NECOMAtter: Curating Approach for Sharing Cyber Threat Information

Takuji Iimura*, Daisuke Miyamoto*, Hajime Tazaki*, Youki Kadobayashi°
*The University of Tokyo, Japan ° Nara Institute of Science and Technology, Japan

ABSTRACT

In this paper, we design and implement a novel system for connecting cyber threat information. The objective is to improve the information and its analysis results with machine intelligence assisted by human intelligence. This paper illustrates the system named NECOMAtter based on these assumptions, and summarizes our contributions in order to develop actionable knowledge.

Categories and Subject Descriptors

D.2.2 [**Software Engineering**]: Design Tools and Techniques

General Terms

Security

Keywords

Curation service, Cyber threat information, Web

1. INTRODUCTION

Analyzing data and deriving threat insights are becoming far more important than ever, which requires collective analysis among experts with diverse background. A lot of research has been carried out in the anomaly detection and detection of specific event, although very few attempt has been made to assist "connecting the dots".

We argue that existing effort on machine intelligence should be assisted by human intelligence, and vice versa. There is critical lack of systems-oriented research here: how are we going to deliver findings from machine intelligence to human operators, and how human operators, or possibly another machine intelligence, connect these dots?

It is worth noting that there might have been simplistic assumptions on how we are going to construct systems out of multiple detection algorithms and analysis algorithms: operators interact with each security appliance through web interface, and interactions among multiple security appliances should be done through Security Information and Event Management, which is supposed to connect the dots and somehow derive insights on threats.

NECOMAtter tries to refactor the interactions among multiple analysts, multiple algorithms and multiple datasets by revisiting the challenge from systems perspective and making entire interactions more compositional, as the formations of campaigns and threats are also highly compositional.

It is also worth noting that we should not be tricked by the fallacy of universal single detector – as threat landscape rapidly evolves, the effectiveness of particular detector will change, and as with any complex problem in the real world, any attempt to build a single algorithm that works universally against arbitrary combination of phenomena will fail.

We thus aim to build a system that encourages analysts to improvise on the analytics according to rapidly developing situation – by encouraging analysts to collaborate with other analysts and algorithms in problem-specific manner. We argue that this improvisation of machine intelligence and human intelligence has not been seriously considered in modern cybersecurity context, and that such improvisations can be assisted by refactoring the way algorithms and heuristics are packaged.

CFI '14, Jun 18-20 2014, Tokyo, Japan.
ACM 978-1-4503-2942-2/14/06.
http://dx.doi.org/10.1145/2619287.2619306.

2. DESIGN OF NECOMATTER

In this section, we propose the design of NECOMAtter, a system for connecting cyber threat information to uncover the causal connection of each information.

Our key idea is to develop a content curation system for the information. Content curation is defined as the act of discovering / collecting the web contents, and presenting them into a context of with organization and annotation. In the case of Twitter [3], curators can use web curation services such Chipstory [1] and Togetter [2] to summarize posted articles. If we could deem various cybersecurity information as various tweets, development of curation services might be necessary to enrich cybersecurity information by summarizing and/or filtering cyber threat information.

2.1 Overview

In NECOMAtter, every cybersecurity information will be "tweeted" while the system refers the design of Twitter. Each tweet is read and written by NECOMAtter tweet bots as well as NECOMAtter users. The roles of tweet bots are posting cybersecurity information and reacting tweet message posted from other tweet bots. For example, a DoS detection system posts the attackers' source addresses, a bot

prevention system also tweets the related information to the reported addresses.

Our developed NECOMAtter have to equip the function of the curation. Each NECOMAtter user has own timeline, and reading other users' timelines therefore needs lot of time. Our developed curation tool named NECOMAtome facilitates collecting and sharing of tweets about a specific topic.

Figure 1 shows a screenshot of NECOMAtome, while a user is drag and dropping a couple of tweets (left side) to a new list (right side) to generate a summary of a specific event.

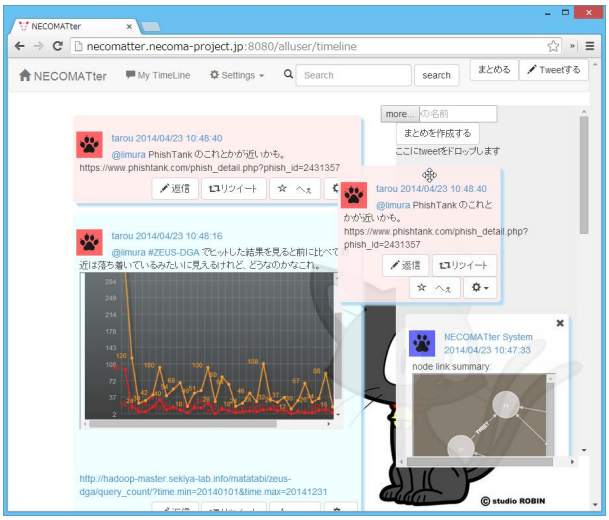

Figure 1: NECOMAtome

2.2 Accumulation of information

NECOMAtter is designed to store the summary of the information in the NECOMAtter system. Since we assumed that the size of information tends to be big, it is not easy to store all information in the system. On the other hand, storing no information is not feasible; in this situation, each information is retrieved whenever a NECOMAtter user requests regardless of various information sources. In comparison to the storing information, search-ability for information would be lowered. Our approach is hybrid, storing the summary of the information, and showing some pointers, e.g., URL links, for a user who requests detailed information.

For the data accumulation, NECOMAtter is not recognized data types. Each source provides each information, and it is formatted to own data format. We considered that NECOMAtter could not deal with all data formats including the newly developed data as well as the various existing data. Instead, NECOMAtter is designed to recognize all data as plain text format.

2.3 User interface

We also borrow the concept of timeline, a real-time list of tweets on Twitter. In the case of NECOMAtter, there are also various types of cyber threat information, and each NECOMAtter users have different demands for the information. For supporting the users to retrieve their suitable information along with the demands, NECOMAtter should equip the timeline function which enables to search by tags and to elect by subscribing other NECOMAtter users and/or curated groups of the users.

In addition to that, NECOMAtter timeline can be sorted in chronological order. On monitoring the ongoing events, the useful information tends to be the newest cyber threat information rather than the old one.

3. CONCLUSION

This study proposed NECOMAtter, a system for connecting each information to uncover the causal relationship of the information. We assumed that the difficult point is to develop actionable knowledge from various types, context, format and sources of the information. To deal with the problem, our study focused on utilizing the idea of the web curation services. While we deemed the various threat information as various posted article in web, the curation services would enrich the information.

In the NECOMAtter system, we speculate that many ideas from Twitter [3] system can be useful in cybersecurity context. Every cybersecurity information is tweeted by NECOMAtter users and/or tweet bots, and the users can create their own timeline. The tweet bots can react tweets to provide additional information toward connecting each information. The curation tool was implemented as a subsystem of NECOMAtter, named NECOMAtome, and all users can use the curation service in order to categorize cybersecurity information, share the information, filter unnecessary information, and highlight previously unseen relationship that extend. our understanding of cyber threats.

We also considered that NECOMAtome can provide a holistic view for understanding malicious campaigns. Assuming if there were tweet bots for DoS detection, malware prevention, spam detection, phishing detection, and other prevention services for cyber attacks. Within the curation service, the fragment of the cyber threat information would be aggregated in NECOMAtome. We expected that the holistic view plays an important role for extracting actionable knowledge from the threat information.

In our future work, we will assess our hypothesis by validating the effectiveness of our curating approach. As well as the validation, we will also implement many tweet bots to enrich the information, and consider the access control in regard to the context of the cyber threat information.

Acknowledgment

This research has been supported by the Strategic International Collaborative R&D Promotion Project of the Ministry of Internal Affairs and Communication, Japan, and by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 608533 (NECOMA). The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Ministry of Internal Affairs and Communications, Japan, or of the European Commission.

4. REFERENCES

[1] Togetter, Inc. Create stories from Tweets. - Chirpstory. Available at: `http://chirpstory.com/`.

[2] Togetter, Inc. Togetter. Available at: `http://togetter.com/`.

[3] Twitter, Inc. Twitter. Available at: `https://twitter.com/`.