Innovations in Computing Technology and Applications
Volume 2 | Year: 2017 | ISBN: 978-967-0194-95-0

Threat Actors Profiling Towards Targeted Critical Infrastructures By Adapting The Grounded-Theory Approach Framework

Mohamad Syahir bin Abdullah, Anazida binti Zainal

Department of Computer Science, Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia

mohamadsyahir94@gmail.com, anazida@utm.my

Abstract— A critical infrastructure is one of the most vital elements that help in generating and increase the growth rate of the nations' economic. All of the critical sectors play an important role not only in economic areas but also in other essential areas for a nation such as the national security, telecommunication and more. Nowadays, all these sectors have connected their system with the Internet network to comply with all the increasing demands and compete with their competitors. However, this situation has open up the possibilities and increases the risk for their system's vulnerabilities to be exploited by the cyber attackers. There are some increasing numbers that have been reported regarding the attackers' activities towards the critical sectors from time to time. The loss that included a lot of money and properties has been recorded more often. A nation that has or still not been affected need to take some countermeasure to solve this problem. Therefore, by implementing Cyber Threat Intelligence (CTI) analysis, we can classify the attacks that have been done by the threat actors so that we can see the trends and pattern of their activities. Moreover, this research used the existing framework that will be adapted and modified to suit the actors' characteristic as this research will focussing more onto the external actors before we can profile all the threat actors.
Keywords- *component; threat actor; critical infrastructure; profiling; framework*

1. Introduction

Threat actors are known as an entity that can cause havoc in the Information Technology (IT) era and possess a huge threat to the entire cyber world itself. Many forms of attacks have been planned and executed by these actors throughout the years. They have been involved in various cybercrimes and campaigns with many different agendas. One the reason they able to act freely is because there are no particular law to restraint them from doing so [1]. The threat actors can be differentiating into several types. One of the most commonly recognized actors is the group or organization of hackers known as the hacktivist. The hacktivist groups come with different names and agendas. For example, one of them is the infamous Anonymous. The hacktivist are known in helping advance political causes, inveigh against the corporate domination of telecommunications and mass media, the rapid expansion of dataveillance, and the hegemonic intrusion of the "consumer culture" into the private lives of average citizens [2]. Driven by various reasons that they considered it important and suit their motives, they have been terrorising the cyber world for some time.

In this few years, a lot of loss involving time, money and resources has been recorded. One of the most affected parts by the hacktivist is the critical infrastructure sectors. Critical infrastructures such as the telecommunication sector, health care sector, energy sector and other sectors play an important role for generating income for a country. However,

"Complicating matters is the fact that most critical infrastructures are in private hands"[3]. It somehow quite difficult to monitor all the sectors together thus increases the chance of the sectors to be exploited by the irresponsible people. For example, United State of America alone has almost 16 critical infrastructure sectors whereby each sector has its own organization and ways to handle the cyber threat problems. This number show how big the system is and how it is not dependent to each other. Contrary, the critical infrastructures of our technological society require proper operation of interconnected systems of systems to ensure our way of life runs smoothly and safely. Therefore, to ensure that all vulnerabilities and threats are able to be secured, a certain way is needed to monitor and record all kind of activities or attacks done by the malicious actors.

2. Literature Review

The urge to have a profiling for the threat actors is quite common in the criminal justice system. To identify a certain actor, some medium that enable to assess their related information is needed. In this case, the research is using a profiling method to profile the threat that they possess.

2.1. Security Concepts and Relationships

A threat profile can consist of information regarding the threat actors, the critical assets, and the threat scenarios [4]. By adding a third dimension that include an assessing advance persistent threats (APT) which will be useful to recognize the relation between the threat scenarios and its whole campaign.

Identifying the risk and all its relation will be needed later when the research wants to profile the threat actors. Having to know this relationship will let the affected organization to organize the information of threat into a more standardizes format [4]

2.2. Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE)

A method of information security risk evaluation that help an organization to make an information-protection decision regarding the risk to their critical assets that based on confidentiality, integrity, and availability [5]. OCTAVE focusing on the cyber-attack, thus will discussing only the relation that relate with the network access and not with the physical access.

2.3. Grounded-Theory Approach Framework

There are four main areas which are the Catalyst (reason or motive), Actor Characteristic (potential threat actor), Attack Characteristic (type of attack), and Organization Characteristic (affected area). This framework serves as a good platform in distinguishing the threat where it showed to us a proper relationship among the areas and help in profiling the attacks easier. However, as this research will focusing more towards the external actors.

3. Research Methodology

The research framework comprises of three phases comply with all three research objectives. This work consists of various steps to construct it.

Phase 1 : Modification of Attributes from the Existing Framework

The purpose of this phase is to identified the related framework to be used in threat actor classification and to identified the suitable attributes to be implemented in the modified framework. The research started by identifying the existing framework and study that framework to gain a better understanding. Previously in Chapter 2, three frameworks have been reviewed which are the Security Concepts and Relationships framework [4], OCTAVE framework [5] and the Grounded-Theory Approach framework (Nurse et al., 2014). The output produced from this phase is that a suitable framework for profiling the threat actors has been identified which is the Grounded-Theory Approach framework.

The chosen framework will later be adapted and undergo a modification to the suit the actors' characteristic as the previous framework only profiles the insider actors thus, some characteristic from the insider actors need to be change in this research. The suitable attributes gained from the literature review are listed and compared together. The attributes then selected after going through the selection process.

Phase 2 : Process data using clustering technique

Phase 2 aim is to manage the dataset obtained and cluster the data using a suitable clustering technique. Firstly; the dataset are modified where the data is being changed from some abbreviated term to a full word and filter the unneeded data that is not include in the research scope. This process produced a clean data to be used in the clustering.

To decide which clustering is the best, the research studied all the suitable clustering techniques available to decide whether it is k-means or fuzzy c-means algorithms to be used in this research. All the clustering technique are tested using R with the data that have been cleaned earlier. After the best clustering technique has been identified which is the EM clustering technique, process to cluster the data or the activities of the threat actors towards the critical infrastructures is started. The clustered data attributes are justified accordingly to the related framework as stated in Phase 1.

Phase 3 : Adapting the Framework and Develop the Dashboard

To achieve the final objective, phase 3 has adapted the attributes that have been selected into the modified framework. Also, in the final phase, the research has developed a dashboard to properly display the results obtained. Using the specified Tableau software, a dashboard is created that will show the profile of the attackers. The results of this research will be display by using a visualization format to let other people understand the result easily. The graphical presentation will able to grasp the attention of the people. Therefore, in this research, the designed dashboard create with the Tableau software is used to show the pattern and trend for the activities of the threat actors towards the critical infrastructures.

4. Research Design and Implmentation

To obtain the suitable attributes that will be used to modified framework for Characterizing Insider Attacks Based on Grounded-Theory Approach. There are several solutions proposed to complete the selection of the attributes which are through the literature review and from doing the clustering on the datasets.

4.1. Attributes Selection from Literature Review

The attributes are gathered from the literature reviews such as from the original framework, threat actor profile and OCTAVE. The comparison between all the selected attributes are as follows:

Table 1: Comparison of the Attributes

	Original framework	Threat Actor profile	OCTAVE profile
Catalyst	Participating event	Motive	Motive
Actors Characteristic	Personality characteristic	Name	Actor
	Historical behaviour	Relationship	
	Phycological state	Capability	
	Motivation to attack	Intent	
	Attitude towards work	Description	
	Skill set	ID	
	Opportunity		
Attack Characteristic	Attack	Action	Access
	Attack objectives	Outcome	Outcome
	Attack steps		
	Attack goal		
Organization Characteristic	Assets	Targeted victim	Assets
	Vulnerabilities	Targeted Asset	
		Region of operation	

4.2. Obtaining Attributes through Clustering Technique

To avoid confusion when accessing or using the data, the incomplete or abbreviated value is being modified. There are two attributes that contain values that have been abbreviated. These two attributes are Attack Class and Country. For Attack Class attributes, the values are short formed to CC, H, CW, and CE while for Country attributes, all the value are representing the actual country code set by the International Organization for Standardization (ISO) where the name of the country is abbreviated into two or three letters (wordatlas.com,2016). All the abbreviated value will be modified to its actual word by using the feature of 'find and replace' in Microsoft Excel.

Data is also filtered to remove unnecessary data that not in this research scope. In order to solve the problem where some data is not a part of the research scope, the unneeded data will be filtered. To make the data filtering easier, this research has used RapidMiner Studio tool. There are two attributes that contain some unneeded value which are in Author and Target Class. For Author attribute, this research mainly covered about the external actors only such as Hacktivist and Nation-founded hacker groups while for the Target Class attribute, this research is focusing on industry or organization that part of national critical infrastructures. Thus, using RapidMiner Studio, only the required Author and Target Class data are identified and stored in the new Microsoft Excel file.

Clustering is used to cluster the data of the threat actors' activities towards the critical infrastructures. The number of clusters obtained can justify whether the attributes from the datasets are relevant and have some relationship to the framework or not. To complete this

process, this research has used several clustering processes in order to get the best result such as k-means and Expectation-Maximization (EM) techniques. As the data used are in categorical form, it is appropriate to change it into numerical value. Using the *transform()* function, all the values from all the stated attributes are being assigned into different ID as R can process numerical value better than categorical value. After the labelling process is completed, some of the chosen attributes will be used to justify its importance and relation to the framework.

Next, this research will plot the three attributes using the *mclust()* function that follows the EM clustering principal. It will read from the file and cluster the result in four clusters. We also want it to show the plot in graph form and the summary of the clustering process. The following code has been implemented:

```
"library(mclust)        fit <- Mclust(test2,4) plot(fit)
summary(fit)"
```

From the Mclust() function, there are some different representation form that can be observed which are through density, classification, BIC, and uncertainty.

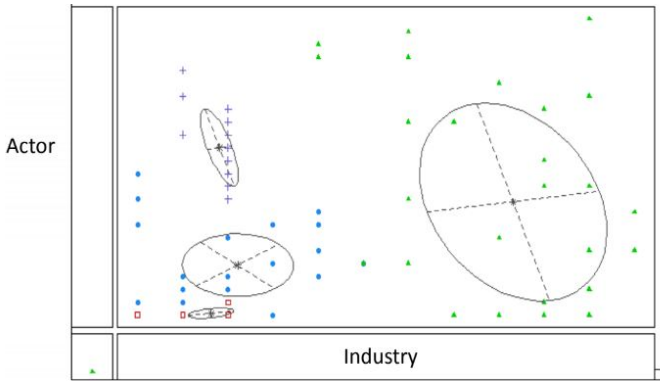

Figure.1. EM Clustering

It can be said that the four clusters can represent the four motives or Attack Class which are hacktivism, cybercrime, cyber warfare, and cyber espionage. In the modified framework, these four motives can be considered in Catalyst phase where it shows the initial intention for the actors to attack the targeted industry. From the figure 4.5, the plot for ActorID and IndustryID does show the distinctive groups that can be defined as certain actors are focusing only on certain industry which is true considering that some actor like Anonymous that mostly act on hacktivism motive that mostly focusing on Government sector of a country. The clusters do suit the behavior of the threat actors towards the industry. However, in most cases, the actors might use different methods to initiate an attack as can be seen on the plot for AttackID and ActorID which has some group of clusters that overlap on each other's. Meanwhile in the summary of the cluster, the number of incidents in each cluster can be seen which has quite similarity to the actual numbers in the Attack Class attribute.

5. Result, Analysis And Discussion

Based on the original framework, there are four main areas which are Catalyst, Actor Characteristic, Attack Characteristic, and Organization Characteristic. To suit all the attributes into all the areas, the attributes that have been selected will be reorganized and some redundant attributes will be dropped or merge before being implemented in the respective areas. Some of the areas do have unbalanced number of attributes due to different ways of characterising the threat actor's activities towards the critical infrastructure and there are also some attributes that bring the same meaning but only with different naming such as in the Catalyst area, it is intended for the factor that trigger the attack (Nurse et al., 2014)

which share the same meaning as Motive and Attack Class attributes. So, this kind of attributes are merged together. While in the Actor Characteristic area, some attributes such as Attitude Towards Work will be dropped as it does not meet the scope of the research. Therefore, the final attributes that can be concluded from the previous table to be implemented in the new modified framework are as follows:

Table.2. Selected Attributes

Area	Attributes
Catalyst	Motive
Actors Characteristic	Actor Name
	Date
Attack Characteristic	Attack
	Description
	ID
Organization Characteristic	Targeted Asset
	Industry
	Country

The modified framework is re designed and can be seen in the following figure:

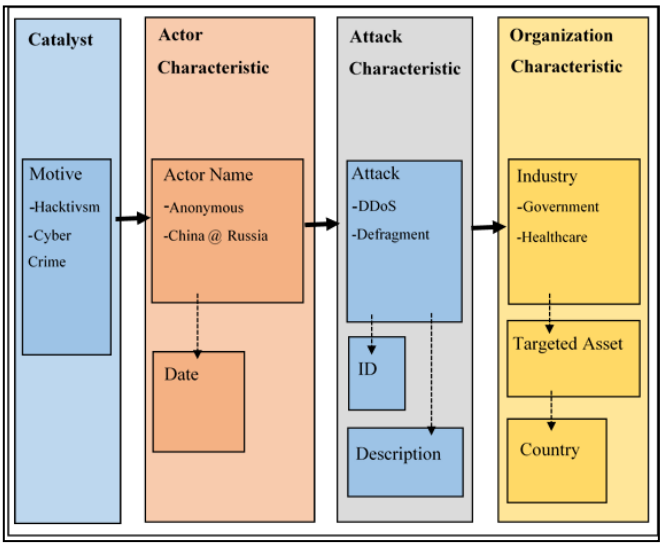

Figure 2. Modified Framework

The framework explains that the Catalyst or the Motive is the initial reason that contribute to the certain actor in the Actor Characteristic to take malicious action or start

the cyber-attack, as in Attack Characteristic towards the intended sector or industry of the critical infrastructure.

Then, using Tableau, many interactive ways to represent the data can be achieve and it also can connect to previous software that this research is using which is R Studio Tableau can provide much cleaner visual to display the result obtained from the processed data. Figure below show the overall visualization:

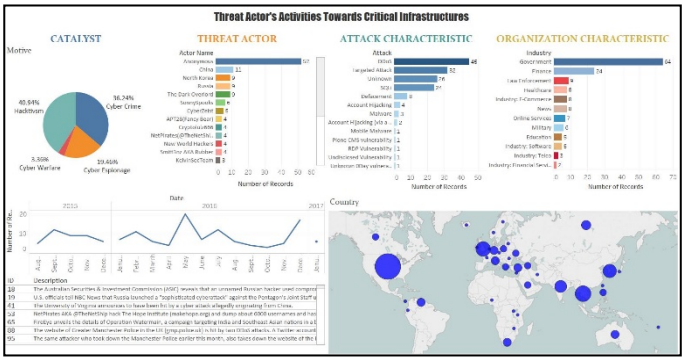

Figure 3. Threat Actor Profiles

The above row of the dashboard shows the main areas inside the framework which are Catalyst, Threat Actor, Attack Characteristic and Organization Characteristic. The bottom part of the dashboard display the Date of the incidents happen, ID and Description of the attack and finally the Country affected by the attack. One of the advantages of Tableau is that each of the component are responsive to each other as when a component is selected, it will link to other related information inside the dashboard and the others component will responsively show the required information. Using this feature, the profile of the attacker can be easily shown in the following figure.

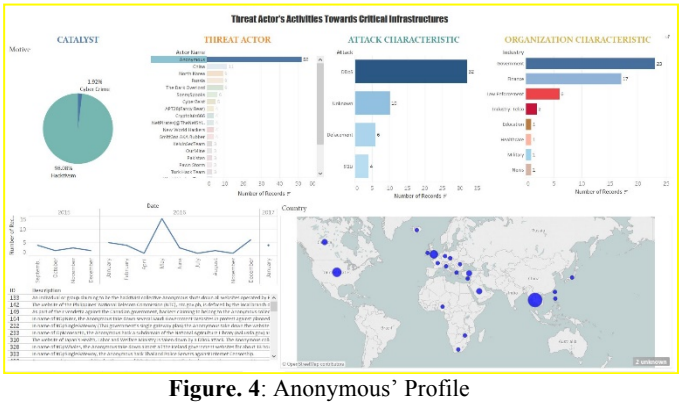

Figure. 4: Anonymous' Profile

For Anonymous, Hacktivsm serve as a major catalyst for them to initiate a certain attack which suits them as they are mostly known for their hacktivist group. Then, the most frequent attack technique used by Anonymous is DDoS while the three most attacked Industry are Government, Finance and followed by Law Enforcement. From the dashboard, the trend line of the actor's activities throughout the year can also be observed where Anonymous most active period is during May 2016. The details information about the attacks that have happened can be read in Description pane which also including the ID of

the attack. Finally, the last information that can be achieved from this dashboard is the Country affected from the action of this threat actor. From the world map, it shows that the Anonymous operated on most of the country in this world where some of the activities are more centered in countries across the Europe.

6. Conclusion

From the dashboard, lot of thing can be learned and observed such as the motive of the attacks, the methods used by the threat actors, the industry involved, country affected, trends or timeline of the attacks and description of the incidents. The graphical representation let other people with non-technical knowledge to understand the characteristics and behavior of the threat actors easier thus enable them to take precaution steps to avoid or prevent from such incidents to happen in their sectors of the critical infrastructures. Thus, this dashboard should provide important insights in keeping the industries in a secure cyber environment. However, the accuracy of the analysis still mostly depending on the data collected. Thus, the source of the data should be reliable to assist in making a good analysis.

References

[1] P. Ray, G. Stephens, L. Kwan, P. Ray, and G. Stephens, “Towards a Methodology for Profiling Cyber Criminals Towards a Methodology for Profiling Cyber Criminals,” no. January 2008, pp. 1–9, 2016.

[2] M. Manion and A. Goodrum, “Terrorism or civil disobedience,” ACM SIGCAS Comput. Soc., vol. 30, no. 2, pp. 14–19, 2000.

[3] K. Geers, “The Cyber Threat to National Critical Infrastructures: Beyond Theory,” Inf. Secur. J. A Glob. Perspect., vol. 18, no. 1, pp. 1–7, 2009.

[4] S. Irwin and S. Northcutt, “InfoSec Reading Room Creating a Threat Profile for Your Organization Creating a Threat Profile for Your Organization,” Creat. a Threat Profile Your Organ., p. 32, 2014.

C. Alberts and A. Dorofee, “Octave Sm * Threat Profiles,” pp. 1–14.