An Evaluation of Network Survivability under the Effect of Accumulated Experience from Sophisticated Attackers

Pei-Yu Chen[1,2,*] and Frank Yeong-Sung Lin[1]

[1] Department of Information Management, National Taiwan University, Taipei, Taiwan, R.O.C
[2] Institute for Information Industry, Taipei, Taiwan, R.O.C.
{d96006,yslin}@im.ntu.edu.tw

Abstract. This paper is focused on the resource allocation of network attack and defense with mathematical programming and to optimize the problem. It adopts a concept, discount coupon, to describe the attack behavior of taking advantage of accumulated experience from his previous attack actions of minimizing future attack cost. The attacker obtains free experience before he launch an attack or from a compromised node which could further reduce the cost of an attack. The attacker's objective is to minimize the total attack cost, while the core node is compromised and the network could not survive. Here, by transforming with node splitting into a generalized shortest path problem and applying the algorithm to optimally solve it.

Keywords: Internet Security, Attack Behavior, Accumulated Experience, Network Survivability, Resource Allocation, Node Splitting, Generalized Shortest Path Problem, Optimization.

1 Introduction

A number of recent high profile attacks such as the Google Aurora, Stuxnet, RSA hack and attacks have been traced on the Internet [1]. Stuxnet is clearly an example of a stealthy virus developed by an adversary that spent a great deal of time and money on research and development. Consindering these characteristics, it is targeted and persistent leading to leakage of vital state secrets and critical damage, which have been labeled as Advanced Persistent Threats (APT). APTs is with high dependent about the knowledge, information, and experience of target network. The adversary picked a specific victim target and make good use of information system vulnerabilities in advance. They act opportunistically, looking for poorly protected targets and therefore easy to breach. A sophisticated attacker only spends a little effort or it can be even view as automatic to accomplish information collecting or experience accumulating when facing a taregeted network. In [2], the authors conduct a small scale study of Internet attack processes using several fixed honeypot in a ten-month period. It was found that the increasing numbers of attacks on non-stop basis compromised

* Corresponding author.

James J. (Jong Hyuk) Park et al. (eds.), *Future Information Technology*, Lecture Notes in Electrical Engineering 309,
DOI: 10.1007/978-3-642-55038-6_106,

personal computers running automated robots to collect information. Moreover, the ports of one or more target systems were probe following a given sequence [3]. In the majority of prior theoretical work, behavior of attackers is modeled as exogenous and its guiding principles remain unclear. Prior research offers examples of quantitative modeling used in evaluating attackers' behavior [4] and of explicit models of that behavior [5,6].

The traditional evaluation of systems or infrastructures was analyzed in terms of a binary state: safe or compromised [7]. It is insufficient to describe the ability of recovering or continual service providing. The definition of network survivability has been discussed for many years. Among these various definitions of network survivability, Ellison and et al. [8] proposed that "capability of a system to fulfill its mission, in a timely manner, in the presence of attacks, failures, or accidents." To the best of our knowledge, there is little research about network attack-defense with mixed nonlinear integer programming to optimize a resource allocation problem. The attack behavior about taking advantage of accumulated experience to minimize the total attack cost is considering in this model. We proposed a mathematical formulation to describe the effect of accumulated experience of attacker. And the problem is transformed with the mathematical technique of node splitting and solved Generalized-Reverse-Dijkstra algorithm to solve [9].

2 Problem Description and Formulation

This research considers network survivability in terms of protection of the "core node" in which organizations store their most significant data or run the most critical service to their business. The main purpose of investing in security is to defend against attacks, acquiring a proper understanding of attackers' behavior is an important step toward better security practices. In [5], they pointed out that the quantitative modeling of survivability for validation or measurement should be based on detailed intruder models. Since intruders are humans or are controlled by humans, they not only learn but accumulate experience. Thus the unlikely intruder of today becomes the most probable intruder of tomorrow. These evidences show that the attacker may only collect so useful information form nods to reduce his future attack costs.

Because of the valuability of this node, attackers do their best to compromise it. As attack resources are limited, the attacker needs guidelines about how to make use of the experience and his budget to compromise a node is addressed through this problem. In this paper, the attacker somehow gain free experience before he launch an attack or from a compromised node which could further reduce the cost of an attack. The objective is to minimize the total attack cost from an attacker's perspective, while the core node is compromised and the network could not survive. Thus, the minimum attack cost could be also viewed as the evaluation of the robustness of a network under intentional or malicious attack from defenders' perspective.

2.1 Problem Description

Given that both the attacker and the defender have information about the targeted network topology. Meanwhile, the attacker has complete information about the

defender's budget allocation. Though it is almost impossible for the attacker to know everything about the target network, the problem is described as a worst case scenario with specific assumptions and problem objectives in the following sections. In general, researchers focus on the node failure but link failure, which are more common to the real world; therefore, only node attacks are considered in this research.

2.2 Problem Formulation

The attacker behavior with mathematical programming problem is modeled as following. The given parameters and decision variables are shown in Table 1.

Table 1. Given Parameters and Decision Variables

Given parameter	
Notation	Description
V	The index set of all original nodes
L_1	The index set of all original links
L_2	The index set of all original nodes , which are artificial links
L_3	The index set of all artificial links connect to artificial origin or destination
W	The index set of all given critical Origin-Destination pairs *(s, n)*, where s is the source node, and *n* is the core node
$p_{(k)}$	The index set of 1-st node to k-th node on path p, where $p_{(k)} \in V$
M	A large number that represents the link disconnection
ε	A small number that represents the link connectedness
d_i	The discounted factor between [0, 1] that represents the effect of accumulated experience at the compromised node i without paying an extra fee, where $i \in V$
P_w	The index set of all candidate paths of the O-D pair w ,where $w \in W$
δ_{pl}	An indicator function, which is 1 if link l is on path p, and 0 otherwise, where $l \in L_1 \cup L_2 \cup L_3$, $\forall p \in P_w$
b_i	The budget of defense resources that allocated to node i, where $i \in V$
$\hat{a}_i$	Threshold of an attack cost leading to a successful attack, which is a monotone increasing function of b_i
Decision variable	
Notation	Description
y_l	1 if link l is compromised, and 0 otherwise, where $l \in L_1 \cup L_2 \cup L_3$
t_{wl}	1 if link l is used by the O-D pair w, and 0 otherwise, where $l \in L_1 \cup L_2 \cup L_3$, $w \in W$
c_l	Cost of link l, where $l \in L_1 \cup L_2 \cup L_3$
x_p	1 if path p is chosen, and 0 otherwise, where $\forall p \in P_w$

The problem is then formulated as the following minimization problem:
Objective function:

$$\min_{y_l} \sum_{p \in P_w} \left(\sum_{i \in V} \hat{a}_i \prod_{i \in p_{(k-1)}} d_i \right) x_p , \tag{IP 1}$$

Subject to:

$$c_l = y_l M + \varepsilon \qquad l \in L_2 \tag{IP 1.1}$$

$$\sum_{l \in L_1 \cup L_2 \cup L_3} t_{wl} c_l \le \sum_{l \in L_1 \cup L_2 \cup L_3} \delta_{pl} c_l \qquad \forall p \in P_w, w \in W \tag{IP 1.2}$$

$$\sum_{p \in P_w} x_p \delta_{pl} = t_{wl} \qquad \forall w \in W, l \in L_1 \cup L_2 \cup L_3 \tag{IP 1.3}$$

$$\sum_{p \in P_w} x_p \delta_{pl} \le y_l \qquad \forall w \in W, l \in L_1 \cup L_2 \cup L_3 \tag{IP 1.4}$$

$$\sum_{p \in P_w} x_p = 1 \tag{IP 1.5}$$

$$x_p = 0 \text{ or } 1 \qquad \forall p \in P_w \tag{IP 1.6}$$

$$y_l = 0 \text{ or } 1 \qquad l \in L_1 \cup L_2 \cup L_3 \tag{IP 1.7}$$

$$t_{wl} = 0 \text{ or } 1 \qquad \forall w \in W, l \in L_1 \cup L_2 \cup L_3 . \tag{IP 1.8}$$

The objective function is to minimize the total attack cost; the attacker minimizes the objective value by deciding which path will be attacked. IP 1.1 describes the definition of the link cost, which ε if the link functions normally, and $M + \varepsilon$ if the link is broken. IP 1.2 requires that the selected path for each O-D pair, w, should be the minimum cost path. IP 1.3 is the relations among t_{wi}, x_p and δ_{pl}. The auxiliary set of decision variables, t_{wi}, was used to replace the sum of all $x_p \delta_{pl}$. The substitution is to further simplify the problem solving procedures. IP 1.4 requires that the compromised link is equal to the total chosen path p. IP 1.5 and IP 1.6 jointly require that exactly one path is selected between each given O-D pair. IP 1.7 determines whether each link l is compromised, or not. IP 1.8 determines whether each link l is used to from a shortest cost path by O-D pair, w, or not.

3 Solution Approach

We model strategic interaction between attackers and defenders by mathematical programing and examine it with a optimal solution approach. Here, we demonstrate the generalized shortest path problem [6] to resolve the proposed model, which is a directed graph $G = (N, A)$ with n nodes and m arcs. There are two attributes on each arc m, which is represented by cost function $c: A \to R$ and weighted function $w: A \to R$. In [7], the weighted w is in multiplicative way of their problem. It can be optimally solved in polynomial time using their proposed Generalized-Reverse-Dijkstra

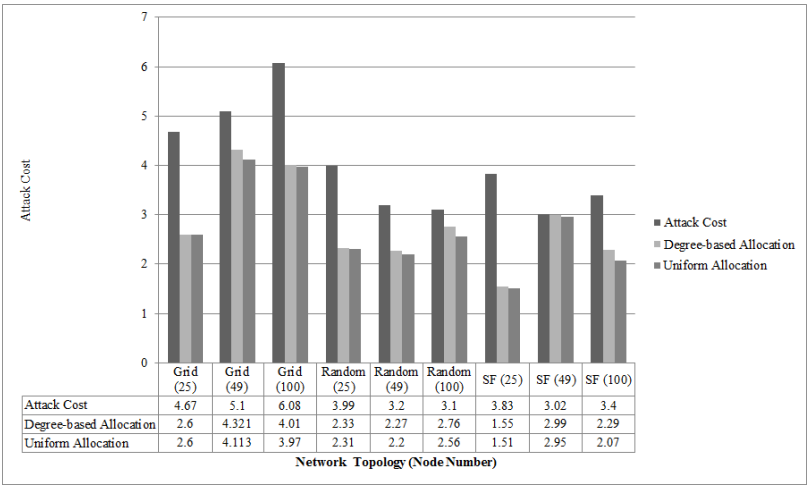
algorithm. Here, we use the weight to represent a discount factor of accumulated experience of sophisticated attacker. The accumulation of discount factors on a path might be additive or multiplicative [7]. The accumulation of discount factors is multiplicative is assumed here. We demonstrated a solution process as following.

Lemma 1 Given a budget allocation strategy, a topology, $G = (V, L)$, and critical O-D pairs, W, the formulation of the model can be optimally solved by the Generalized-Reverse-Dijkstra algorithm with the node splitting method. The time complexity is $O(|V^2|)$.

Proof. By adopting the node splitting technique, a node is divided into two independent sub-nodes, i' and i'', connected by an artificial link L',. The attributes of an artificial link L' inherit from original node, i.e., the attack cost and the discount factor. The related attributes, discount factors and attack costs, of the other links L are 1 and 0, respectively. We then transform $G = (V, L)$ into $G' = (V', L')$. Using the Generalized-Reverse-Dijkstra algorithm, we can then find the shortest path with the minimal cost in G'.

4 Experiments

We adopt three types of network topology as attack target: grid network, random network and scale-free network. To determine which budget allocation policy is more effective (i.e. resulting in a higher attack cost) under different cases, we design two initial budget allocation policies, uniform and degree-based, with different scale of a network. The former is to distribute the defense budget evenly to all nodes, the latter is to allocate budget to each node according to the percentage of a node's degree.



	Grid (25)	Grid (49)	Grid (100)	Random (25)	Random (49)	Random (100)	SF (25)	SF (49)	SF (100)
Attack Cost	4.67	5.1	6.08	3.99	3.2	3.1	3.83	3.02	3.4
Degree-based Allocation	2.6	4.321	4.01	2.33	2.27	2.76	1.55	2.99	2.29
Uniform Allocation	2.6	4.113	3.97	2.31	2.2	2.56	1.51	2.95	2.07

Fig. 1. Attack Costs of Different Networks and Allocation Strategies

The attack cost of different network topologies, node numbers, and budget allocation strategies is demonstrated in Figure 1. The Generalized-Reverse-Dijkstra algorithm heuristic and both budget allocation strategies perform well in all conditions (approximately 20% improvement on average)is shown, and degree-based budget allocation strategy is batter than uniform allocation. However, the difference between two budget allocation strategies is not significant. One possible reason is that in under the heuristic, the extracted defense resources are allotted to compromised nodes according to certain strategies. These nodes selected by the attacker already indicate which nodes are more important, although node degree is related to the importance of nodes. As a result, the gap between two strategies is small.

5 Conclusion

This paper focuses specifically on the behavioral impact of security practices and contributes to the growing body of work focused on applying the economic approach to security investment decisions, especially when those decisions take into account strategic interaction between different parties involved. The effects of accumulated experience and provide an evaluation of the robustness of a information system network from sophisticated attacker is consindered. This problem is then formulated as a mix integer mathematical model. The sophisticated attacker chooses a node as the starting node of the target network, and finds a minimal attack cost path. Moreover, this problem is successfully transformed into a revised shortest path problem, a generalized shortest path problem, which algorithm shows a pseudo-polynomial time.

However, this model can be further extended to a more complex attacker behavior. There are critical nodes of an organization. The attack target shall be one or more nodes. On the other hand, the defender may also have some active behavior to reduce attack impacts on the target network. For example, the defender may to try to fix the compromised nodes or reallocate the defense resource to increase the network survivability. Those interactions of the defender and attacker are still need to be discussed in future work.

Acknowledgments. This research was supported by the National Science Council of Taiwan, Republic of China, under grant NSC-102-2221-E-002-104.

References

1. McAfee, Advanced Persistent Threats, McAfee (2010)
2. Dacier, M., Pouget, F., Debar, H.: Attack Processes Found on the Internet. In: NATO Symposium IST-041/RSY-013, Toulouse, France (April 2004)
3. Bhuyan, M.H., Bhattacharyya, D.K., Kalita, J.K.: Surveying Port Scans and Their Detection Methodologies. The Computer Journal 54, 1565–1581 (2011)
4. Nicol, D.M., Sanders, W.H., Trivedi, K.S.: Model-based Evaluation: From Dependability to Security. IEEE Transactions on Dependable and Secure Computing 1(1), 48–65 (2004)

5. McDermott, J.: Attack-Potential-Based Survivability Modeling for High-Consequence Systems. In: Proceedings of the 3rd IEEE International Workshop on Information Assurance, pp. 119–130 (March 2005)
6. Ortalo, R., Deswarte, Y., Kaaniche, M.: Experiments with Quantitative Evaluation Tools for Monitoring Operational Security. IEEE Transactions on Software Engineering 25(5), 633–650 (1999)
7. Ellison, R.J., Fisher, D.A., Linger, R.C., Lipson, H.F., Longstaff, T.A., Mead, N.R.: Survivable Network Systems: An Emerging Discipline, Technical Report CMU/SEI-97-TR-013, Software Engineering Institute, Carnegie Mellon University (1997)
8. Ahuja, R.K., MAgnagti, T.L., Orlin, J.B.: Network Flows. Prentice Hall, Englewood Cliffs (1993) ISBN 978-0136175490
9. Batagelj, V., Brandenburg, F.J., Mendez, P.O.D., Sen, A.: The Generalized Shortest Path Problem, The Pennsylvania State University (July 2000)