TOPOLOGY MAP ANALYSIS FOR EFFECTIVE CHOICE OF NETWORK ATTACK SCENARIO

Hidema Tanaka

National Defense Academy of Japan
Hashirimizu 1-10-20 Yokosuka, Kanagawa Japan 239-8686

ABSTRACT

In general, network attack should be prohibited and information security technology should contribute to improve the trust of network communication. Almost network communication is based on IP packet that is standardized by the international organization. So, network attack does not work without following the standardized protocols and data format. Therefore, network attack also leaks information concerning adversaries by their IP packets. In this paper, we propose an effective choice for network attack scenario which counter-attacks adversary. We collect and analyze IP packets from the adversary, and derive network topology map of the adversary. The characteristics of topology map can be evaluated by the Eigen value of topology matrix. We observe the changes of characteristics of topology map by the influence of attack scenario. Then we can choose the most effective or suitable network counter-attack strategy. In this paper, we assume two kinds of attack scenarios and three types of tactics. And we show an example choice of attack using actual data of adversary which were observed by our dark-net monitoring.

KEYWORDS

Network attack, Dark-net Monitoring, Topology map, Adjacency matrix, Laplacian matrix, Total accessibility matrix

1. INTRODUCTION

Network attack is not a special threat today, and its purpose and technologies are evolving complicate rapidly. APT (Advanced Persistent Threat) is seen frequency nowadays, and organizing new adversary groups are the typical problem. The members of adversary group disperse worldwide or are mal-distributed in a specific area (such as the country). The former case has the possibility that the group belongs to the worldwide terrorism organization. On the other hand, the latter case has high possibility that the group has governmental support. In this paper, we focus on the activity of the adversary group which exists in specific country. Needless to say, almost network attack uses IP packets. The specification of IP technologies is determined by the international standardized groups such as IETF [11] and ISO [12], and details of them are open to the public. As the result, IP packets used in network attack have the information concerning to the action of adversary at the same time. So, there are many security projects based on these facts such as Honey pot project [1], Dark-net monitoring [9] and so on. Almost existing projects are used for observation and analysis of network attack trend in the worldwide scale. From the view point of analysis of IP packet, these security projects are regarded as a passive use of this information. On the other hand, our motivation stands on the point of active use of information in IP packets to apply counter-attack.

As already mentioned above, we focus on the adversary group in the specific country. The IP packets from adversaries have information of network infrastructures, in the area. Therefore, we can analyze the topology map of the target area by collecting and analyzing IP packets from there. The characteristic of topology map can be estimated by the Eigen value of the matrix that is derived from the topology map. The research field of network dynamics develops the analysis method using Eigen value of topology map. Using these Eigen values, we propose a method of effective choice for network counter-attack strategy, which shows the most effective and suitable

DOI: 10.5121/ijcnc.2017.9608

one. Network attack such as DDoS, also changes the topology map and its characteristics. Focusing on this fact, we assume two kinds of attack scenarios and three kinds of tactics. To evaluate our proposal attack, we demonstrate using actual data obtained by our dark-net monitoring. Note that we cannot show all the details because they have many sensitive topics.

There are some previous works, which focus on the topology map analysis for network security. However, for example [6], all of them are researched for the purpose of developing defense technology, and there is no previous work for attack strategy. In this point, our work is very epoch-making one since we focus on the counter-attack.

2. PRELIMINARIES

2.1. OUTLINE

The characteristics of the network can be estimated by topology map analysis [5][8][14][26]. The topology map is expressed by some methods. In this paper, we take two kinds of method, which apply integer matrix; Adjacency matrix [19] and Laplacian matrix [24]. The Eigen value of each matrix shows the characteristic of topology map. In this paper, we focus on two types of characteristics; "Spread speed" and "Convergence". "Spread speed" denotes the characteristic, which shows easiness of communication. "Convergence" denotes the characteristic, which shows easiness of settling of information.

As an example of previous works using such Eigen values of topology map, there is a chain-reaction bankruptcy analysis of bank-transaction [15]. In this work, they derived some topology maps of bank-transactions and calculate their Eigen values. Using these Eigen values, they made it clear that only bankruptcy of mega-banks is not always the cause of the financial crisis. Network dynamics is the research field that analyzes a phenomenon using such characteristics of the network. In this paper, we have applied the basic technique of network dynamics to develop the method of network attack.

2.2. ADJACENCY MATRIX

Let G be a topology map with n nodes. Then G can be expressed as $n \times n$ Adjacency matrix A. Let $A_{i,j}$ $(1 \leq i,j \leq n)$ be an element of matrix A as follows.

$$A_{i,j} = \begin{cases} 1 \text{ if } i \text{ is adjacent to } j, \text{ and} \\ 0 \text{ if } i \text{ is not adjacent to } j. \end{cases} \tag{2.1}$$

Note that $A_{i,i} = 0$ because of $A_{i,i}$ denotes the link to itself. Let degree of node i be the Hamming-weight of i-th row (or i-th column). From the symmetry of matrix A, a condition of $A_{i,j} = A_{j,i}$ holds (i-th row and i-th column denote same adjacency of i-th and j-th node). The node, which has the largest degree, is defined as "hub-node". Let λ be the Eigen value of A, which is derived following characteristic equation.

$$\det(\lambda I - A) = 0 \tag{2.2}$$

Since the characteristic equation has the n-th degree, eigenvalue can have different $m(1 \leq m \leq n)$ values. Let $\lambda_{\max}(A)$ be the maximum value of λ. Then the value of $\lambda_{\max}(A)$ shows the characteristic of the connection density among hub-nodes. And it indicates the characteristic of "Spread speed" of topology map.

2.3. LAPLACIAN MATRIX

A topology map G also can be expressed by Laplacian matrix L. Let $L_{i,j}(1 \leq i, j \leq n)$ be an element of matrix L as followings.

$$L_{i,j} = \begin{cases} d_i \text{if } i = j, \\ -1 \quad \text{if} i \text{ is adjacent to } j, \text{ and} \\ 0 \quad \text{if} i \text{ is not adjacent to } j. \end{cases} \tag{2.3}$$

if i is adjacent to j, and eq. (2.3) if i is not adjacent to j, where d_i denotes the degree of i-th node. The Eigen values of L are also derived by the same way of Adjacency matrix, using eq.(2.2). So, we have $m(1 \leq m \leq n)$ different values for L as follows.

$$0 = \lambda_1(L) \leq \lambda_2(L) \leq \cdots \leq \lambda_{\max}(L) \tag{2.4}$$

The minimum value $\lambda_1(L)$ is always equals to zero. The second minimum value $\lambda_2(L) > 0$ shows algebraic connectivity of topology map. When $\lambda_2(L)$ has large value, the topology map has high connectivity. The maximum value $\lambda_{\max}(L)$ shows the difficulty of connection delay. The synchronization of topology map can be evaluated by the ratio $R = \lambda_2(L)/\lambda_m(L)$. When R has large value, it indicates the characteristic of "Convergence" of topology map.

3. BASIC IDEA

3.1. BACKGROUND

Nowadays, almost network communication is based on IP packet technology. The specification of IP packet is standardized and open to the public. Figure 1 shows the contents of IP packet structure and we can find that IP packet has much information in its header; protocol, source IP address, destination IP address, timeout and so on. In general, every network attack does not work when they do not follow the specification of IP packet. From this fact, two big topics are focused; defense topic and attack one.

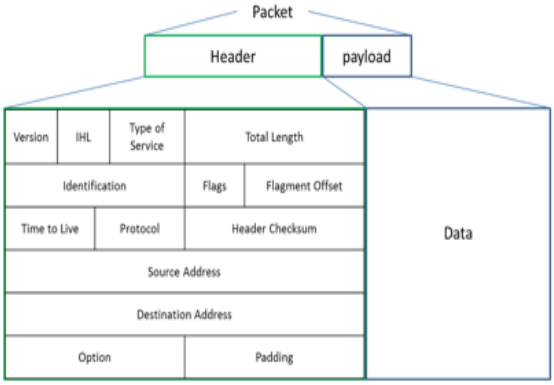

Figure 1. Contents of IP packet

In the defense topic, there are many projects to observe network attack using information in malicious IP packets. Among them, "Dark-net monitoring" is common to use in relatively large organizations such as governmental institutes, universities, enterprises and so on. "Dark-net" is local network space whose global IP addresses are not used. Therefore, IP packets, which arrived at IP addresses in Dark-net, are regarded as malicious action. Today, the analysis of Dark-net access (Dark-net monitoring) is regarded as a defense method to detect network attacks. There are many projects of world scale Dark-net monitoring, such as Nicter [16], Norse [23] and so on.

For the attack topic, we need to hide true IP address; forged IP address, spoofing, springboard and so on. In many cases, springboard PCs known was "bot-net" is common attack method and we can also find such access in Dark-net monitoring. In fact, there are many methods, which detect

springboard PCs and find out true malicious IP address [13] [21][25]. However, even if springboard PC is intentional or accidental, in this paper, we suppose those springboards PCs that execute persistent access to Dark-net are adversaries.In our proposal method, we observe Dark-net and collect IP address from the attackers. Then we classify them using country information in IP address and derive topology map of the adversary [10]. Therefore, our proposal method requires Dark-net monitoring operations in own organization.

3.2. DERIVING TOPOLOGY MAP

The "*trace route* command" is available on almost modern computer systems. It is a network diagnostic tool for displaying the route to given IP address. There are some existing results using trace route command to analyze network [2][4][22]. As shown above, IP address and IP packet have much information concerning to adversary. Our purpose is to derive network topology map attacking us. In our strategy, malicious IP addresses monitored in Dark-net are classified adversary group by categorizing their packets. To do this procedure, we collect different malicious IP addresses from same country or region. Then we execute trace route them, and we estimate the topology map of the target area. We call such topology map as the malicious topology map in the followings.

```
traceroute to ×.×.×.× (×.×.×.×), 30 hops max, 60 byte packets↓
 1  XXX.XXX.XXX.XXX  0.819 ms  0.821 ms  1.140 ms↓
 2  XXX.YYY.XXX.XXX  4.778 ms  4.787 ms  4.787 ms↓
 3  X.XX.XXX.XXX  13.239 ms  13.249 ms  13.249 ms↓
 4  XXX.XX.XXX.XX  13.247 ms  13.246 ms  13.245 ms↓
 5  XXX.XXX.XXX.X  123.796 ms  123.808 ms  123.808 ms↓
 6  XXX.XXX.XXX.XXX  135.251 ms  135.318 ms  135.302 ms↓
 7  XXX.XXX.XXX.XXX  135.324 ms  135.407 ms  135.394 ms↓
 8  X.XX.XXX.XX  191.892 ms  183.468 ms  183.445 ms↓
 9  XX.XXX.XX.XX  285.373 ms  285.369 ms  285.358 ms↓
10  XXX.XX.XXX.XX  286.124 ms  286.014 ms  285.999 ms↓
11  X.XXX.XXX.XXX  285.678 ms  287.925 ms  288.023 ms↓
12  X.XXX.X.XXX  288.006 ms  287.722 ms  287.419 ms↓
13  X.XXX.X.XXX  286.129 ms  284.697 ms  284.683 ms↓
14  * * *↓
15  X.XXX.XX.XXX  302.102 ms  301.368 ms  287.371 ms↓
16  X.XXX.XXX.XXX  295.975 ms  295.952 ms  285.994 ms↓
17  * * *↓
18  XX.XX.XX.XXX  288.504 ms  288.740 ms  288.484 ms↓
19  XX.XXX.XX.XXX  335.176 ms  334.767 ms  334.761 ms↓
20  XXX.XXX.XXX.XXX  344.397 ms  343.860 ms  343.853 ms↓
21  ×.×.×.×  345.501 ms  340.605 ms  339.085 ms↓
```

Figure 2. Example result of *traceroute* command

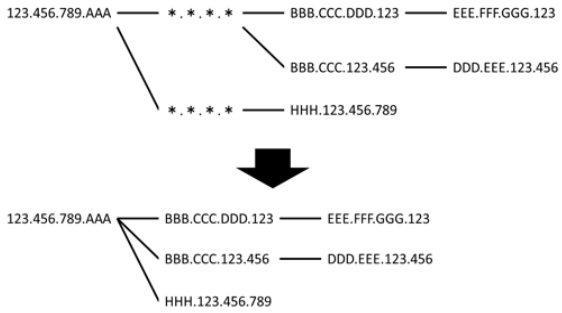

Figure 3. Estimating malicious topology map

However, the results of trace route command do not always show all IP address on the route. Figure 2 shows an example the of result of trace route command (note that symbol "X" denotes some numbers). In this figure, in line 14-th and 17-th, the symbol "* * *" denotes no answer from the server or the router on the path. In this case, unfortunately, we cannot know its IP address. First, we make a temporary topology map holding these unknown IP address. Then we delete these unknown IP addresses from temporary topology map, and we derive resultant topology map such as Figure 3. We define such resultant topology map as the malicious topology map. Actually,

there is an open project to estimate the detailed Internet topology map such as CAIDA [3]. However, our purpose does not follow their service policy. So, note that we derive a topology map by our own method. If we can get cooperation with the organization such as CAIDA, it is obvious that we can get precise malicious topology map easily.

The malicious topology map can be expressed by Adjacency matrix and Laplacian matrix. Therefore, as shown in section 2, we can analyze its characteristics by its Eigen values.

3.3. OUR STRATEGY

The threat scenarios of network attack are complicated and various, and they are evolving in every second. In this paper, we focus on following two types.

Scenario-1. Spread of malware and disinformation

Scenario-2. Concentration and confusion of information sharing

Scenario-1 is easy to understand and typical case of network attack, so we omit the details. The purpose of Scenario-2 is to generate the differentials in information sharing between the target area and others and make confusion among them. This scenario is also based on the one of the important characteristics of Internet technology such as immediacy of information sharing. By using these characteristics, we can generate a threshold of intentional diffusion of information. This scenario is similar to spreads of rumor in SNS, but it is different from such scenarios in the point that the difference in the spread of different information is generated deliberately. The effectiveness of these attack scenarios can be estimated by the characteristics of malicious topology map. Therefore, the effectiveness of Scenario-1 is related to the characteristic of "Spread of speed" and Scenario-2 is related to "Convergence" respectively [7][17].

On the other hand, network attack has various tactics such as DDoS attack, XSS, down of services, constructing rogue servers, and so on. These tactics have influence on the topology map and can change its characteristics. Therefore, the attacker can choose attack scenario and discuss its effectiveness by selecting tactics. In this paper, we consider following three tactics and simulate its effectiveness against change of characteristics of malicious topology map.

Tactics-1. Down of node

Tactics-2. Construction of agent node

Tactics-3. Combination of Tactics-1 and Tactics-2

Tactics-1 can be achieved by the well-know attack such as DDoS. Tactics-2 can be achieved by using IP address, which is not well managed.

There are some problems such as slow down of communication speed and feasibility, with the attack execution. And the choice and location of agent node have a big influence on the effectiveness of the strategy. These problems influence effectiveness and feasibility of strategy, however, they are individual problems for every actual malicious topology map, and we discuss this problem in section 7. In this paper, we search for the optimal attack effectiveness by brute force search, so, we limit the size of malicious topology map within our computer can analyze. Therefore, we limit the maximum number of nodes is 100.

3.3. EIGEN VALUE AND EFFECTIVENESS OF TACTICS

Let #n be a number of nodes and #l be a number of links in a topology map. Since the maximum value of #l means to make the complete graph with n nodes, the condition of $\#l \leq {}_nC_2$ holds. From the viewpoint of our tactics, we cannot make number of links than the condition. From this fact,

we can find that the maximum Eigen value is determined by the optimal tactics; $\lambda_{max}(A) = n - 1$ and $\lambda_{max}(L) = n$. We can see the relation between the value of (#*n*, #*l*) and each tactic, as follows.

Tactics-1. decreases #n and decreases #*l*

Tactics-2. increases #n and increases #*l*

Tactics-3 . holds #n and increases #*l*

Table.1 Correlation of number of nodes (n), number of links(l), Eigen value(λ) and tactics

n	l	λ	
—	—	—	
	↗	↗	Tactics-3
	↘	↘	
↗	—	—	
	↗	↗	Tactics-2
	↘	↘	
↘	—	—	
	↗	↗	
	↘	↘	Tactics-1

The changes of values of (#*n*, #*l*) and each Tactic are summarized in Table 1. In this table, the symbols "—", "↗" and "↘", denote unchanged, increase and decrease respectively. Note that we assume the number of down nodes equals to the number of generation of agent nodes in Tactics-3. From these facts, we expect followings.

Expectation-1. Tactics-1 will be useless for Scenario-1. When decrease of the maximum Eigen value is smaller than decrease of minimum Eigen value, Tactics-1 will be useful for Scenario-2.

Expectation-2. Tactics-2 will be useful for Scenario-1. When the increase of maximum Eigen value is smaller than increase of minimum Eigen value, Tactics-2 will be useful for Scenario-2.

Expectation-3. The effectiveness of Tactics-3 will be inferior to Scenario-1 than Scenario-2. On the other hand, Tactics-3 is most effective for Scenario-2, because it adjusts the balance of relation between maximum Eigen value and minimum one to maximize the value of *R*.

As the results, we can conclude as followings.

- It is enough for Scenario-1 to execute only Tactics-2 simulations.
- It is necessary for Scenario-2 to execute all of the tactics to search for most effective tactics.

In the followings, to discuss our expectations, we execute all tactics for both scenarios.

4. PROPOSAL METHOD FOR CHOICE OF COUNTER-ATTACK SCENARIO

Our proposal counter-attack scenario choice is defined as the combination of scenario and tactics shown in section 3.3. Since we have two kinds of scenarios and three types of tactics, we have total six patterns of counter-attack strategy. In fact, our proposal attack strategy satisfies following purposes.

Purpose-1. Choose an attack scenario and search for the most effective tactics.

Purpose-2. Choose an attack scenario and change the malicious topology map by the tactics.

Purpose-3. Search for the most effective scenario to the given malicious topology map and decide the attack tactics.

Purpose-1 and Purpose-2 can be regarded as a part of tactics from the viewpoint of the counter-attack operation. Our proposal counter-attack strategy will contribute such concrete purpose, however, these cases are too specific to describe in this paper. On the other hand, Purpose-3 is only executing some retaliate. The detail of Purpose-3 is ambiguous but it is general for almost counter-attack. Therefore, in the followings, we stand the position of Purpose-3.

The procedure of our proposal strategy is as follows.

Step-1. Collect IP addresses from the target area (malicious IP group).

Step-2. Execute trace route command for malicious IP group.

Step-3. Derive the malicious topology map.

Step-4. Execute simulation of Tactics-1 ~ Tactics-3 for both scenarios.

Step-5. Select the best result in Step-4 as the scenario.

In Step-1, we can use access log, Dark-net monitoring and so on. We used our Dark-net monitoring log since it does not need to extract malicious accesses in our experiments. An important point here is to collect many IP addresses as possible. The huge number of IP address helps to derive the malicious topology map correctly. We call these IP addresses as malicious IP group.

In Step-2, it is desirable to execute trace route command from more than one different place. And for even same IP address, it is desirable to execute changing time and a day of week sometimes. Because of the network traffic will change over time and a day of week, so there is the possibility that network routing changes. As the result, it is possible to get newer different IP addresses on the route and to derive more precise topology map.

In Step-3, we take the method shown in section 3.2. The details of Step-4 and Step-5 are shown in section 5. In this paper, we estimate computational complexity as the number of calculation of Eigen values. So, the computational complexity of Step-4 is determined by the number of total nodes (N) in the malicious topology map, the number of attack target nodes (n), the number of agent nodes (m) and the number of links from each agent node (l), as follows.

Tactics-1: $$C_1 = {}_N C_n \tag{4.1}$$

Tactics-2: $$C_2 = \sum_{i=1}^{m} {}_{(N+i-1)} C_l \tag{4.2}$$

Tactics-3: $$C_3 = {}_N C_n \times \sum_{i=1}^{m} {}_{(N+i-1)} C_l \tag{4.3}$$

5. EXAMPLE EXECUTION OF OUR COUNTER-ATTACK STRATEGY

5.1. STEP-1: COLLECT IP ADDRESS AND DARK-NET MONITORING

In the monitoring period (March 1st ~ 21st, 2013), we recorded total 1,654,925 of malicious access for our Dark-net. Among these accesses, there are 1,093,859 different IP addresses. Using the country information of IP address, the access numbers of each country are summarized in Table 2. In the followings, we focus on Country-C and Country-B because the size of topology map is adequate for our computer simulations. Note that we have no other intentions at all. We show the details of the procedure for Country-C mainly, and only results are shown in Country-B.

Table 2. Access numbers of each countries

Country	Access number	Independent IP address
Country-A	757,775	553,689
Country-B	75,785	53,390
Country-C	8,728	3,674
Country-D	3,896	2,089
Total	1,654,925	1,093,859

5.2. STEP-2: TRACE ROUTE

We executed trace route for 3,674 different IP addresses with the parameter as follows.

$$\$ \text{ trace route } -\text{I } -\text{n } -\text{m30} IP\_address$$

Using these parameters, we can get maximum 30 IP addresses on the route for target IP addresses. Note that we focus on the IP addresses in the Country-C. For the restriction in our network environment, we execute them from the only single start point, and we did not execute them changing time and a day of the week. As the result, we got 2,119 of new IP addresses in Country-C. We omit IP addresses, which do not exist in the result of trace route or isolate in the temporary topology map. Thus, we have 2,119 nodes with 3,819 links, which is smaller than the initial recorded 3,674 IP addresses. We needed about 2 days for this process.

In the same way, we executed same procedure against Country-B whose number of initial recoded IP addresses is 53,390. As the result, we got the resultant topology map of 17,684 nodes with 24,163 links. Since we had network troubles in experiment period, we needed about one month for this process.

5.3. STEP-3: ESTIMATION OF TOPOLOGY

Using the estimation method shown in section 3.2 for the resultants of traceroute, we have the malicious topology map of 2,119 nodes with 3,819 links as shown in Figure 4. But this topology map is too large for our computer environment to execute the proposal method. Therefore, we limited the number of nodes to 100 and focused on the nodes in the metropolitan area using the information of IP locator and *WHOIS* database. As the results, our target malicious topology map of 100 nodes with 187 links is derived as Figure 5.

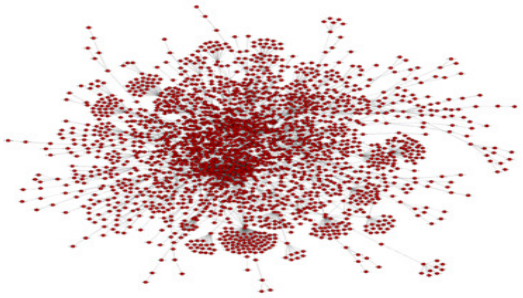

Figure 4. Malicious topology map in Country-C

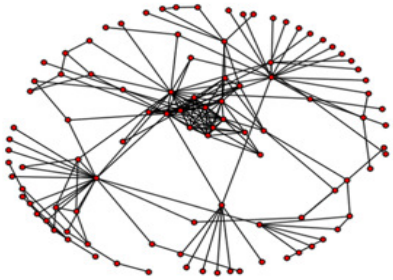

Figure 5. Malicious topology map with 100 nodes (metropolitan area in Country-C)

In the same way, Figure 6 shows malicious topology map of Country-B. By the same reason of Country-C, we derived the target malicious topology map of 100 nodes with 712 links, as shown in Figure 7.

Comparing Figure 5 and Figure 7, in spite of the same number of nodes in each other, the difference between them is obviously. We can find that Country-B has two huge high-density clusters however; Country-C has only one big hub-node and sparse topology map. The number of links causes this difference and it will have influence on the choice of attack strategy.

5.4. Simulation Of Tactics And Results

The initial values of target malicious topology map of Country-C are $\lambda_{max}(A) = 10.0785$ and $R = 0.005487$. Because of limited of specification of our computer environment (Table 3), we set the parameters of each tactics as follows.

N =100, n=100, m=1 and l=2.

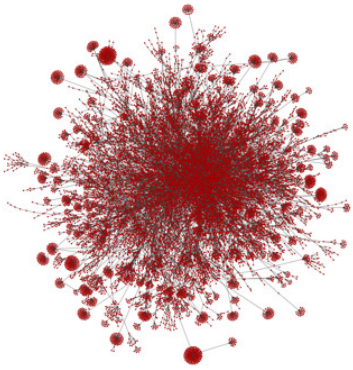

Figure 6. Malicious topology map in Country-B

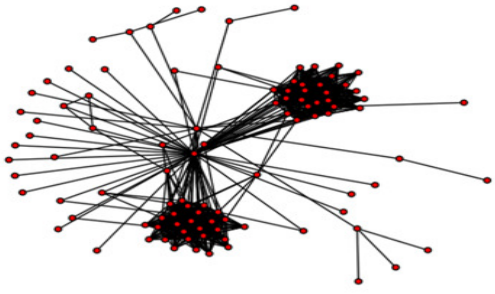

Figure 7. Malicious topology map with 100 nodes (metropolitan area in Country-B)

Table 3. Specification of our computer environment

OS	Windows 7 Professional 64bit
Compiler	python 3.3.5
CPU	Intel(R) Core(TM) i7-3770 CPU 3.40GHz
Memory	16.0GB

The computational cost and simulation time for each scenario and tactics are summarized in Table 4. In the same way, we executed our proposal strategy to Country-B. The initial values of target malicious topology map of Country-B are $\lambda_{\max}(A) = 24.2098$ and $R = 0.002853$. The number of nodes and the condition of tactics decide the computational cost. Since the conditions are same, the computational cost for Country-B is same as one of the Country-C (Table 4).

The results for Country-C show in Figure 8, Figure 9 and Table 5. And the result for Country-B shows Figure 10, Figure 11 and Table 6. Note that we omit IP addresses of target nodes; down nodes and agent node because they are sensitive information. We can derive following strategies for each country.

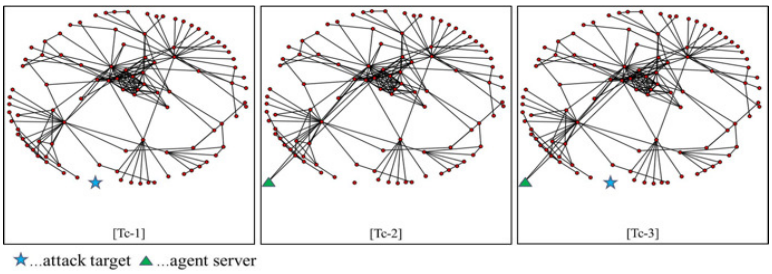

Figure 8. [Scenario-1] Spread of malware and disinformation (Country-C)

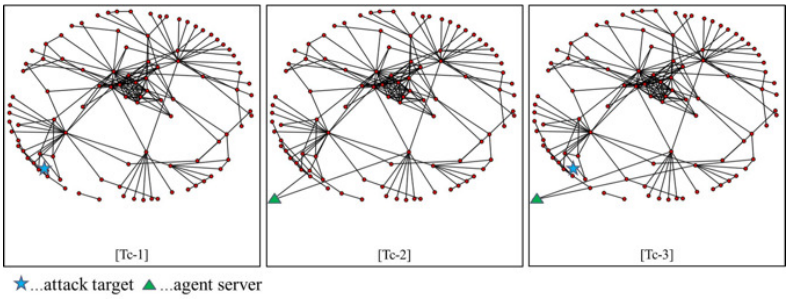

Figure 9. [Scenario-2] Concentration and confusion of information sharing (Country-C)

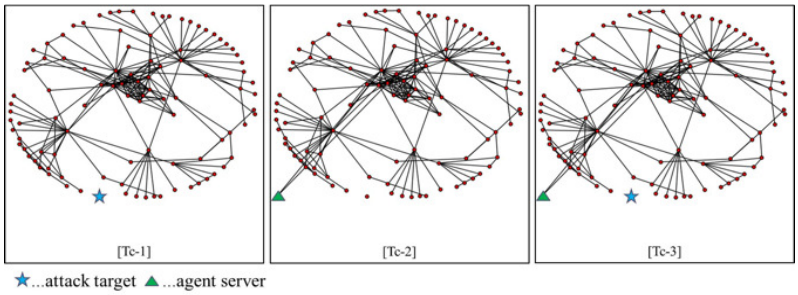

Figure 10. [Scenario-1] Spread of malware and disinformation (Country-B)

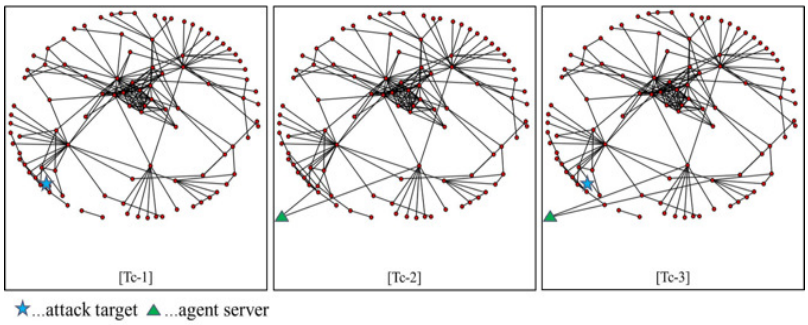

Figure 11. [Scenario-2] Concentration and confusion of information sharing (Country-B)

Table 4. Computational cost and simulation time for Country-C

	Scenario-1		Scenario-2	
	Computational complexity	Time[s]	Computational complexity	Time [s]
Tactics-1	100	1.4	100	2.0
Tactics-2	4,950	68.0	4,950	100.1
Tactics-3	485,100	21,651.9	485,100	27,699.4

Table.5 $\lambda_{\max}(A)$ and R of Initial topology and each Tactics (Country-C)

Map	$\lambda_{\max}(A)$	R
Initial topology map	10.0785	0.005487
Tactics-1	10.0785	0.005950
Tactics-2	10.1152	0.006329
Tactics-3	10.1152	0.007122

Table.6 $\lambda_{\max}(A)$ and R of Initial topology and each Tactics (Country-B)

Map	$\lambda_{\max}(A)$	R
Initial topology map	24.2098	0.002853
Tactics-1	24.2098	0.012527
Tactics-2	24.2165	0.003553
Tactics-3	24.2165	0.013652

<Strategy for Country-C>

Scenario-1. The effectiveness of Tactics-2 and Tactics-3 are same. And the effectiveness of Tactics-1 becomes same as the initial value. From these results, we can confirm that Expectation-1 and Expectation-2 shown in section 3.4, are almost right. As the result, we can conclude that the most effective tactics for Scenario-1 are Tactics-2 against Country-C.

Scenario-2. We can conclude that Tactics-3 is the most effective for Scenario-2 against Country-C. Therefore, we can confirm that Expectation-3 is almost right. Since it is not appropriate to show concrete IP addresses, we omit details however; Tactics-3 is able to success to divide the malicious topology map into three areas.

<Strategy for Country-B>

Scenario-1. The effectiveness of Tactics-2 and Tactics-3 are same. And the effectiveness of Tactics-1 becomes same as the initial value. These results are same as the cases of Country-C. Also from these results, we can confirm that Expectation-1 and Expectation-2 are almost right. As

the result, we can conclude that most effective tactics for Scenario-1 is Tactics-2 against Country-B.

Scenario-2. We can conclude that Tactics-3 is the most effective for Scenario-2 against Country-B. Therefore, we can confirm that Expectation-3 is almost right. Since the same reason for Country-C, we show only the result that Tactics-3 can success to divide the malicious topology map into six areas.

6. CONSIDERATION ABOUT EXAMPLE RESULTS

6.1. TACTICS-1 DOES NOT BECOMES SMALLER THAN THE INITIAL VALUE.

In Expectation-1, we expected that the results of Tactics-1 becomes smaller than the initial value, so there is no effectiveness in Scenario-1 with Tactics-1. Fortunately, both results of Country-C and Country-B hold the initial values. When such condition holds, we will be able to execute Scenario-1 and Scenario-2 simultaneously, because the target of Scenario-1 will not disturb the effectiveness of Scenario-2. The analysis of the feasibility of a concurrent execution of Scenario-1 and Scenario-2 is our future work.

6.2. TACTICS-3 IS THE MOST POWERFUL.

It is obvious that the condition of Tactics-3 for attacker is most advantageous. Therefore, it realizes more than 10% of improvement is estimated compared with the initial value of Scenario-2. However, there are some big problems such as huge computational cost, feasibility for realistic attack and so on. These problems are discussed in section 8. For Country-B, Scenario-2 with Tactics-1 has the second effectiveness that is much better than third one; it is little inferior to the best one. It is obvious to execute Tactics-1 is very easier than Tactics-3. Therefore, Scenario-2 with Tactics-1 can be the more suitable strategy for a case.

6.3. COMPUTATIONAL COST FOR TACTICS-3

Also, mentioned above, the computational cost for deriving Tactics-3 is huge. To solve this problem, we try to derive Tactics-3 using the results of Tactics-1 and Tactics-2. In Scenario-1, we will be able to derive Tactics-3 using them. Because the target server is same as Tactics-1 and the generated links as same as Tactics-2. Another computer experiments also show the same results. So we can conclude that Tactics-3 for Scenario-1 can be derived using results of Tactics-1 and Tactics-2. But, we cannot find out any relations among these results in Scenario-2. We conclude that it is necessary to execute separately in Scenario-2. Development of the method to reduce the necessary computational cost for Tactics-3 in Scenario-2 is our future work.

7. SOLUTION FOR PARAMETERIZATION OF ATTACK-TOLERANCE OF EACH NODE

7.1. BASIC IDEA

In the previous, we have set security levels of all nodes zero. Actually, such condition is not realistic. On the other hand, we cannot judge the security level of which node is low or high from the topology map or matrixes derived from Step-2 and -Step3.

To solve the problem, we assume that the node with the higher degree has high-security level and vice versa. There are some estimation methods for the degree of a node, we focus on Total accessibility matrix T [20]. In particular, Total accessibility matrix shows not only the degree but also accessibility according to degrees of neighbour nodes. Therefore, even if two nodes have the same degree, their accessibility may be different from each other. This characteristic is adequate for our purpose.

7.2. TOTAL ACCESSIBILITY MATRIX

For network G, let d be a diameter and A be (n×n) adjacency matrix, Total accessibility matrix T is calculated as follows,

$$T = A + A^2 + A^3 + \cdots + A^d. \tag{7.1}$$

Let $T_{i,j}$ be an element of T and V_i be accessibility of node i, we can calculate accessibility V_a of node a as follows,

$$V_a = \sum_{i=1}^{n} T_{i,a}\,. \tag{7.2}$$

In this paper, we have set target node group whose accessibility is less than the threshold $\mathbb{V}$ which is determined by the attacker.

7.3. EXAMPLE EXECUTIONS AND RESULTS

From adjacency matrix A, we calculate diameter d. In our examples, we obtain d=4 for Country B, and d=8 for Country C (Note that they are the cases for 100 nodes topology map). We focus on the significant change in increment of accessibility, we judge $\mathbb{V}$=3.06×10^5 from the view point of reasonable ratio of number of backbone nodes with high security and one of endpoint nodes with weak security. In the same view point, we set $\mathbb{V}$=1.75×10^7 for Country C. We determined these values from the sharp value change and the ratio of the number of nodes to the whole.In other words, we assumed that the important node in actual topology map is less than 30%. Therefore, about 70 nodes can be selected as attack targets because they are 100 nodes topology map of metropolitan area.

< TACTICS-2 FOR SCENARIO-1 >

From the above results, it is enough for verification of the effectiveness of Scenario-1 only to execute Tactics-2.We made agent node create links from the smaller value of accessibility. The results are shown in Table 7. In the case of Country-B, we can be successful in improving the effectiveness using 4 nodes. However, In the case of Country-C, sufficient effectiveness is not obtained. Focusing on the value of accessibility, the result of section 5 selects maximum and the second nodes (6.5203×10^5 and 4.2790×10^5 for Country-B and 1.7579×10^8 and 1.7399×10^8 for Country-C). These are nodes of the backbone and are expected to have high security and attack will be impossible. On the other hand, since the value of accessibility is limited in these experiments, feasibility is considered high.

Table 7. Results of Tactics-2 for Scenario-1 ($\lambda_{\max}(A)$)

	Country-B	Country-C
Initial value	24.2098	10.0785
Section 5.	24.2165	10.1152
2 links	24.2122	10.785
3 links	24.2150	10.786
4 links	24.2180	10.787

< TACTICS-1 FOR SCENARIO-2 >

The results are shown in Table 8. In the case of “one node down”, the same results are obtained. Therefore, we can conclude that Tactics-1 for Scenario-2 against Country-B and Country-C will be feasible. As can be seen from the destruction results of multiple nodes, destruction of a node with a small accessibility is effective for improving the value of R.

Table 8. Results of Tactics-1 for Scenario-2 (*R*)

	Country-B	Country-C
Initial value	0.0029	0.0055
Section 5.	0.0125	0.0060
one node down	0.0125	0.0060
two nodes down	0.0140	0.0067
three nodes down	0.0144	0.0077

< TACTICS-2 FOR SCENARIO-2 >

The results are shown in Table 9.The result of section 5 selects 5.5300×10^3 and 3.9134×10^5 for Country-B and 4.9643×10^6 and 2.4272×10^7 for Country-C. In this way, by connecting a node with a small accessibility and a large node, the value of *R* can be greatly improved. On the other hand, in the combination of small accessibility nodes, not much improvement can be expected without three or more linking. However, since the linking of the two nodes is also improved over the initial value, we can judge that the effectiveness of Tactics-2 for Scenario-2 is high.

Table 9. Results of Tactics-2 for Scenario-2 (*R*)

	Country-B	Country-C
Initial value	0.0029	0.0055
Section 5.	0.0036	0.0063
2 links	0.0035	0.0062
3 links	0.0038	0.0068
4 links	0.0041	0.0072

< TACTICS-3 FOR SCENARIO-2 >

The results are shown in Table 10. It is obviously advantageous, and it turns out that the effectiveness is clear. Even attacks using smaller accessibility nodes are sufficiently high effectiveness and high feasibility. It took the case of "3 nodes down and generating 4 links" for the most computation time and required about 3 days. In these experiments, this is also a realistic calculation time, and we can conclude that this tactical attack search is sufficiently practical.

We analyzed the influence of the degree and accessibility of chosen nodes on the improvement of *R*-value. As the result, we discovered that the rate of improvement of *R* becomes higher when nodes of the higher order with small accessibility are down. This means that down of the hub-node which has many links to end point nodes are effective. This is inconsistent with intuition, and it is possible to isolate a partial network efficiently.

7.4. ANALYSIS AND CONSIDERATION

From the above results, it can be confirmed that the estimation of the security level of each node by the value of accessibility calculated from Total accessibility matrix *T* is valid. Although details are omitted, by confirming the node to be attacked using actual topology map and analyzing them using WHOIS etc., we confirmed that the security level of these target nodes are not really high actually and they are suitable as attack targets. Similarly, to the analysis of bank bankruptcy shown in [15], we also found that the higher improvement of attack effect can be expected by attacking multiple small accessibility nodes rather that the attack to the central node with large accessibility.

Table 10. Results of Tactics-2 for Scenario-2 (R)

		Country-B	Country-C
	Initial value	0.0029	0.0055
	Section 5.	0.0136	0.0071
# of down	# of link		
1	1	0.0141	0.0069
	2	0.0141	0.0083
	3	0.0142	0.0085
2	1	0.0138	0.0081
	2	0.0144	0.0084
	3	0.0145	0.0092
3	1	0.0132	0.0086
	2	0.0146	0.0089
	3	0.0147	0.0097

On the other hand, it is also clear that the attack effect greatly depends on the shape of topology map.In particular, the difference between Country-B and Country-C is clear, and the effect is more pronounced toward Country-B.This result is considered to be a more effective feature than the attack against the defense policy.In other words, it can be predicted that Country-C, which tends to be dispersed in each node, has higher attack resistance.Such analysis is our future work.

8. DISCUSSIONS AND CONCLUSION

In this paper, we have proposed propose a method of choice for network counter-attack strategy using topology map analysis and show example executions. Since network attack bothers our usual operation, we believe such action should be prohibited. However, network attack also brings information concerning to adversaries, so we should observe them effectively. Our motivation is based on these facts. Using our proposal strategy, we can derive tactics that determine the position of target server and agent server, to execute the scenario. However, our proposal method does not enable to make an estimation of the actual attack effect. To make proposal method as more practical strategy, we need to solve following problems.

PROBLEM-1.ANALYSIS OF ACTUAL ATTACK RESULTS AND OPTIMUM VALUES OF $\lambda_{MAX}(A)$ AND R.

A relation between attack result and value of $\lambda_{max}(A)$ and R should be analyzed. Since the number of nodes and links determines the maximum values of them, they decide the topology map definitely. Thus, we can also derive tactics from the difference between the initial topology and resultant topology with maximum values. So, we can derive an optimum value of $\lambda_{max}(A)$ and R theoretically, however, there is no realistic meaning. Because the maximum values can be derived from only the complete graph. Even if the attacker has infinite powerful conditions, it is unrealistic to change the initial topology map to the complete graph. Therefore, we can conclude that the estimation of optimum values of $\lambda_{max}(A)$ and R is useless. On the other hand,in this paper, we estimate attack effect only in comparing with the initial value of $\lambda_{max}(A)$ and R. But it is not clear how increase from initial value is contributing to the attack result. The analysis of it is also our future work.

PROBLEM-2. ANALYSIS OF FEASIBILITY OF TACTICS-2 AND -3 IN REAL NETWORK ENVIRONMENT.

We face two problems in Tactics-2 and Tactics-3 as follows.

SETTING OF AGENT NODES

There are many un-managed IP addresses such as Dark-net. In particular, the cases which student group used IP address without notice, and operate phishing servers are reported, at some

universities [18]. From this fact, it will be easy to set agent nodes if we do not specify the location. Therefore, a set at the most effective location may be impossible, but we can conclude that this problem can be solved.

GENERATION OF LINKS

After the set of agent nodes, we need to generate links. There are two ways to realize it. One is to establish physical communication lines or construct new network infrastructure. Another is to forge routing tables. The former way is powerful but we cannot expect its feasibility. The latter way is realistic. Though we will need to forge many routers and their tables, the feasibility will be high for the same reason of above. In particular, when attack scenario and tactics are decided beforehand, the execution will be easy.

REFERENCES

[1] H. Artail, H. Safa, M. Sraj, L. Kuwatly, Z. Al-Masri, "A hybrid honeypot frame- work for improving intrusion detection systems in protecting organizational networks," Journal Computers and Security, Vol.25, No.4, Page.274-288, (2006)

[2] D. Bilo, L. Guala, S. Leucci, G. Proietti, "Network Creation Games with Traceroute-Based Strategies", Structural Information and Communication Complexity," LNCS Vol.8576, Springer, Page.210-223, (2014)

[3] Center for Applied Internet Data Analysis, http://www.caida.org/ (last access 2016-1-15)

[4] L. Dall'Asta, L.Alvarez-Hamelin, A.Barrat, A.Vazquez, A.Vespignani, "Traceroute-Like Exploration of Unknown Networks: A Statistical Analysis," Combinatorial and Algorithmic Aspects of Networking, LNCS Vol.3405, Springer, Page.140-153, (2005)

[5] M.Faloutsos, P.Faloutsos, C.Faloutsos, "On power-law relationships of the Internet topology," Computer Communication Review, Vol.2, pp.251-262, (1999)

[6] L. K. Gallos, R. Cohen, P. Argyrakis, A. Bunde, S. Havlin, "Stability and Topology of Scale-Free Networks under Attack and Strategies," Physical Review Letters, Vol.94, No.18, Page.188701.1-188701.4, (2005)

[7] S.Go'mez, A.Di'az-Guilera J.Go'mez-Gardenes, C.J.Pe'rez-Vincente, Y.Merono, A.Arenas, "Diffusion Dynamics on Multiple Networks," Physical Review Letters, Vol.110, No.2, Page.028701.1-028701.5, (2013)

[8] Y. Hayashi, "Robust Information Communication Networks based on Network Scientific Approaches," IEEJ Journal, Vol.130, No.5, Page.293-296, (2010)

[9] D. Inoue, M. Eto, K. Yoshioka, S. Baba, K. Suzuki, J. Nakazato, K. Ohtaka, K. Nakao, "Nicter: An incident analysis system toward binding network monitoring with malware analysis," Information Security Threats Data Collection and Sharing 2008, Page.58-66, (2008)

[10] The Internet Assigned Numbers Authority (IANA), http://www.iana.org/, (last access 2016-1-27)

[11] Internet Engineering Task Force (IETF) RFC: 791 INTERNET PROTOCOL, https://www.ietf.org/rfc/rfc791.txt (last access 2016-1-15).

[12] ISO/IEC 10731:1994 Information technology – Open Systems Interconnection – Basic Reference Model – Conventions for the definition of OSI services

[13] K. Kisamori, A. Shimoda, T. Mori, S. Goto, "Analysis of Malicious Traffic Based on TCP Fingerprinting," IPSJ Journal, Vol.52, No.6, Page.2009-2018, (2011)

[14] F. Luca, B. Paolo, G. Mario, "Interplay of network dynamics and heterogeneity of ties on spreading dynamics," Physical Review E, Vol.90, No.1, Page.012812.1-012812.9, (2011)

[15] A. Namatame, R. Zamami, "Systemic Risk on least susceptible network," Artificial Economics and Self-organization, LNEMS Vol.669, Springer, Page.245-256, (2013)

[16] National Institute of Information and Communications Technology, JAPAN (NICT), "nicterweb," http://www.nicter.jp/(accessed 2016-1-15).

[17] R.Pastor-Satorras, E.Smith, R.V.Sole, "Dynamical and correlation properties of the Internet," Physical Review Letters, Vol.87, Page.028701, (2000)

[18] Private communication with the security managers in the network security conference. They include many of customer information, we cannot open details.

[19] O.Rojo, R.Soto, "The spectra of the adjacency matrix and Laplacian matrix for some balanced trees," Linear Algebra and Its Applications, Vol.401, No.1-3,Page.97-117, (2005)

[20] Taaffee EJ, Gauthier HL (1973) Geography of Transportation. Prentice-Hall, Upper Saddle River, Ch. 5

[21] D. Takeo, M. Ito, H. Suzuki, N. Okazaki, A. Watanabe, "A Proposal of a Detection Technique on Stepping-stone Attacks Using," Connection-based Method,IPSJ Journal, Vol.48, No.2, Page.644-655, (2007)

[22] Y. Tomita, A. Nakao, "Inferring an AS Path from an incomplete Traceroute," The Journal of the Institute of Electronics, Information and Communication Engineers,Vol.109, No.273(NS2009 103 - 119), Page.17-22, (2009)

[23] U.S.A, Norse corporation, http://www.norse-corp.com/,(last access 2016-1-15).

[24] W.C.Wu, "On Rayleigh-Ritz ratios of a generalized Laplacian matrix of directed graphs," Linear Algebra and Its Applications, Vol.402, No.1-3, Page.207-227,(2005)

[25] R. Yokota, R. Okubo, N. Sone, M. Morii, "The affect of the honeypot on the dark-net observation, part 2," IEICE technical report, Vol.2013-GN-88, No.16, Page.1-4,(2013)

[26] Q.Zhou, Z.Li,"Empirical determination of geometric parameters for selective omission in a road network," International Journal of Geographical Information Science, Taylor & Francis, Vol.30, No.2, Page.263-299, (2016)

AUTHOR

Hidema Tanka is an associate professor of National Defense Academy Japan. His main research area is analysis of cryptographic algorithm, code theory, information security and cyber warfare and its domestic laws.