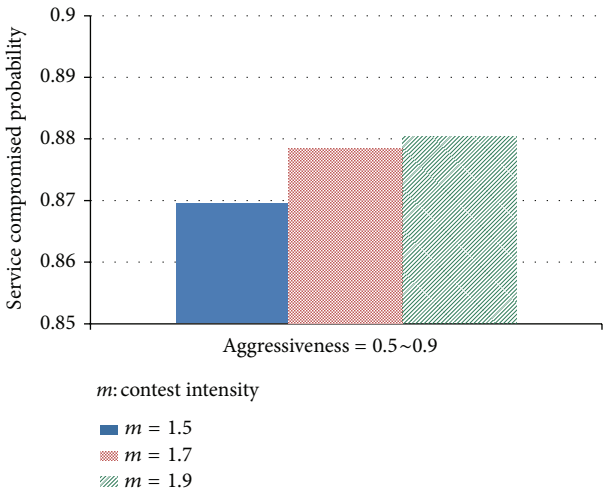

Figure 14: Service compromised probability of aggressive attacker under $m < 1$ circumstance.

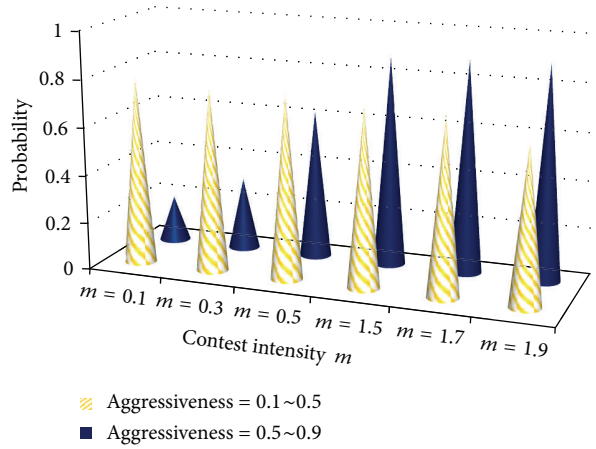

Figure 15: Service compromised probability of aggressive attacker under $m > 1$ circumstance.

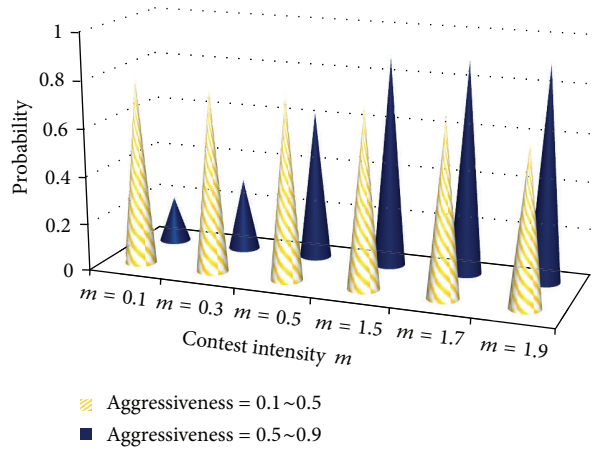

Figure 16: Comparison of service compromised probability under diverse contest intensity and aggressiveness.

compromised probability slows down when contest intensity becomes higher. This is because although aggressive attackers prefer to gain an edge by spending more attack resources to compromise a target, each compromise costs significant resources. Thus, aggressive attackers run out of budget very quickly, and the increasing trend is not exponential.

4.2. Proactive Defense Is Advantageous under "Winner Takes All" Circumstance. According to previous discussions, the effectiveness of resources invested is significant under "winner takes all" circumstances. For less aggressive attackers, the performance is poor since they only invest few resources on an attack. Aggressive attackers tend to spend a larger quantity to compromise a target. If a defender raises the quantity of defense resources on important nodes, it effectively weakens attackers.

Consequently, regardless of whether defenders face either aggressive or less aggressive attackers, a proactive defense performs better in deterring attackers.

4.3. Reactive Defense Is Advantageous under "Fight to Win or Die" Circumstance. Similarly, under the "fight to win

or die" circumstance the effectiveness of defense resources is insignificant. No matter how many proactive defense resources are invested, there is only limited influence on the objective function value. In other words, less aggressive attackers can compromise a target even by only investing scarce attack resources. For aggressive attackers, since they prefer spending a large quantity of resources on compromising a victim, they easily run out of resources before a target service is compromised.

Based on this analysis, it is advantageous for the defender to apply reactive defense mechanisms against malicious attackers under "fight to win or die" circumstance.

5. Conclusions

This paper models an attack and defense scenario that involves high amounts of randomness as mathematical formulations where attackers are assumed to have incomplete information. It further considers a virtualization environment for helping relate these results to recent trends in cloud computing.

According to the simulation results, the outcome of a contest is influenced not only by the quantity of defense resources invested on each node but also by the contest intensity. An attacker's aggressiveness is introduced as a new dimension, and meaningful results are discovered. Effective defense strategies are proposed in the discussion of the results.

For future works, collaborative attacks should be taken into consideration since the attack strategy discussed in this paper is for individual attacks. In other words, there is only one attack targeting a victim's network at a time. Thus, no synergy is considered. A more complicated collaborative attack pattern is worth further study.

Acknowledgment

This work was supported by the National Science Council, Taiwan (Grant nos. NSC 102-2221-E-002-104 and NSC 101-2218-E-011-009).

References

[1] IBM Internet Security Systems X-Force research and development team, "IBM X-Force 2010 Mid-Year Trend and Risk Report," *IBM*, August 2010.

[2] R. J. Ellison, D. A. Fisher, R. C. Linger, H. F. Lipson, T. Longstaff, and N. R. Mead, "Survivable network systems: an emerging discipline," Tech. Rep. CMU/SEI-97-TR-013, 1997.

[3] S. Roy, C. Ellis, S. Shiva, D. Dasgupta, V. Shandilya, and Q. Wu, "A survey of game theory as applied to network security," in *Proceedings of the 43rd Annual Hawaii International Conference on System Sciences (HICSS '10)*, January 2010.

[4] M. N. Lima, A. L. D. Santos, and G. Pujolle, "A survey of survivability in mobile Ad hoc Networks," *IEEE Communications Surveys and Tutorials*, vol. 11, no. 1, pp. 66–77, 2009.

[5] Z. Ma, "Towards a unified definition for reliability, survivability and resilience (I): the conceptual framework inspired by the handicap principle and ecological stability," in *Proceedings of the IEEE Aerospace Conference*, pp. 1–12, March 2010.

[6] F. Xing and W. Wang, "On the survivability of wireless ad HOC networks with node misbehaviors and failures," *IEEE Transactions on Dependable and Secure Computing*, vol. 7, no. 3, pp. 284–299, 2010.

[7] S. Skaperdas, "Contest success functions," *Economic Theory*, vol. 7, no. 2, pp. 283–290, 1996.

[8] G. Levitin and K. Hausken, "False targets efficiency in defense strategy," *European Journal of Operational Research*, vol. 194, no. 1, pp. 155–162, 2009.

[9] K. Hausken and G. Levitin, "Protection vs. false targets in series systems," *Reliability Engineering and System Safety*, vol. 94, no. 5, pp. 973–981, 2009.

[10] G. Levitin and K. Hausken, "Preventive strike vs. false targets and protection in defense strategy," *Reliability Engineering and System Safety*, vol. 96, no. 8, pp. 912–924, 2011.

[11] J. Hirshleifer, "Conflict and rent-seeking success functions: ratio vs. difference models of relative success," *Public Choice*, vol. 63, no. 2, pp. 101–112, 1989.

[12] J. Hirshleifer, "The paradox of power," *Economics and Politics*, vol. 3, pp. 177–200, 1993.

[13] J. Archer, A. Boehme, D. Cullinane, P. Kurtz, N. Puhlmann, and J. Reavis, "Top Threats to Cloud Computing V 1.0," *Cloud Security Alliance*, March 2010.

[14] H. Debar, F. Pouget, and M. Dacier, "White paper: 'Honeypot, Honeynet, Honeytoken: Terminological issues'," Institut Eurécom Research Report RR-03-081, 2003.

[15] B. Cheswick, "An evening with berferd in which a cracker is lured, endured, and studied," in *Proceedings of the USENIX Conference*, pp. 163–174, USENIX, 1992.

[16] C. Seifert, I. Welch, and P. Komisarczuk, "Taxonomy of honeypots," Tech. Rep. CS-TR-06/12, 2006.

[17] M. H. y López and C. F. L. Reséndez, "Honeypots: basic concepts, classification and educational use as resources in information security education and courses," in *Proceedings of the Informing Science and IT Education Conference*, 2008.

[18] Y. Huang, D. Arsenault, and A. Sood, "Closing cluster attack windows through server redundancy and rotations," in *Proceedings of the 6th IEEE International Symposium on Cluster Computing and the Grid (CCGRID '06)*, May 2006.

[19] Y. Huang, D. Arsenault, and A. Sood, "Incorruptible self-cleansing intrusion tolerance and its application to DNS security," *Journal of Networks*, vol. 1, no. 5, pp. 21–30, 2006.

[20] M. Smith, C. Schridde, and B. Freisleben, "Securing stateful grid servers through virtual server rotation," in *Proceedings of the 17th International Symposium on High Performance Distributed Computing (HPDC '08)*, pp. 11–22, June 2008.

[21] F. Y.-S. Lin, Y.-S. Wang, and P.-H. Tsang, "Efficient defense strategies to minimize attackers' success probabilities in honeynet," in *Proceedings of the 6th International Conference on Information Assurance and Security (IAS '10)*, pp. 80–85, August 2010.

[22] F. Y.-S. Lin, Y.-S. Wang, P.-H. Tsang, and J.-P. Lo, "Redundancy and defense resource allocation algorithms to assure service continuity against natural disasters and intelligent attacks," in *Proceedings of the 5th International Conference on Broadband Wireless Computing, Communication and Applications (BWCCA '10)*, pp. 206–213, November 2010.

[23] F. Cohen, "Managing network security: attack and defence strategies," *Network Security*, vol. 1999, no. 7, pp. 7–11, 1999.

[24] S. Nagaraja and R. Anderson, "Dynamic topologies for robust scale-free networks," *Bio-Inspired Computing and Communication*, vol. 5151, pp. 411–426, 2008.