Journal of Information Hiding and Multimedia Signal Processing ©2014 ISSN 2073-4212
Ubiquitous International Volume **5**, Number **4**, October **2014**

Threat Analysis of Cyber Attacks with Attack Tree+

Ping Wang and Jia-Chi Liu

Department of Information Management
Kun Shan University,
949 DaWan Road, Yongkang District, Tainan 710, Taiwan
pingwang@mail.ksu.edu.tw; momolu22@gmail.com

Received January, 2013; revised February, 2013

Abstract. *Defenders have developed various threat risk analysis schemes to recognize the intruder attack profile, identify the system weakness, and implement the security safeguards to protect the information asset from cyber-attacks. Attack trees (AT) technique play an important role to investigate the threat analysis problem to known cyber-attacks for risk assessment. For example, protection trees and defense Tree were used to analyze the system weaknesses against network threat. However, existing AT-based scheme provided a converse thinking to counter against attacks, ignored the dynamic interactions between threats and defenses and lacked the defense metrics for probabilistic analysis to real cyber-attack cases. Accordingly, the present study proposes a new method for solving threat analysis and risk assessment problem by means of an improved Attack–Defense Tree (ADT) scheme. Especially, defense evaluation metrics using Attack Tree+ for each node for probabilistic analysis is used to assisting defender validate the simulated attack results. Finally, a case of threat analysis of Zeus attack is given to demonstrate our approach.*

Keywords: Threat analysis; Risk assessment; Attack–Defense Trees; Attack Tree+

1. **Introduction.** The problem of identifying the attack profile of possible hackers over the Internet is referred to as the Threat Analysis (TA) Problem [10]. Tree structures have been widely used for exploring the potential attack profile by analyzing all the possible attack paths in the threat list utilizing a FTA (Fault Tree Analysis) approach. Generally, the TA focuses on evaluating the competition between attack and defense actions to determine the feasible defense strategies based on attack profile. Compared to TA, risk assessment more emphases collecting sufficient system vulnerabilities information to evaluate the risk level of asset, given a constraint on both the probability of attack occurrence and the potential impact loss. As described before, tree structure have been applied for exploring the attack profile based on FTA (Fault tree analysis) thru discovering all possible vulnerabilities associated with attack action (namely attack paths of threat list) since 1990s. Available Threat Analysis schemes with Attack Trees (AT) [4], such as Defense Trees (DT) [13], Protection trees (PT) [5] and Attack Response Tree (ART) [14]are capable of identifying the risk level and threats of an information asset via accumulating the system vulnerabilities, the corresponding impacts and estimating the attack costs and defense costs. Available threat analysis schemes for risk assessment, such as DT, PT and ART provide a means of stating the theoretical defense costs and lowering the risk, but do not reasonably answer the critical questions regarding: (i) cost-effective solution of defense solutions (ii) decide the suitable nodes to put safeguards in place. Notably, Kordy et al.(2010) proposed a new tree structure, namely Attack–Defense Trees (ADT) [2] to model

the interactions between attacks and defenses using game theory for arbitrary alternation between these two types of actions. Practically, ADT suffers from two facts: (i) only two notations are used to specify the complex attack defense scenarios. (ii) absence of the defense metrics for probabilistic analysis to real cyber-attack cases. Accordingly, an improved ADT scheme (iADTree), incorporating the defense thinking of ACT scheme [1], is proposed to investigate the threat analysis for APT (Advance Persistent Threat) attacks, allowing defender discover the possible defense policies to select the countermeasures associated with each of attack path. In the proposed approach, probabilistic analysis with defense evaluation metrics for each nodes is used to assisting defender analyze the attack sequence taking into account the proponent and opponent attitude. The effectiveness of the proposed approach was evaluated by a set of metrics for mitigating new cyber threats.

In developing the model proposed, there are three important aspects of focusing on our work: (i) explore the ROA(Return Of Attack) of targeted goal, (ii) examine the effect of ROI(Return Of Investment) with countermeasures, (iii) evaluate the possible countermeasure in accordance with the overall defense cost of the responding to these specified attacks and (iv) determine an appropriate allocation of the limited resources to achieve the best possible protection of the network security against cyber threats.

The remainder of the paper is organized as follows. The proposed model is introduced in Section 2. Section 3 takes an example to illustrate the method. Section 4 discusses how to select the optimal countermeasure to defense practically. Section 5 draws the conclusions.

2. An Analysis Model for Attack Profiles and Countermeasures.

2.1. **Basic attack modeling.** To appropriately selecting the proper safeguard under the circumstance of interleaving attacks, our model is capable of describing the attack profile, estimating the metrics of each node, and deciding the appropriate solution of countermeasure. In ADT, there are two basic types of events: attack node and defense node. It is too simple for two notations to specify the complex attack scenarios. Thus, the present study redefines the notation: attack event is break into two sub-events: detection (e.g., network exploits) and attack; defense event is separated into deception (e.g., honeypot) and countermeasure (fix vulnerabilities of host) as depicted in Fig.1 and Table 1.

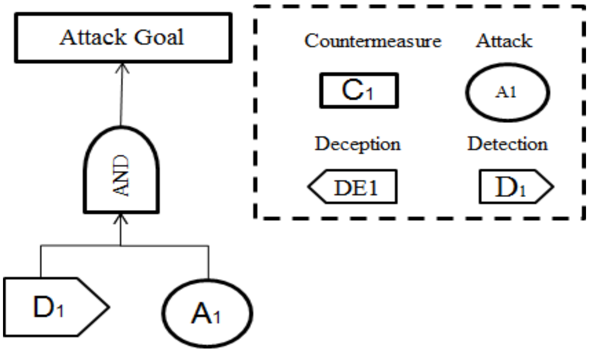


FIGURE 1. Notation of an iADTree

Basically, an iADTree can be consists of (i) a detection event and an attack event, (ii) multiple detection events and an attack event, (iii) multiple detection events, a deception, an attack and multiple countermeasures. In an example of attack and defense scenario depicted by iADTree, it is shown as Fig.2.

In real attack and defense scenario, assume an attacker wants to compromise the valuable asset of enterprise to steal privacy information from database server via the FTP

Table 1. Meaning of notation

Action	Examples	Notation
Attack	Registry modification, open ports	A
Detection	DNS query, port scan	D
Deception	Honeypot deployment	DE
Countermeasure	Vulnerability fix, safeguards put in place	M

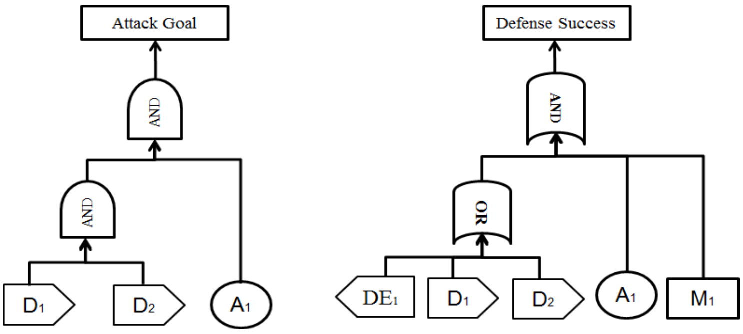


Figure 2. Attack and defense scenarios depicted by iADTree

server, the attack scenarios and its corresponding profile is depicted as shown in Fig.3 and Fig.4

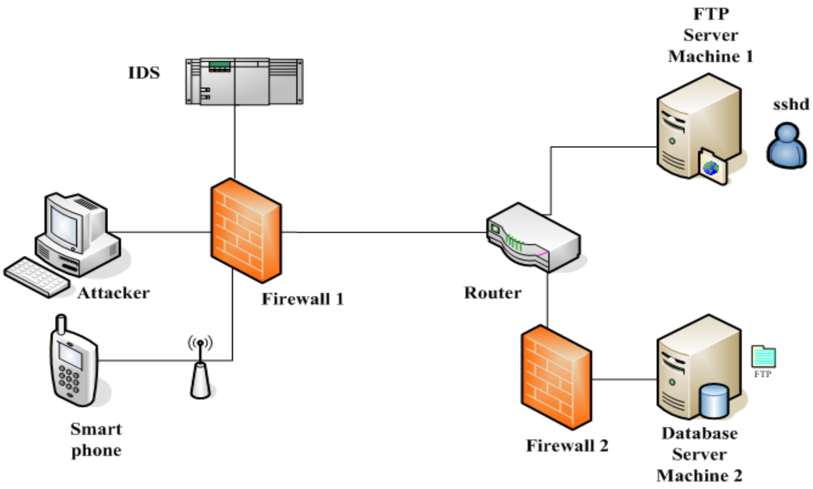


Figure 3. Attack scenarios of stealing privacy information

For probabilistic analysis, defender need estimate the probability of attack success for each node in iADTree. In Fig.1, Fig.2(a)(b) and Fig.4, the probability of attack success at the goal can be derived by Eqs.(1)˜(4), respectively.

$$p(t) = p_{A_1}(t)(1 - p_{D_1}(t)) \tag{1}$$

$$p(t) = p_{A_1}(t)(1 - p_{D_1}(t))(1 - p_{D_2}(t)) \tag{2}$$

$$p(t) = p_{A_1}(t)[1 - p_{D_1}(t) + (1 - p_{D_2}(t) + (1 - p_{DE_1}(t))](1 - p_{M_1}(t)) \tag{3}$$

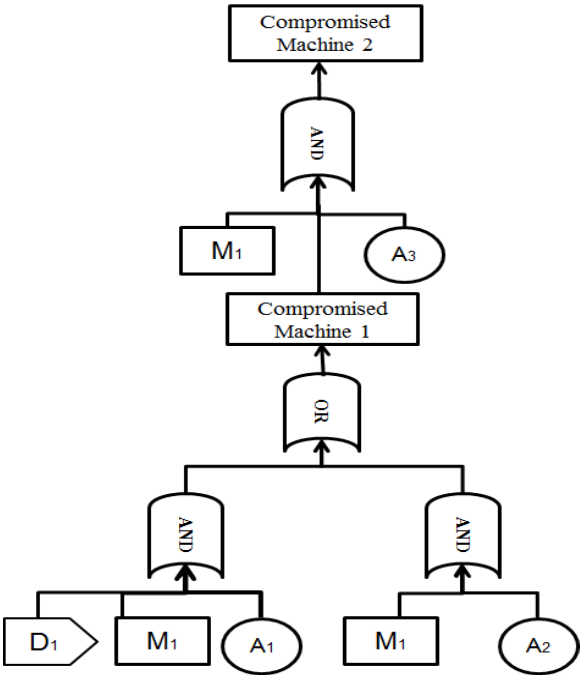


Figure 4. Attack and defense scenarios with iADTree

$$p(t) = [p_{A_1}(t)(1 - p_{D_1}(t))(1 - p_{M_1}(t)) + p_{A_2}(t)(1 - p_{M_1}(t))]p_{A_3}(t)(1 - p_{M_2}(t)) \quad (4)$$

2.2. **Estimating the success probability.** In solving the TA problem, a fundamental difficulty exists in assessing the success probability of basic attack actions, such as $p_{A_1}(t)$. Inspired by attack analysis concept on Intrusion Detection System (IDS), the success probability of attack occurrence is solved using 'episode frequency rules' [6] thru accumulating and associating the alert events as follows. Generally, episodes are partially ordered sets of events. The frequent episode rule is used to discover the specific event sequences as a means of appropriately estimating the probability of attack occurrence.

Given an event sequence *s =(s; Ts; Te)* and a window width win. Let time window of an episode *w=(w; Ts; Te)*. The support degree of an episode is defined as the fraction of windows where the episode occurs. In other words, given an event sequence s and a window width win, the support degree of an episode (α) in s is

$$\sup(\alpha) = p_i(\alpha, \mathrm{s}, \mathrm{win}) = \frac{|\{\alpha \, \mathrm{occurs\ \ in} \, \omega\}|}{|\{\mathrm{W}(\mathrm{s}, \mathrm{win})\}|} \quad (5)$$

Once $\sup(\alpha)$ has obtained, it can be used to predict $p_{A_1}(t)$ that describe connections between attack events in the given event sequence (i.e., signature).

2.3. **Attack and defense actions for threat analysis.** Suppose threat i is assumed to be composed of q basic attack actions *(k=1,. . . ,q)*, the metrics associated with the leaf nodes in the tree structure are calculated using AND-gate and Or-gate formulae of FTA[11] as shown in Table 2. The analysis of iADTree is constructed by the start from attacker's actions (leafs) thru recursively occupied sub-goals until attacker's goal (root node), as illustrated in Fig.3(a).

In Table 2, the probability of success of threat *i* (p_i) represents the chances of threat *i (i=1,...,m)* successfully hacking into the system, and has a value in the interval [0,1].

Meanwhile, the attack cost *(c_i)* represents the manpower cost required to carry out the attack, and is stated in terms of U.S dollars. The impact associated with a specific threat is measured on the scale of [1˜10], where a higher value indicates a more severe loss. Finally, the defense cost *(d_i)* represents the cost of defending against specific threat i and comprises both the security hardware cost and the defense manpower. Note that for simplicity, the man-hours used in evaluating the attack cost and defense cost, respectively, are converted to dollars at the rate of \$100 per man-hour.

Table 2. Rule set for iADTree metrics

	Non-leaf node k	
	AND	OR
Probability of success p$_o(t)$	$\prod_{k=1}^{q} p_k(t))$	$-\prod_{k=1}^{q}(1-p_k(t))$
Attack cost C$_o(t)$	$\sum_{k=1}^{q} c_k(t)$	$\underset{k}{\forall}\, Min(c)_k$
Impact l$_o(t)$	$\sum_{k=1}^{q} l_k(t)$	$\underset{k}{\forall}\, Max(l)_k$

In evaluating the performance of iADTree, two important metrics, i.e., *ROA* and *ROI* modified from [1] associated with each of the nodes is evaluated as follows: Return On Attack *(ROA)* of a non-leaf node *(k)*for specific threat i at time t can be evaluated by aggregating the attack cost *(c_k)*, the success probability *(p_k)* and the impact loss *(l_k)* as (see Fig.5)

$$A_k(t) = \frac{p_k(t) \cdot l_k(t)}{c_k(t)} \tag{6}$$

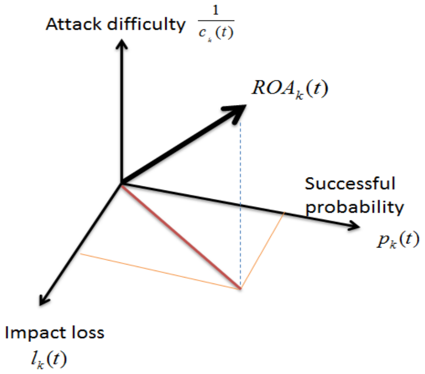


Figure 5. Affecting factors of ROA

For the root node, the overall *ROA* of the entire system at time t can be evaluated by either AND-gate or Or-gate computation

$$A_o(t) = \underset{k}{\forall} \max ROA_k(t) \tag{7}$$

$$A_o(t) = \underset{k}{\forall} \max ROA_k(t) \tag{8}$$

After analyzed ROA associated with a specific threat i, defender adopts the safeguards during the countermeasure stage (i.e., *t+1*) to decrease the *ROA* of attack. Thus, *ROI* at each non-leaf node is calculated as

$$I_k(t+1) = \frac{ROA_k(t+1) - ROA_k(t)}{d_k(t+1)}, \tag{9}$$

where dk represents the defense cost.

Having assigning values to the leaf nodes, the metrics are propagated up the tree until the goal node metrics are determined thru the link of AND-gate and OR-gate logic. Finally, the *ROA* and *ROI* value is obtained in each node.

Attacker's goal is to obtain the maximum result of *ROA* in terms of two resource constraints, i.e., attack cost ACk and the attack time ATk,

$$\underset{i}{\forall}, MaxROA_o(t), \tag{10}$$

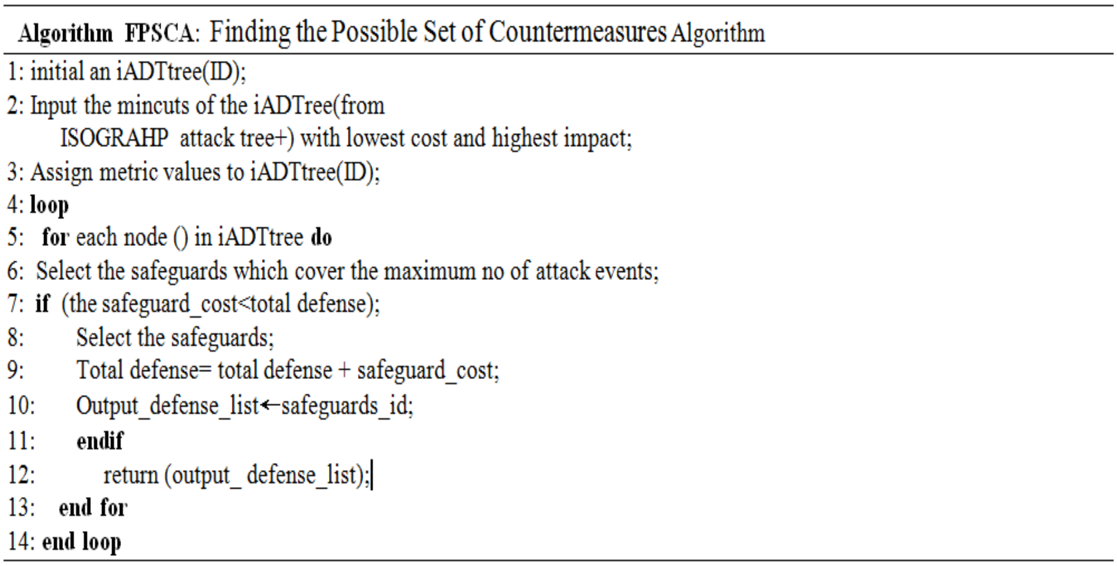
$$s.t.\quad \begin{cases} \sum_{k=1}^{q} c_k(t) \leq AC_k, \\ \sum_{k=1}^{q} \Delta at_k(t) \leq AT_k, \end{cases}$$

To eliminate the ROA to acceptable level, defender wants to Max ROI to ensure the effectiveness of countermeasure by minimizing the defense cost DCk within the time constraint of defense DTk, i.e.,

$$\underset{i}{\forall}, MaxROI_o(t), \tag{11}$$

$$s.t.\quad \begin{cases} \sum_{k=1}^{q} d_k(t) \leq DC_k, \\ \sum_{k=1}^{q} \Delta dt_k(t) \leq DT_i, \end{cases}$$

Input: parameters of attack actions of cyber threats and a pool of possible safeguards
Output: Suggested defense mechanisms given in a budget constraint



```
Algorithm FPSCA: Finding the Possible Set of Countermeasures Algorithm
initial an iADTtree(ID);
Input the mincuts of the iADTree(from
      ISOGRAHP  attack tree+) with lowest cost and highest impact;
Assign metric values to iADTtree(ID);
loop
  for each node () in iADTtree do
  Select the safeguards which cover the maximum no of attack events;
  if  (the safeguard_cost<total defense);
       Select the safeguards;
       Total defense= total defense + safeguard_cost;
       Output_defense_list←safeguards_id;
       endif
          return (output_ defense_list);
   end for
end loop
```

Figure 6. Algorithm FPSCA

For clarity, the defender generally may select the minimum number of defense mechanisms deployed in order to cover as many attack events of attack path in the iADTree as possible. Then the problem converts to discover the smallest possible set of countermeasures that contains at least one countermeasure from each minicuts of attack tree. [1] To achieve the above goal, the detailed algorithm for finding the possible set of countermeasures is described by PDL as Fig.6.

3. **Cyber Security Application over APT Attacks.** In the present study, APT (Advanced Persistent Threat) with Zeus attacks is used as the means of illustrating the threat analysis process. Typically, APT (Advanced Persistent Threat) attacks use a wide range of intelligence-gathering techniques and use malware tools to launch actual attacks. The primary goal of such attacks is to achieve a specific task rather than an immediate financial gain. In addition, APT attacks are usually carried out using the coordinated efforts of well-skilled, motivated, and organized individuals. Consequently, these hosts are vulnerable to attack, produce an impact loss to negative reputation. An example of a node containing APT attacks is that the potential threats caused by experienced hacker would attack the confidential file to a typical cloud service network. Zeus attacks[15], a new zero day PDF exploit the attack profiles of hackers on Cloud Computing services, will be analyzed. First identified in July 2007 when it was used to steal information from the United States Department of Transportation, it became more widespread in March 2009. In 2010, there were reports of various attacks by Zeus such that the credit cards of more than 15 unnamed US banks were compromised. The threat analysis process referring to [10] is constructed using the following four-step procedure.

Step 1: Understand of the system vulnerability

Generally, the recognized security vulnerabilities of computer have been investigated, examined and reported. For example, Mitre Corporation maintains a list of disclosed vulnerabilities in a system called Common Vulnerabilities and Exposures (CVE), where vulnerabilities are scored using Common Vulnerability Scoring System (CVSS). Vulnerability issued by US-CERT for Zeus botnet can be referred to [3].

Step.2: Collect the information of recognized attacks

Once the system vulnerabilities are identified, defender may focus on the issues of understanding possible network attacks that hackers maliciously attempt to compromise network security, as well as discovering attack profile with the probability of an event occurrence and its impact.

Step 2.1 Collect the malware

Deploy honeypot Dionaea [7] at switching edge node in the camp networks, log the alerts, and capture the payloads.

Step 2.2 Signature analyses

In the present study, attack profile is validated by CWSandBox and SandNets [9] in a dynamic malware analysis environment supported by Testbed@NCKU project [15]. Defender can examine the details of attack sequence to discover the possible attack profile. After collected the information from the aforementioned three sources of cyber-attacks, defender constructs the iADTree and predict the success probability of malware infection and hacker attack.

Step 3: Perform iADTree analysis

The metrics for intermediate and goal nodes shown in Table 2 operates on lower level nodes beginning with the leaf nodes. The partial iADTree in Figure 7 (see Fig.8) shows how an attacker might intrude into the servers for gain the root privilege thru exploiting IE vulnerability. After assigning values to the leaf nodes, the metrics are propagated up the tree until the goal node metrics are determined using Attack Tree+ tool [8], as shown in Fig. 9. Four major attack paths in Fig.9 are depicted in Table 3. Notably, the attack path, D2→C1→B1→A1 is the highest risk of threat occurrence which is selected as the first priority of safeguards to be implemented.

Table 3. The attack paths with support and confidence

Target	Attack,source	Attack paths
1	D_2	$D_2 \rightarrow C_1 \rightarrow B_1 \rightarrow A_1$
	D_1	$D_1 \rightarrow C_1 \rightarrow B_1 \rightarrow A_1$
	C_3	$C_3 \rightarrow B_2 \rightarrow A_1$
	C_3	$C_2 \rightarrow B_2 \rightarrow A_1$

Fig.10 illustrates that a possible way is given as a malware with the minimum attack cost is to gain root access in a host (see orange nodes). This attack path is the first priority of security controls to be deployed.

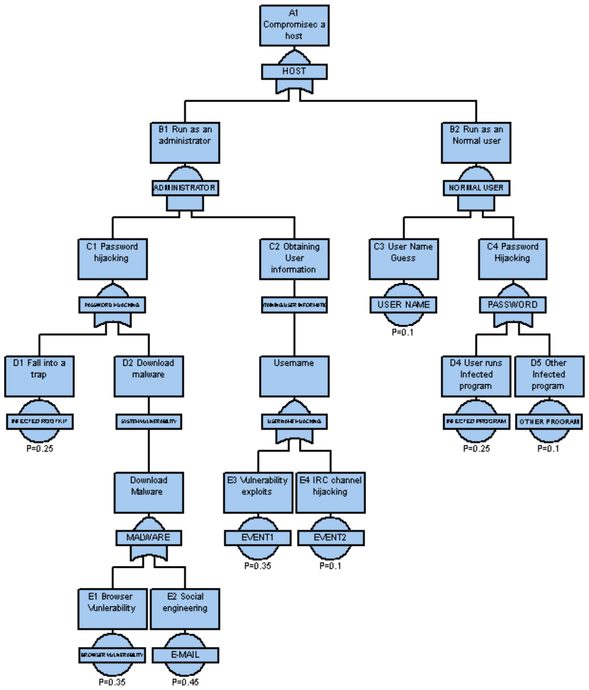


Figure 7. Compromised a host

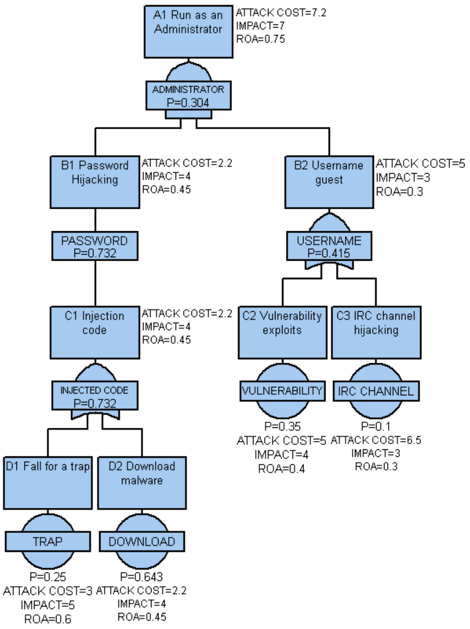


Figure 8. Attack path with attack cost

Step.4: Countermeasure analysis In practice, implement each security control can lower the distinct ROA values to attacker and increase the corresponding attack cost. Two crucial parameters and , are defined to specify the protection capability of safeguards as shown in Table 4.

where α represents the capability ratio of lowering the impact,β means the increasing ratio of the attack cost. Obviously, the higher α and is the better choice. For this example, S1 is selected as an illustration case with setting(α=0.325, β=0.40). After implemented with safeguard S_1 against infected code (see green blocks D_3) in Fig.8, the metrics are analyzed and filled in Table 5. Table 5 illustrates that assigning the safeguards will cost 5.6k indirectly con-verted to increase 40% defense cost to attack cost and eliminate 25% impact effect of attack to infected code(C_1). Consequently, the final ROA (A_1) will decrease from 0.75 to 0.543, impact loss declined from 7.0 to 6.7, success probability falls from 0.304 to 0.0167 and impose attack cost increase from 7.2k to 14.7k. The ROI of countermeasure on node D3 is 0.207. Similarly, other APT attacks could be analyzed

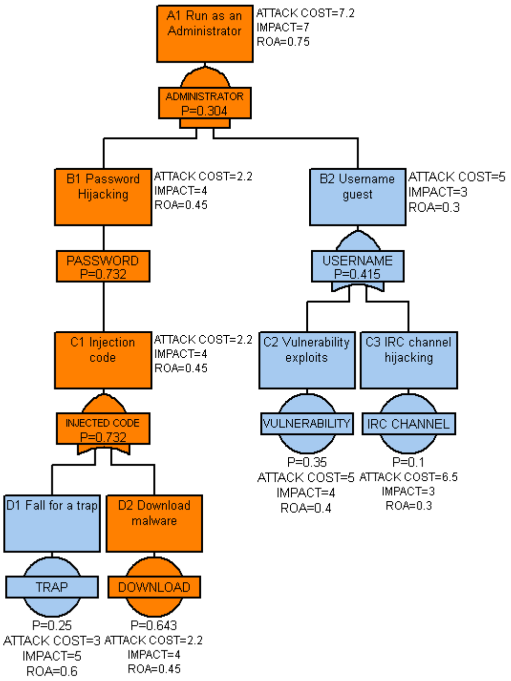


Figure 9. Path of the minimum attack cost

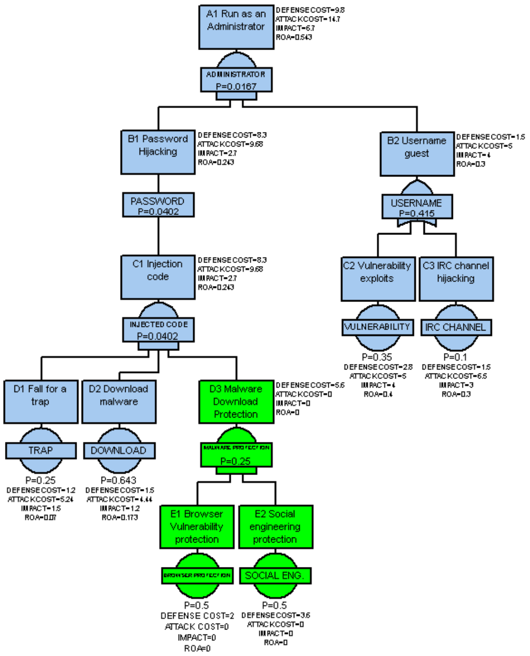


Figure 10. The attack path with defense actions

Table 4. Parameters for protection capability of safeguards

Safeguards, ID	α	β	ROA
			Before:0.45
			After: 0.243
			Before:0.51
			After:0.340
			Before:0.56
			After:0.220
			Before:0.35
			After:0.175

following in this case, such as Gh0stNet discovered in March 2009, is another example of APT to attack at least 11,296 hosts of south Asian.

Table 5. Metrics of attack path for IE vulnerability

Metric	Value
ROA_o (root)	0.75 → 0.543
Success probability (root)	0.304 → 0.0167
Attack cost (root)	7.2k→ $14.7k\$dollars$
Impact (root)	7.0 → 6.7
Defense costk	5.6k $ dollars
ROI_k	0.207

4. **Discussion.** How to implement the security controls in a cost-effective way for defender to cover all the attack and detection events, i.e., mincuts of iADTtree is an interesting research issue. To effectively enhance network survivability as considered with increasing the cost of attacks, defender need seriously evaluate the defense strategies using trade-offs between ROI and defense cost in a limited defense resources. There are many possible defense strategies to be chosen based on different defense logic, such as (i) select the safeguards which cover the maximum value of risks with the minimal defense cost, (ii) select the safeguards which cover the maximum number of attack events with minimal defense cost, and (iii) select the safeguards which cover the maximum number of attack paths (i.e., minicuts) with the minimal defense cost.

Fig.10 shows the defender constantly chooses the minimum attack cost associated with countermeasure implemented. At some cases, the highest ROA or impact loss will be primarily chosen to maximizing the security controls. In contrast to [2], iADTree approach take advantages on analyzing the interactions of attack and defenses, holding better flexibility by incorporating countermeasures.

5. **Conclusions.** In evaluating the performance of iADTree in cloud security services, different scenarios in accordance with a set of assessment metrics are examined using iADTree scheme for threat analysis of cyber security. It allows defender to convert defense cost with attack cost, and estimate the impact losses for the evolution of a system's security concerns. Furthermore, iADTree can simulate the interactive scenarios of attack and defense, consequently evaluate the required costs and examine the risk for a specified threat in an objective way. Finally, the proposed method improves the precision of the threat analysis and risk rating by facilitating defenders to make a right decision while exploring possible attacks.

Acknowledgment. Acknowledgment. This work was supported partly by TWISC@ NCKU, National Science Council under the Grants No: NSC101- 2221-E-168-034 and NSC 101-2219-H-168-001.

REFERENCES

[1] A. Roy, D. Kim and K.S. Trivedi, Cyber security analysis using attack countermeasure trees, *Proc. of Cyber Security and Information Intelligence Research Workshop (CSIIRW2010), ACM*, textit Oak Ridge, TN, USA, 2010.

[2] B. Kordy, S. Mauw, S. Radomirovic, P. Schweitzer, Foundations of attack–defense trees, *LNCS. Springer*, *Heidelberg*, vol. 6561, pp. 80-95, 2011.

[3] *DHS, CVSS, U.S. Department of Homeland Security library*, February 2005, available at http://www.dhs.gov/interweb/assetlibrary/NIAC05.pdf

[4] B. Schneier, Attack trees: Modeling security threats. *Dr. Dobbs Journal*, Dec 1999.

[5] K. S. Edge, G.C. Dalton II, R.A., Raines, R.F., Mills, Using attack and protection trees to analyze threats and defenses to Homeland security, *MILCOM* 2007, pp. 1–7, 2007

[6] H. Mannila, H. Toivonen,and I. A. Verkamo, Discovery of frequent episodes in event sequences, *Proc. of Data Mining and Knowledge Discovery*, vol. 1, no. 3, pp. 259-289, 1997.

[7] Honeynet Project, honeypot Dionaea, available at http://dionaea.carnivore.it/

[8] ISOGraph, attack tree+ , available at http://www.isograph-software.com/2011.

[9] J. Stewart, Behavioral malware analysis using Sandnets, *Computer Fraud & Security*, vol. 2006, no.12, pp. 4-6, December, 2006.

[10] Microsoft Corporation, Threat Analysis & Modeling, vol. 212, 2007, available at http://www.microsoft.com/en-us/download/details.aspx?id=14719

[11] M. Rausand and A. Hyland, System reliability theory; Models, statistical methods and applications, *Wiley*, 2004.

[12] S. A. Zonouz, H. Khurana, W. H. Sanders, and T. M.Yardley, RRE: A game-theoretic intrusion response and recovery engine, *Proc. DSN*, pp. 439–448, 2009.

[13] S. Bistarelli, M. D. Aglio, and P. Peretti, Strategic games on defense trees. *LNCS, Springer*, *Heidelberg*, vol.4691, pp. 1–15, 2007.

[14] Symantec, Zeus: King of the bots (PDF), available at http://www.symantec.com/ content/en/us/enterprise/media /security_response/ whitepapers/ zeus_king_of_bots.pdf.

[15] Testbed @TWISC, available at http://testbed.ncku.edu.tw/index.php3.