DOI: 10.5176/2251-3043_4.4.345
ISSN:2251 - 3043 ; Volume 5, Issue 1; 2016 pp. 1-7

Analysis of Organizational Vulnerability using Social Network Analysis and Attack Graph

Supachai Kanchanapokin and Associate Prof. Sirapat Boonkrong
King Mongkut University of Technology, North Bangkok

Abstract— **The purpose of this paper is to develop an approach to analyze organizational vulnerability caused by its employees. The proposed approach is adapted from general attack graph analysis approach and social network analysis approach. The attack graph, which is relationship graph in this proposed approach, is created from organization's email logs and virus reports. The relationship graph is analyzed using shortest path analysis to discover all possible attack paths start from risky employees to target employee, and then grouped by path length for further actions based on security policy. The proposed approach was tested using datasets that are limited to only one month with assumption that weight on all edges are equal. This paper suggested further study to improve accuracy of the proposed approach using other mathematical methods such as shortest path analysis with weight or Markov Chains. The proposed approach could also be used by security audit in risk assessment process.**

Keywords-attack graph; social network analysis; shortest path; graph theory; IT policy; security policy

I. INTRODUCTION

Cyber-attacks have been known since the 1960s with an objective to take control of the computer systems in an organization. Since then, analysts have attempted to find the way to determine risk from the vulnerabilities in the network and computer systems of an organization so they can mitigate or prevent them from being attacked. There are several tools and techniques that help analysts study the vulnerabilities. One of the most favorite tool is the attack graph approach which is the analysis based on tools and techniques developed from graph theory [1]–[9]. This tool helps analysts visualize all vulnerabilities and all possible attack paths so they can find patterns of attacks and improve security systems. However, the vulnerability of the organization is not only the computer network and computer systems, but also the people in the organization. The current attack graph approach has yet focused on the vulnerabilities caused by employees.

Attackers today are now using social engineering as a major attack to an organization because it attacks the people within the organization directly. Some employees in an organization have bad behaviors when using computers or IT systems which create vulnerabilities. For example, they click on a URL or double click on an attached file which was emailed from an unknown sender. Attackers exploit vulnerability of an employee to take control over his/her computer and try to get to other employees in the organization. It keeps going until the attacker gets to the target employee to collect critical information of the organization and takes control of the whole computer systems. If we know who causes or can be exploited in social engineering attacks, it should be possible find the way to prevent them on the first hand. We can say that an organization must improve their protection technology together with eliminating risk behaviors of employees in its organization. Since different behaviors create different security issues, organization should also be able to take specific actions for specific issue as well [10].

The problem is that, not like network and computer system vulnerability analysis, there is no structural approach to determine risk of an organization from vulnerability of people or employees. Since the vulnerability of people are caused by bad behaviors when using IT systems, an analysis of human behaviors is required. However, such analysis has been taken in sociology.

Similar to security of network analysis, Social Network Analysis (SNA) utilizes graph theory in the study too. The main ideas of this article is to find structural approach that integrate SNA and Attack Graph Network Analysis to analyze vulnerabilities caused by people in an organization.

II. RELATED WORKS

Newman has categorized networks into 4 groups [11]; social networks, information networks, technological networks, and biological networks. As stated in previous section, the development of new approach in this article were from several studies in two types of networks, the social network analysis in social networks and attack graph analysis in information networks. In this section, we are giving background of works that lead to our study. Those works are related to the computer network analysis and social network analysis.

A. Attack Graph Analysis

Attack graph is a graph-based approach that provides all possible ways that attackers will successfully gain control over organization computer network. The approach is widely used in vulnerability analysis of computer networks [1]–[5], [12]. The graph in attack graph is created from the attack model that represents network configuration and its vulnerabilities in an organization. However, different works present attack graph differently. In Ammann's attack graphs, vertices of the graph

represent state or privileges over a network component that attackers could obtain by exploiting vulnerability in the network systems, while edges represent actions to be taken to obtain each states. Sheyner [5], [6] use vertices to represent both states and specific attacks whether it is detectable or undetectable by intrusion detection system, while edges represent order of the attacks. Some may include attributes other than network attributes into the graph. For example, Dantu [2] and Williams [7] include attacker profile into each vertices on the attack graphs.

Attack graph is analyzed using several mathematical methods so analysts could find which attack paths are likely to happen and how its risks affect an organization. Phillips [8] suggested that single shortest path algorithms can find the lowest cost attack path which is the most possible path that attacker may choose. Zhang [9] and Smith[13] proposed Probabilistic Risk Assessment (PRA) method which uses probabilistic measures to determine system reliability. Madan [14] uses semi-Markov processes (SMP) to quantify security attributes of the intrusion tolerant system. This method helps to describe attacker's behavior and solve SMP. It also helps to calculate a security measure called "Mean Time to Security Failure (MTTSF)" which will be used to plan ahead on attacks.

In the method proposed by Jha [3] and Sheyner [5], attack graphs will be interpreted via Markov Decision Processes to compute the probabilities of intruder success for each attack. Mehta [4] suggested that the ranked attack graphs calculated by two algorithms, PageRank and Ranking States of attack graphs, are valuable for a system administrator as it allows them to estimate the security level of the system and provide a guideline for choosing appropriate corrective or preventive measures.

Williams [7] incorporate attacker's behaviors into attack models and construct attack profiles. The attack profiles are formed with assigned rating and will be used to compare with user profile and identify who is the attacker base on its rates. This approach will improve the robustness of the collaborative systems. Dantu [2] also uses attacker profile to create attack graphs and labelling attack path with behavior attributes identified in the attacker profiles. Attacker profiles are used for calculating the trust of a given attack path using probabilistic estimation (Bayesian statistic). As a result, network analyst can create patch for network security device according to the analysis.

As per the works above, all attack graphs are created from organization's network configurations. Its analysis also focuses on computer networks vulnerability, not the vulnerability of the people in an organization. This paper intends to find approach to fill this gap.

B. Social Network Analysis (SNA)

Similar to the computer network studies, "Social Network Analysis", or SNA also utilizes graph theory to generate mathematical model of the networks. Analysts can create graph to model social network which is known as "social graph" and analyze it. Moreno [15] used circles and lines to represent studied group in his work which was the first time graph was used in social study. Harary [16] improved SNA by using matrices as data structure of social graph. Since then matrices have become fundamental to social graph study.

SNA is the analysis of relationships between members in a social network. We can use a social graph to represent social network for analysis. In social graph, vertices represent people in society and edges represent relationship between people in the society. Social graph can be created from several sources of information. Newman [11] has founded that networks of different sources has different properties. For example, a network of emails has properties of directed graph while a network of telephones has properties of undirected graph.

Email has been used as source information to generate social graph in many researches. Newman [17] used email networks to study the spread of computer virus. He found that email viruses spread in organization network differently from human viruses because it was directed, unlike human network which was undirected. He also suggested that controlling email viruses problem in an organization shall use network structure to help identifying risk node in the network. However, his studies are on a few behaviors and properties of networks. There are more to study on the behaviors and functions of the network that will help analysts understand more of the real world networks. Ebel [18] also studied email network of Kiel University in the same year and found that the email network was scale-free and exhibited Milgram's small-world effect. Scale-free network shows some properties that have not been found in random network. In Scale-free network, there will be some nodes that act as hubs, while in random network, hubs are not allowed [20]. In 2003, Tyler [19] used email networks to find the structure of community within an organization. He used betweeness centrality to discover group of community within an organization.

Social network graph of email and attack graph have some common similarities. Both graphs are directed and have some vertices that act like hubs. This means that we may develop attack model and attack graph using social network graph, and analyze the model and its graph to determine risk of the whole organization as well. This assumption will be studied and described in the following sections.

C. General Attack Graph Analysis Approach

Attack graph analysis has been proposed and studied for many years. Those studies follow similar approach to prepare and analyze the attack graph in those studies. In order to understand the approach of attack graph analysis, we have summarized general attack graphs analysis approach as illustrated in the Figure 1.

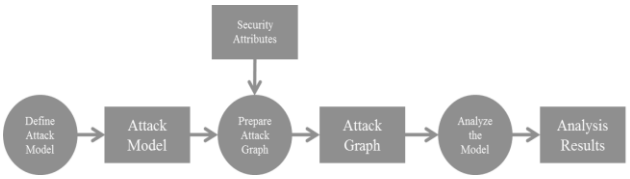

Figure 1. General approach of attack graph analysis

Each step in the above approach represented by circles and output of each step are represented by rectangles. These steps are described below.

1) Define Attack Model: In this step, analyst identify components of the environment that will be studied and, if needed, form mathematical formula to represent those components in the model. Objectives of the study also be defined in this step. The output of this step is Attack Model.

2) Prepare Attack Graph: In this step, the network will be visualized with graph using information gathering from the studied group. The graph represents logical relationship between components and its problems in the attack model, including with appropriate attributes. The output of this step is Attack Graph.

3) Analyze the Model: The graph from previous step will be analyzed using mathematical methods to determine that objectives of the analysis is met. In attack graph analysis, the common objectives of the analysis is to find the most probable attack path from the vertex of the start state to the end state of the attack.

Both attack model and attack graph are defined and created based on security attributes of a studied computer networks. There are tools available for analyst to create the attack model and attack graph. For example, NuSMV from Carnegie Mellon University is the favorite tool used by many researchers such as Jha [3], Sheyner [5], [6] and Ammann [1], while MulVAL was used by Ou [21]. Mathematical model(s) will be applied into the graph via these tools to identify which attack path is the most risky for the organization. Analyst will then find proper solution to mitigate it beforehand.

However, the general approach of attack graph analysis and its tools are not able help create attack model and attack graph of the studied employees network in an organization. The next section will explain how we adapted the general approach of attack graph analysis so it can help analyze attack model and attack graph based on employees' network.

III. The Proposed Approach

The similarity between scale-free social network graph and attack graph mentioned in previous section leads to the assumption that it is possible to associate social network analysis with attack graph analysis to determine risk of organization from user aspect. By integrating general attack graph analysis approach in Figure 1 with social network analysis approach, an approach for the study of risk analysis can be proposed (See Figure 2).

The approach in Figure 2 has been introduced by Kanchanapokin and Boonkrong [22] as an approach for further study of organization vulnerability using integration approach of attack graph and social network analysis. It has some changes from the methodology in Figure 1. First, the name "attack model" and "attack graph" have been changed to "relationship model" and "relationship graph". We have changed the attack model to relationship model to imply the step of creating social network model and social network graph of the organization so it can be understood what and how each member of an organization relates to each other. Our proposed approach has two phases. The first phase covers steps to define relationship model and prepare relationship graph from available information in the studied organization. Employee's risk properties were also added into the relationship model to generate a relationship graph. This can then be used to analyze possible risk that the organization might have taken from the risk of individual employee. These attributes could also be calculated from available information in organization such as email log and virus/malware log of each user node. As in this paper, we calculated risk attribute using antivirus report on employees' computer and put into employees' vertices generated from email log.

The second phase, we find attack path using graph shortest path analysis. We grouped all risky employees by the length of their shortest path found on individual into levels so we can apply appropriate security policy to them.

Below is explanation of each step.

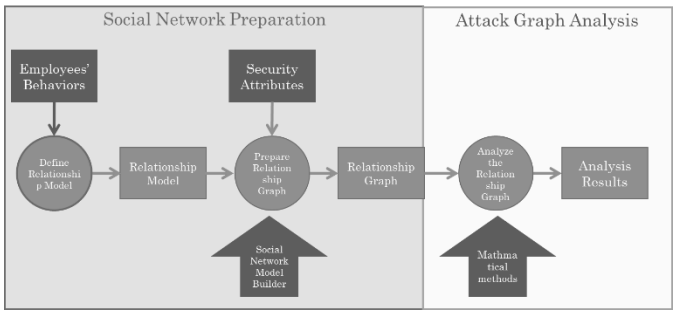

Figure 2. The proposed approach[22]

A. Defining the Relationship Model

Generally, the environment we studied was an organization that comprised of employees who had a chance to be compromised by an attacker. Those employees could be grouped into an employee primary attack, other employees, and an employee who would be the final target of the attack. Figure 3 below represents the environment of the organization we mentioned above.

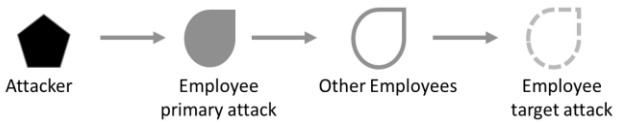

Figure 3. Environment of the studied organization

The employee primary attack represents the employee who has bad behavior that makes him or her vulnerable to an attack. An adversary could then exploit the vulnerabilities to collect information or gain control of his or her computer to get to next employees on the network. The interaction between employee primary attack and other employees are represented by a link between them. A linkage between two employees implies that both have interaction in some aspect. The employee target attack, also known as the destination of the attack, holds the most important information of the organization, or has authority to access the organization's information systems. The model of employees and its relationship is shown below in Figure 4.

Figure 4. Relationship model

In the Relationship Model illustrated in Figure 4, employees in the model are tagged with risk score R_i so they can be

determined who is likely to be an employee primary attack. The risk score is calculated from any information available for the organization. That means different organizations would provide a different risk score. The focus of this paper is on the bad behaviors in IT usage, such as frequently sending email with attached files to unauthorized email addresses [22], or the use of memory stick that has been infected with a virus on company computer. Therefore, the risk score represents employees' bad behaviors for the studied organization.

The interaction between two employees is tagged with a value that defines the importance of the relationship between them. This is known as a relationship score. This comes from the fact that an employee may consider interaction with some employee more important than another employee. Relationship score can be justified by the number of interactions between an employee and other employees next to him. For example if employee 1 sends an email to employee 2 more often than to employee 3 then the relationship score between 1 and 2 will be greater than the relationship score between 1 and 3.

Let $w_{i,j}$ be the relationship score between employee i and j. If the number of interactions between employee 1 and 2 is greater than the number of interactions between employees 1 and 3 then $w_{1,2}$ is greater than $w_{1,3}$. However, the number of interactions can be a lot more difference on each employee. We may not use the number of interaction between employees as a relationship score directly. Instead, we may use the ratio of interactions from an employee and other employees next to him.

Let $w_{i,j}$ be the relationship score from employee i to j and there are m employees next to employee i. If $f_{i,j}$, $f_{i,j+1}$, $f_{i,j+2}$, … $f_{i,m}$ be the number of interactions from employee i to j, $j+1$, $j+2$, …, m then $w_{i,j}$ is as follow.

$$w_{i,j} = \frac{f_{i,j}}{\sum_{j}^{m}(f_{i,j},f_{i,j+1},f_{i,j+2},\ldots,f_{i,m})} \quad (1)$$

We can say that the relationship score ($w_{i,j}$) is the ratio of interactions from employee i to employee j , or the possibility of interaction that may occur on employees i and all employees next to employee i.

In this paper, we have attempted this approach by extracting information from the real environment, and transformed it into the model. This is explained in the next section.

B. Preparing Relationship Graph

Social Network Analysis requires social network graph. The social network graph is network of vertices that represent people, and edges that represent relationship between those people. In order to build a social network for organizational analysis in this paper, we used vertices to represent employees, edges to represent interaction between employees and called it a relationship graph. We could build a relationship graph from interactions among employees in an organization. In working environment, employees interact to other employees via phone call, internal memo, business talking, meeting, email, etc. Those interactions imply their relationship.

In this paper, a relationship graph was created with information extracted from an organization's email log. The email log provided email addresses of employees which were identified as vertices' name. The email log also provided sender email addresses and recipient email addresses that informed us that there was a relationship between the two employees. Moreover, an edge between the two employees was used to define their relationship.

Let v_i be the vertices of employee i and $e_{i,j} = (v_i,v_j)$ be the directed edge from employee i to j. The relationship graph $G = (V, E)$ where V be a set of vertices of n employees and E be the set of directed edges is defined as follows:

$$G = (V, E) \quad (2)$$

$$V = \{v_i \mid i = 1 \text{ to } n\} \quad (3)$$

$$E = \{e_{i,j}=(v_i,v_j) \mid i,j = 1 \text{ to } n\} \quad (4)$$

Suppose the email address of each employee extracted from the email log are Emp1, Emp2, … , Empn then V and E are represented as:

$$V = \{Emp1, Emp2, \ldots Empn\} \quad (5)$$

$$E = \{(Emp1,Emp2), (Emp1,Emp3), \ldots\} \quad (6)$$

In the relationship model mentioned in previous section, the relationship graph has risk score attributes R_i on each vertex and relationship score $w_{i,j}$ on each edge. In order to generate a relationship graph, the vertices' name, edge, and all attributes must be transformed into graph data so it can be generated using a graph tool such as Pajek or Gephi. In this paper, Pajek was chosen as a graph tool for our analysis because it is easy to use and it provides common analysis tools such as shortest path analysis and shortest path analysis with weight.

In general, graph data has vertices record set and edge record set that are V and E in equation (5) and (6). The vertices record set comprises of vertex ids and vertex names. It can also include vertex attribute(s) such as sex, age, or score if needed. The edge record set comprises of two vertex ids on each line that form an edge record. For directed graph, the first vertex id is the start of an edge and the second vertex id is the end of an edge. It means the first vertex sent email to second vertex if the graph data represents relationship of email. Like vertices record set, the edge record set can also include attribute(s) such as weight or distance if needed. Figure 5 shows an example of vertices record set and edge record set, and relationship graph generated from these record sets.

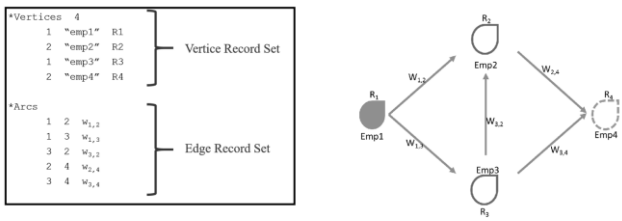

Figure 5. Sample graph data and its relationship graph

Once the relationship graph has been generated, the graph can then be analyzed using the method explained in the next section.

C. Analysing the Relationship Graph

In this step, the relationship graph from previous section is analyzed using mathematical methods such as shortest path analysis, Markov analysis or Bayesian analysis. This step could help us find groups of risky employees and possible attack paths on the relationship graph.

In order to analyze the relationship graph, it is important to find a vertex that would be an employee primary attack. As

defined in the relationship model, the employee primary attack is an employee who has the highest risk score *R* and the employee target attack is an employee who responsible for crucial information of the organization.

By using Pajek, we can scope our analysis from the whole relationship graph into a subgraph that consists of only vertices and edges containing relationship between the primary and target employees. Possible attack paths from one vertex to the other can then be found.

In Figure 6, suppose that the attack path starts at vertex Emp1 and ends at vertex Emp4. The possible attack paths could be Emp1-Emp2-Emp4 or Emp1-Emp3-Emp4 or Emp1-Emp3-Emp2-Emp4. However, it can be considered that the shorter route an attacker can take to get to Emp4 from Emp1, the more risk the organization will have. This means that the attack path which is the most risky for this organization is the shortest path between Emp1 and Emp4, which is Emp1-Emp3-Emp4 or Emp1-Emp2-Emp4, in this case.

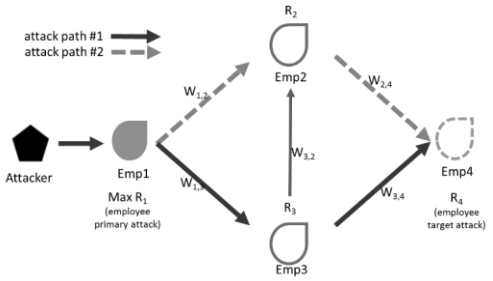

Figure 6. Finding possible attack paths on relationship graph

Employee with highest risk score might not cause the most risk of an organization if there is no path from this employee to target employee. In order to determine which employee can cause the highest risk to an organization, path lengths of all employees to the destination must be compared with one another.

By comparing the lengths of attack paths of the different employee primary attacks, it is possible to determine which vertex generates the higher or even the highest risk to the whole relationship graph. As a result, the organization is able to execute proper actions on risky employees differently.

In any graph or organization, it is likely that more than one path of the shortest length can be found. When this occurs, the one path with the highest risk can be identified by applying other mathematical methods such as shortest path with weight or Markov chains. The relationship score ($w_{i,j}$) on edges can be used in this case. In this paper, we are just going to verify the proposed approach so we will use shortest path without weight method.

IV. Testing the Approach

A. Prepare and Calculate the Relationship Graph

In this paper, our relationship model comprises of an employee primary attack, an employee target attack and other employees in between them. Those employees form a social network graph with vertices and edges representing relationship between them.

In this study, a social network graph of the organization's e-mail activities logs was prepared and formed. Each employee is represented by an individual email account and the transmission of an email (both sending and receiving) is represented by an edge. The log of March 2013 was chosen for the study and transformed into graphs data containing information of 8,858 vertices and 26,404 edges. From this step, a relationship graph was generated via Pajek.

The risk score can be calculated using various information related to an employee such as number of spams, number of virus detected, or number of emails sent to unauthorized accounts outside the organization [22]. The verification of the proposed method, in this particular example, is done by calculating the risk score from the number of malicious code detected on each computer which was extracted from the virus protection monthly report on the same period of the email log (March 2013). In order to illustrate the proposed approach, this paper uses one risk property which is the number of viruses found on the computer of employee *i*. The value R_i can then be calculated as follows.

$$\mathrm{R_i} = \frac{\text{number of virus found on employee i}}{\text{total number of virus found}} \quad (7)$$

The value R_i or the risk vector is then assigned to each vertex. The relationship scores, $w_{i,j}$ were also included into the graph data. They were calculated by counting the number of emails sent from v_i to v_j. However, as stated in the previous section that shortest path without weight was the method used, the relationship scores are included in the relationship graph without being used in the analysis phase. This is because we would like to illustrate that only the shortest path method can also be used for the risk analysis.

B. Analysis and Discussion

In the studied organization, there were some vertices with risk scores that were assumed to be employee primary attack vertices. By ordering the risk scores of all vertices from highest to lowest, it would inform us which employee to be analyzed first. Table I shows all vertices that could be employee primary attack and employee that would be an employee target attack. Please note that emails were changed to preserve confidentiality of the studied organization.

In this study, it was found that the first 4 employees who had the highest risk scores were the ones with the vertex IDs "5987", "1141", "5119" and "7929" in this order Therefore, they were assumed to be employee primary attack vertices. Moreover, one business director "grs@abc.co.th" (vertex ID: 7000) was chosen to be employee target attack due to his responsibility for business information of the organization.

In this study, it was considered that the employee primary attack, who had the highest risk vector or R_i and had the shortest path length to employee target attack, would generate the most risk to the company. The shortest path length or "shortest attack path" means the least path length (number of hops) between employee primary attack and employee target attack.

By using Pajek to find "All shortest path between two vertices" when the first employee primary attack was "5987" and employee primary attack was "7000", it was found that there was no path between them. Therefore, the shortest path length was 0. This means that even though the vertices ID "5987" had the highest risk score, it does not always mean that there is a path to the destination. Therefore, there is no reason to consider the employee vertex id "5987" in this particular case.

TABLE I. LIST OF EMPLOYEE PRIMARY AND TARGET ATTACK ORDERED BY RISK SCORE

VID	VName	Risk Ratio	Primary/ Target
5987	pop@abc.co.th	0.3810	Primary
1141	cda@abc.co.th	0.2238	Primary
5119	nun@abc.co.th	0.1048	Primary
7929	tat@abc.co.th	0.1048	Primary
6634	suk@abc.co.th	0.0429	Primary
8615	wya@abc.co.th	0.0381	Primary
2938	jas@abc.co.th	0.0333	Primary
6569	sin@abc.co.th	0.0143	Primary
5379	pda@abc.co.th	0.0095	Primary
5534	pch@abc.co.th	0.0095	Primary
8203	uha@abc.co.th	0.0095	Primary
1049	crn@abc.co.th	0.0048	Primary
2211	gan@abc.co.th	0.0048	Primary
5363	pri@abc.co.th	0.0048	Primary
5859	pen@abc.co.th	0.0048	Primary
5981	pon@abc.co.th	0.0048	Primary
7587	sol@abc.co.th	0.0048	Primary
7000	grs@abc.co.th	0.0000	Target

Using the same step with second employee primary attack whose vertex ID was "1141" and the same target, it was found that the shortest path between them was 1. Hence, the shortest path length was 1. This means that vertex"1141" had lower risk score than the vertex "5987", but had caused higher risk to the organization. This is because the vertex "5987" has no path to the target vertex "7000" while the vertex "1141" has path length 1 (or only one hops) to the target vertex "7000".

After all the vertices were processed, the results were summarized in Table II, ordered by its length from lowest to highest, and its risk ratio from highest to lowest. Please note that the vertices with path length of zero were included at the bottom because there should not be direct impact to the target from them.

As per the result in TABLE II, the risky employees can be classified by path length into three groups. We claim that the length of an attack path is more important than employees' risk ratio because it shows how quickly attackers can reach their target.

This way analysts can apply appropriate security policy for each of the employees. For example, security policy for the first group is to immediately suspend their accounts and let MIS officer to clean virus from their computers before reactivating their accounts in new name. The second and third groups would be the same but would be allowed to use old accounts. However, the two employees in the first group who has highest risk ratio (VID:1141, 7929) may be interrogated by security committee for appropriate actions.

TABLE II. ANALYSIS RESULT GROUPED BY LENGTH OF SHORTEST ATTACK PATH

Employee Primary Attack			Employee Target Attack		Length
VID	*VName*	*Risk Ratio*	*VID*	*VName*	
1141	cda@abc.co.th	0.2238	7000	grs@abc.co.th	1
7929	tat@abc.co.th	0.1048	7000	grs@abc.co.th	1
6569	sin@abc.co.th	0.0143	7000	grs@abc.co.th	1
8203	uha@abc.co.th	0.0095	7000	grs@abc.co.th	1
1049	crn@abc.co.th	0.0048	7000	grs@abc.co.th	1
2211	gan@abc.co.th	0.0048	7000	grs@abc.co.th	1
5119	nun@abc.co.th	0.1048	7000	grs@abc.co.th	2
6634	suk@abc.co.th	0.0429	7000	grs@abc.co.th	2
5363	pri@abc.co.th	0.0048	7000	grs@abc.co.th	2
8615	wya@abc.co.th	0.0381	7000	grs@abc.co.th	3
2938	jas@abc.co.th	0.0333	7000	grs@abc.co.th	3
5379	pda@abc.co.th	0.0095	7000	grs@abc.co.th	3
5859	pen@abc.co.th	0.0048	7000	grs@abc.co.th	3
5981	pon@abc.co.th	0.0048	7000	grs@abc.co.th	3
7587	sol@abc.co.th	0.0048	7000	grs@abc.co.th	3
5987	pop@abc.co.th	0.3810	7000	grs@abc.co.th	0
5534	pch@abc.co.th	0.0095	7000	grs@abc.co.th	0

By following the approach in Figure 2, we created the relationship model, relationship graph and analyzed the graph to get the results, just like the general approach in Figure 1 can do. Only now we can include risk score calculated from user behavior into the relationship model and the relationship graph. We have found that the relationship graph can be analyzed using attack graph analysis as well. This knowledge expand the use of attack graph analysis for risk analysis caused by employee as we expected.

V. CONCLUSION

Attacker exploiting vulnerability of people in an organization could damage its organization as much as vulnerability of computer network systems. However, the problem of the current attack graph analysis approach has yet focus on people in an organization. It is believed that the attack graph analysis approach can also be adapted to analyze attack graph of an organization that caused by its people. However, the new approach should be able to put user behavior into the analysis. In this paper, the proposed approach help us created the relationship model, relationship graph and analyzed the graph to get the results, just like the general attack graph analysis approach. But we can now include attributes about user behavior(s) into the relationship model to create the relationship graph. The relationship graph generated from this approach allow us to apply several mathematical methods not only from attack graph analysis, but also from both SNA as well. The results which we can classified by its length, help the organizational to develop appropriate IT policy and actions to handle it specifically. IT security auditor may also use this approach as a tool in risk assessment process to help develop risk findings and recommendations for risk mitigation for the organization as well.

However, we considered only risk scores and shortest attack path length of the graph. We have not include relationship scores in the analysis just to make it less complicated to preliminary test the proposed approach. Since it is possible to find more than one shortest path between two vertices. Finding shortest path using relationship score can help finding exact attack path.

Since we use shortest path analysis which helps us determine the organization risk in some level, it may not enough for complex social network containing many attributes. Analyzing relationship model with other mathematical methods such as Probability Transition Matrix and Markov Chain analysis that can handle more attributes with more accurate in its analysis.

There should be the study to compare mathematical methods using this proposed approach. For example, comparing walk path using relationship score, Markov Chain and Bayesian network. These mathematical methods could help finding more accurate attack path.

Because there are different information sources available in organizations. Analyst should carefully select source(s) that regularly available and easy to be transformed into graph data. Selecting information sources is important when automating the process of this approach to quickly calculate organization risk on time. We suggest analyst customize process according to the organization information sources, management requirements, and security policy of the organization.

REFERENCES

[1] P. Ammann, D. Wijesekera, and S. Kaushik, "Scalable, Graph-based Network Vulnerability Analysis," in *Proceedings of the 9th ACM Conference on Computer and Communications Security*, New York, NY, USA, 2002, pp. 217–224.

[2] R. Dantu, K. Loper, and P. Kolan, "Risk management using behavior based attack graphs," in *International Conference on Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004*, Las Vegas, 2004, vol. 1, pp. 445–449 Vol.1.

[3] S. Jha, O. Sheyner, and J. Wing, "Two formal analyses of attack graphs," in *15th IEEE Computer Security Foundations Workshop, 2002. Proceedings*, Nova Scotia, Canada, 2002, pp. 49–63.

[4] V. Mehta, C. Bartzis, H. Zhu, E. Clarke, and J. Wing, "Ranking Attack Graphs," in *Proceedings of the 9th International Conference on Recent Advances in Intrusion Detection*, Hamburg, Germany, 2006, pp. 127–144.

[5] O. Sheyner, J. Haines, S. Jha, R. Lippmann, and J. M. Wing, "Automated generation and analysis of attack graphs," in *2002 IEEE Symposium on Security and Privacy, 2002. Proceedings*, 2002, pp. 273–284.

[6] O. Sheyner and J. Wing, "Tools for Generating and Analyzing Attack Graphs," in *Proceedings of Formal Methods for Components and Objects, Lecture Notes in Computer Science*, Leiden, The Netherlands, 2003, pp. 344–371.

[7] C. Williams, R. Bhaumik, R. Burke, and Mobasher, Bamshad, "The Impact of Attack Profile Classification on the Robustness of Collaborative Recommendation," presented at the Proceedings of the 2006 WebKDD Workshop, Philadelphia, Pennsylvania, 2006.

[8] C. Phillips and L. P. Swiler, "A graph-based system for network-vulnerability analysis," in *in Proceedings of the 1998 workshop on New security paradigms*, 1998, pp. 71–79.

[9] P. Zhang, L. Min, L. Hopkins, B. Fardanesh, P. C. Patro, J. Useldinger, M. Graham, and D. Ramsay, "Utility application experience of Probabilistic Risk Assessment method," in *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, Seattle, Washington, 2009, pp. 1–7.

[10] T. W. K. Daniel, H. M.h, LAM S. T, MOK Y. C, OEI W. C, T. K.l, and Y. X.l, "Education in IT Security: A Case Study in Banking Industry," *GSTF J. Comput. JoC*, vol. 3, no. 3, Aug. 2014.

[11] M. E. J. Newman, "The structure and function of complex networks," *Soc. Ind. Appl. Math.*, vol. 45, no. 2, p. 90, Mar. 2003.

[12] O. Sheyner and J. Wing, "Tools for Generating and Analyzing Attack Graphs," in *Formal Methods for Components and Objects*, F. S. de Boer, M. M. Bonsangue, S. Graf, and W.-P. de Roever, Eds. Springer Berlin Heidelberg, 2003, pp. 344–371.

[13] C. L. Smith, *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners*. National Information Service, NASA HQ: CreateSpace Independent Publishing Platform, 2012.

[14] B. B. Madan, K. Goševa-Popstojanova, K. Vaidyanathan, and K. S. Trivedi, "A method for modeling and quantifying the security attributes of intrusion tolerant systems," *Perform. Eval.*, vol. 56, no. 1–4, pp. 167–186, Mar. 2004.

[15] J. L. Moreno, *Who shall survive? A New Approach to the Problem of Human Interrelations*. Washington, DC: Nervous and Mental Disease Publishing Company, 1934.

[16] F. Harary, *Graph Theory*. University of Michigan, IL: Addison-Wesley Publishing Company, 1969.

[17] M. E. J. Newman, S. Forrest, and J. Balthrop, "Email networks and the spread of computer viruses," *Phys. Rev. E*, vol. 66, no. 3, p. 035101, Sep. 2002.

[18] H. Ebel, L.-I. Mielsch, and S. Bornholdt, "Scale-free topology of e-mail networks," *ArXivcond-Mat0201476*, Jan. 2002.

[19] J. R. Tyler, D. M. Wilkinson, and B. A. Huberman, "E-Mail as Spectroscopy: Automated Discovery of Community Structure within Organizations," *Inf. Soc.*, vol. 21, no. 2, pp. 143–153, Apr. 2005.

[20] A.-L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.

[21] X. Ou, W. F. Boyer, and M. A. McQueen, "A Scalable Approach to Attack Graph Generation," in *Proceedings of the 13th ACM Conference on Computer and Communications Security*, New York, NY, USA, 2006, pp. 336–345.

[22] S. Kanchanapokin and S. Boonkrong, "Exploring Bad Behaviors from Email Logs," in *Information Science and Applications*, K. J. Kim, Ed. Pattaya, Cholburi, Thailand: Springer Berlin Heidelberg, 2015, pp. 517–524.

Supachai Kanchanapokin is a Ph.D. student of Information Technology Department at King Mongkut university of Technology, North Bangkok. He received his M.Sc. in Information Technology from the Department of Computer Engineering at Kasetsart University. His research interests are in information system security, social network study, IT policy and strategy management, and IT auditing.

Previously, Supachai was an IT professional consultant of several international IT consulting firms, providing IT services such as system implementation, IT strategy and policy development, IT service and operation management. He is now senior director of infrastructure and facility management division at a national research organization in Thailand.

Sirapat Boonkrong is an associate professor and an associate dean of academic and research affairs at the Faculty of Information Technology, King Mongkut's University of Technology North Bangkok (KMUTNB), Thailand. He received his B.Sc. and Ph.D. in Computer Science from the Department of Computer Science at the University of Bath, UK. His main area of research is information and network security.

Previously, Sirapat worked as a researcher at National Electronics and Computer Technology Center (NECTEC) in Thailand. He also has experience in industry as a project manager at an IBM-partnered company. He is currently a full-time lecturer at the Faculty of Information Technology, KMUTNB and is also supervising several Ph.D. students all of whom are in the field of information and network security.