UTILIZING PAST EXPERIENCES OF INCIDENT HANDLERS FOR REALIZING A CBR RECOMMENDER IN IT SECURITY INCIDENT RESPONSE

Wira Zanoramy A. Zakaria, Kilausuria Abdullah and Faiszatulnasro Mohd Maksom
MyCERT, Cybersecurity Malaysia, Level 7, SAPURA@MINES
Jalan Tasik, The Mines Resort City
43300 Seri Kembangan, Selangor, Malaysia
{wira, suria, fais}@cybersecurity.my

ABSTRACT

Incident response is a very important subject in IT security. Due to significant rise in the number of total reported incidents, there is a need for an intelligent based recommender system to assist the Incident Handlers (IH) in responding to cyber threats. This work explores the application of the Case-based Reasoning (CBR) methodology in order to develop a CBR recommender system for assisting IH in handling and responding to cyber security incidents. The architecture of the proposed system and the work done on case representation describing some sample cyber incident category are discussed in this paper.

Keywords: *CBR system, recommender system, case representation, incident response*

1. Introduction

Based on Malaysia CERT (MyCERT) yearly statistics, there are a total of 11918 reported incidents for the year 2014, 10636 cases reported in 2013 and 9986 cases reported in 2012 ("MyCERT Incident Statistics," 2015). From this statistics, it is clear that the number of cyber incident is increasing from year to year. This is due to many reasons: the expansion of the number of Internet users, increased awareness among Malaysian IT users in reporting cyber incidents to MyCERT, the increasing number of threats and cyber criminals in the cyberspace and so on. These incidents are classified into many categories such as cyber harassment, denial of service (DoS), online fraud, intrusion, malware and spam. These reported incidents came from many entities, including individuals, companies, organizations and government agencies.

Incident Handlers are main asset at any security departments. Depends on the organization, usually they are divided into three categories based on their skills and experience - 1st level, 2nd level and 3rd level incident handlers. The 1st level handlers are mostly juniors in incident response. The 2nd levels are the staff with intermediate skills of incident response and each of them basically have a bit of specific skills in certain types of incidents for example phishing and malware outbreak. The 3rd levels are those who have very deep knowledge and experience in all of the cyber threats. At MyCERT, all incident handlers regardless at what level they are, work hand in hand in order to resolve any reported incident.

This paper is organized in the following order: Section 2 described on the problem statement of this work. Section 3 descibed about the research outcome. Section 4 discussed on the general introduction to case-based reasoning. Section 5 defined the proposed system. Section 6 discussed on the conclusion of this work and listed some potential future effort for this research.

2. Problem Statement

Based on MyCERT experience, listed below are a few challenges in handling cyber incidents:

a) The large number of reported incidents especially in cyber crisis period could become a burden for Incident Handlers to response promptly.
b) The ever-changing trends of cyber attacks.
c) Time needed for incident response might depends heavily on the experience and skills. Senior Incident Handlers more likely to response faster than their juniors.
d) Security staffs that are new to incident handling have difficulties to response to cyber incident. New Incident Handlers usually depends on the training and guidance provided by the senior Incident Handlers.
e) There is a gap in centralizing and storing all the valuable experiences of the Incident Handlers. It could become a big issue when the seasoned Incident Handler left the department or change their job scope.

3. Research Outcome

To overcome the shortcomings described in the previous section, there is a need for an intelligent system that can store, retrieve and learn in the domain of security incident response and handling. The proposed system will be able to act as a medium for information sharing in the domain of incident response and promotes efficient incident resolution dissemination among the incident handlers.

Within this research, valuable successful experiences from the incident handlers were retrieved and represented in the form of formatted cases. Important features of the domain of incident handling are identified and it is utilized to build a usable knowledge base or to be precise an incident response case storage.

4. Case-Based Reasoning

Case-based reasoning (CBR) is one of artificial intelligence technique that is mostly being implemented in intelligent systems. CBR falls under the subtheme of machine learning. It is a problem solving approach that makes use of past experiences to resolve the current problem (Kolodner, 1992; Aamodt & Plaza, 1994; Supic, 2012). Past experiences in the particular domain are extracted and it is represented in computer comprehensible format or usually known as a case. The case is an information structure that contains two parts: Problem and Solution (Qu, 2002).

This cases which stores the description of a previous problem coupled with the respective solution, is grouped, indexed and retain within a case storage. Every time the CBR recommender is fed with new problem, it will retrieve the case storage for the matching previous case, extract the solution described within the selected case and reuse the solution for solving the new problem.

As described by CBR cycle in Figure 1 below, in any CBR system, it involves the 4R steps: Retrieve, Reuse, Revise and Retain. In the case retrieval step, matching case(s) relative to the new problem is selected from the case storage. The solution part of the selected case is used in the case reuse step.

If needed, the solution provided by the selected case is modified in the case revise step, in order to fit the effort in solving the new problem. Finally, a new case is created from the information provided by the new problem, and it is saved into the case storage in the case retain step. By the retain step, it is said that CBR system learns a new experience in the respected domain. The case storage of the CBR system or in other words, the experience of the system expands as it solves more problems from time to time.

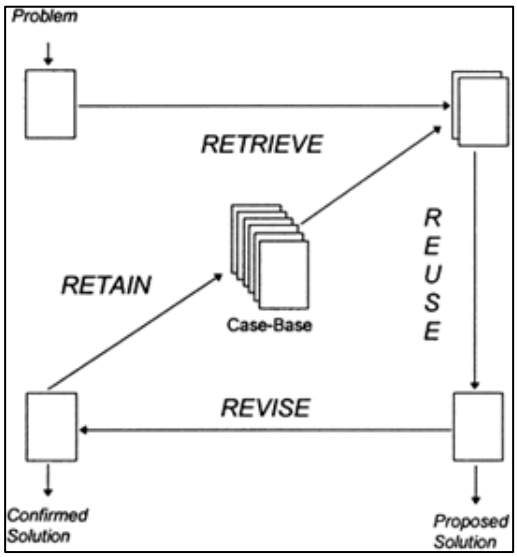

Figure 1: CBR cycle

5. CBR Recommender in Cyber Incident Response

In this research, we proposed a CBR recommender system for assisting IT security Incident Handlers or any CERTs. A CBR recommender system is an intelligent system that is able to make use of previous experiences stored in the form of cases, to derive a computer generated solution recommendation in order to solve a problem. An initial finding of this research also has been described in Wira (2015). In the paper, we discussed about the motivations that drives this research. The knowledge base of this system contains extracted and systematically formatted past experiences of successful incident handling and response. The experiences are represented in the form of indexed cases, which later can be used by the CBR recommender. Figure 2 below shows the flow of the proposed system.

The CBR recommender is feed with a new reported cyber incident. The reported incident is captured and treated as a new case by the CBR system. Then, the system take the new case an runs the 4R steps until finally it produces a list of recommended action and response for that particular cyber incident. This output is used as a solution to resolve the problem described by the new case.

Figure 2: Flow of the proposed system.

5.1 Case Representation for Cyber Incident

To build any CBR system, the most crucial phase is the case representation (Supic, 2012). This part of the work involved selecting the right attributes or features that systematically describe the domain of incident response. As mentioned in Section 4, a case consists of two segments: Problem and Solution. The Problem section for this domain contains the description of the reported incident. The Solution section contains the corresponding action taken by the Incident Handler for that particular reported incident described by the Problem section.

For the domain of cyber incident handling and response, the case representation is built using the attribute-value approach. Each attribute declared in the case is assigned with a particular value either numerical or phrases. Table 1 – 3 below shows the complete structure of the incident response cases. There are a few incident categories taken into consideration in this research – phishing, malware and web defacement incidents.

Case ID: 001/Date/Time **Incident Type: Phishing** • Phishing_URL • Domain_IP • Phishing_email (full email header)
Respond: • check phishing site alive/down • get information from phishtank • analyze full email header to get source of originating IP address • notify to admin, technical and hosting contact

Table 1: Case structure for describing and resolving a phishing incident

Case ID: 002/Date/Time **Incident Type: Malicious document** • Malicious_site_URL • Malicious_document_name • IP_infected • IP_suspected • Log
Respond: • check malicious site alive or down • get information from virustotal • analyze full email header to get source of originating IP address • notify to admin, technical and hosting contact

Table 2: Case structure for describing and resolving a malware related incident

Case ID: 003/Date/Time **Incident Type: Intrusion > Web Defacement** • Defaced_URL • domain IPs /admin contact/technical contact/hosting contact • Domain_server_name • Log
Respond: • check defaced URLs site alive/down • notify to admin, technical and hosting contact

Table 3: Case structure for describing and resolving a web defacement incident

6. Conclusion and Future Work

CBR is proven to work with many other domains, including intrusion detection and spam. From the work done this far, we found out that CBR is also a valid methodology for the domain of incident handling and response. The abundance of past incident data equipped with highly experienced IH talents at MyCERT, assisted a lot in realizing this research work. For the next step of this research, we plan to have this work to be implemented in open source CBR development tool, to build the complete ontology for this domain, to further this research with defining the 4R steps and finally to test this system in the real CERT environment.

References

MyCERT Incident Statistics. (2015). Retrieved from www.mycert.org.my/statistics/2015.php

Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence Review, 6*(1), pp. 3 - 34.

Aamodt, A., Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *Artificial Intelligence Communications, 7*(1), pp. 39-59.

Supic, H. (2012). Representation of cases in group recommender systems by combining users' perceived feature importance weights. *International Conference on E-Learning and E-Technologies in Education (ICEEE) 2012*, pp. 214–218.

Wira, Z. A. (2015). Application of Case Based Reasoning in IT Security Incident Response. *3rd International Conference on Recent Trends in Engineering and Technology (ICRET) 2015.*