The Pennsylvania State University

The Graduate School

A COGNITIVE PROCESS TRACING APPROACH TO CYBERSECURITY DATA TRIAGE OPERATIONS AUTOMATION

A Dissertation in

Information Sciences and Technology

by

Chen Zhong

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

August 2016

The dissertation of Chen Zhong was reviewed and approved* by the following:

Peng Liu
Professor of Information Sciences and Technology
Dissertation Co-Adviser, Chair of Committee

John Yen
Professor of Information Sciences and Technology
Dissertation Co-Adviser

John M. Carroll
Distinguished Professor of Information Sciences and Technology

David Hunter
Professor of Statistics

Robert F. Erbacher
Research Scientist at Army Research Laboratory
Special Member

Andrea Tapia
Associate Professor of Information Sciences and Technology
Director of Graduate Programs, College of Information Sciences and Technology

*Signatures are on file in the Graduate School.

Abstract

Security Operations Centers (SOCs) not only employ various cyber defense technologies to continually monitor and control network traffic, but also rely heavily on cybersecurity analysts to make sense of the resultant network monitoring data for attack detection and incident response. As the network monitoring data are usually generated at a rapid speed and contain a lot of noises, analysts are so far bounded by tedious and repetitive data triage tasks that they can hardly concentrate on in-depth analysis to generate timely and quality incident reports. These difficulties result in a great disparity in force between overwhelmed cybersecurity analysts and aggressive attackers. Therefore, there is an urgent need to liberate cybersecurity analysts from the tedious data analytics to focus on the higher-level cyber situational awareness.

Our work is aimed to reduce analysts' workloads by leveraging analysts' previous cognitive processes of data triage. With this goal in mind, a fundamental step is to trace analysts' cognitive processes at the fine-grained level while they are performing data triage tasks. We defined data triage as a dynamic cyber-human system and proposed a trace representation of a fine-grained data triage cognitive process. Based on the trace representation, we developed a tracing method which integrates automated capture and self-reports to capture an analyst's cognitive process with the minimum interference. An interactive toolkit, named ARSCA, was developed as a specific realization of the tracing method. In collaboration with Army Research Lab, an experiment on professional analysts was conducted in which participants' cognitive processes were traced by ARSCA while they were performing a simulated cybersecurity analytics task. The results of the experiment show that the proposed tracing method is a feasible way to conduct on-the-job cognitive task analysis studies with less influence on the human performance. Besides, the preliminary trace analysis indicates that the collected traces are abundant in

information and shed some insights into the cognitive process of how the participants perform data triage task.

To the best utilization of the captured traces in the experiment, we proposed an automated trace analysis method for constructing data triage rules directly from the traces. The rules were further used to build finite state machines for automated data triage. The biggest challenge in the rule construction for the data triage was how to distinguish the effective data triage operations from the exploratory ones in traces. To solve it, a graph model was proposed to represent both the temporal and logical relationships among analysts' data triage operations. We evaluated the automated data triage systems constructed from the traces by applying them to a large dataset and comparing the data triage results with the ground truth. The results of the study illustrate the rules mined from the traces are useful to conduct effective data triage, which further validates the practical value of the proposed tracing method. Besides, the results also show that the automated system built on the traces from the analysts with better task performance have a better data triage performance, which implies that a better data triage performance can be achieved by tracing and utilizing experts' operations.

In conclusion, an initial step had been taken towards leveraging human analysts' previous cognitive processes to facilitate data triage. Its contribution lies in three aspects. Firstly, the proposed tracing method realizes the possibility of tracing human analysts' cognitive processes in a less intrusive manner while analysts are performing cybersecurity analytics tasks. Secondly, the proposed trace analysis method was shown to be effective and useful in constructing useful data triage rules from the analysts' operation traces in a largely automated way. Thirdly, the constructed data triage rules can be used to construct data triage automation for reducing analysts' workloads.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my special appreciation and thanks to my advisers Dr.Peng Liu and Dr.John Yen. You have been tremendous mentors for me. I would like to thank you for your advice on both research as well as on my career. I can still clearly remember my first conversation with Dr.Liu over the phone five years ago, from which my journey of PhD started. During the past five years, I have been continuously encouraged and inspired by your research enthusiasm. I feel so lucky to have you as a role model as I'm growing as a researcher.

I would also like to thank my committee members, Dr. John M. Carroll and Dr.David Hunter and Dr. Robert F. Erbacher for serving as my committee members even at hardship. I also want to thank you for letting my defense be an enjoyable moment, and for your brilliant comments and suggestions, thanks to you.

I would like to give my special thanks to Dr.Robert F. Erbacher, who is my most important collaborator. His mentorship deepened my understanding in cognitive task analysis and cyber operations. The inspiring discussions with him have improved my research a lot.

I would like to thank every one of my collaborators: Deepak S. Kirubakaran, Renee Etoty, Christopher Garneau, Gaoyao Xiao, Steve Hutchinson, Hasan Cam, Sam Gur, Xiaoxiao Bai, William Glodek and Mingyi Zhao. Without any of them, I cannot make my way through this dissertation.

I would like to thank Dr.Nicklaus Giacobe for his endless advice on teaching when I was his teaching assistant.

I would like to thank all my labmates, colleagues and friends at Penn State. Jun Wang, Fang Dong, Bin Zhao, Tao Zhang, Jun Dai, Xiaoyan Sun, Mingyi Zhao and Gaoyao Xiao all offered me a great help, especially in my first year at Happy Valley. Thanks Kai Chen, Jun Dai for their advise on both life and career. Xin Peng has celebrated five new years together with me when we were away from home. Best wishes to Chi Pnig for her new life and career.

Finally, I give my deepest thanks to my mother Dr.Chunxia Xu and my father

Dr.Ming Zhong for their love, support and encouragement. Words cannot express how grateful I am. Thank you for everything and wish you the best.

Dedication

This dissertation is dedicated to my mother Dr.Chunxia Xu and my father Dr.Ming Zhong for their endless love, which gives me the ability to love others.

Chapter 1

Introduction

1.1 Background and Motivation

1.1.1 A Critical Role of Cybersecurity Analysts

The big breaches of Target, JPMorgan Chase, Sony Pictures, and Primera Blue Cross constantly remind us that the risk of cyber attacks can hardly be underestimated. Many prominent companies, government organizations and military departments have invested a lot of money to construct Security Operations Centers (SOCs) against the increasingly sophisticated cyber attacks. Typically, SOCs are cyber defense systems for 24*7 monitoring, intrusion detection, and diagnosis (on what is actually happening) [78]. In a military setting, CNDSP (Computer Network Defense Service Provider) centers have already been established and operating for quite a few years. SOCs usually employ multiple automated security measures, such as traffic monitors, firewalls, vulnerability scanners, and Intrusion Detection/Prevention System (IDS/IPS), all of which continually generate network monitoring data.

Nowadays SOCs rely heavily on human analysts (i.e., cybersecurity analysts) to make sense of these data to achieve cyber situational awareness (Cyber SA).

More specifically, the following questions need to be answered through analysts' cybersecurity analytics: Whether a network is under attack? How does attacks happen? What will attackers do next?

Analysts are playing a critical and indispensable role because the automated measures are in many cases unable to "comprehend" sophisticated cyber attack strategies even with advanced correlated diagnosis. Specifically, analysts need to conduct a series of analyses, including data triage, escalation analysis, correlation analysis, threat analysis, incident response and forensic analysis [17]. Each stage of analysis involves very complicated analytical reasoning performed by analysts with their knowledge and experience gained through years of on-the-job training. For example, data triage, as the fundamental step, encompasses examining the details of a variety of data sources (e.g., IDS alerts, firewall logs, OS audit trails, vulnerability reports, and packet dumps), weeding out the false positives, and grouping the related indicators so that different attack campaigns (i.e., attack plots) can be separated from each other. Data triage provides a basis for closer inspection in the subsequent analyses to finally generate confidence-bounded attack incident reports. These incident reports will serve as the primary basis for further decision-making regarding how to change current security configuration and how to act against the attacks.

1.1.2 Challenges Faced by Cybersecurity Analysts

Cybersecurity analytics presents the analysts with multiple challenges in today's SOCs. First of all, the network monitoring data, which are being continuously generated, are too overwhelming for analysts to process. Compared with a computer, human brains have smaller orders of magnitudes of data processing throughput, and human beings have such disadvantaged weaknesses as fatigue, anxiety and depression. However, neither the network nor the attack campaign is waiting for the human brains. In addition, the data coming from various sources contain many

false alerts so that analysts need to apply their domain knowledge and experience to make high quality decisions regarding which parts of the data are worth further investigation and what are suspected malicious events to report as an incident. According to a research by the Ponemon Institute in January 2015 [71], the average number of malware alerts an organization receives in a week is 17,000, and only 19% or less of them are reliable. About two-thirds of the time of cybersecurity analysts are wasted for investigating the false alerts. FireEye reported that the average annual operational spending of an organization due to false positives is $17.816 million given the default resource capacity informed through common deployments [28].

Secondly, analysts have to hasten their data triage under the pressure of limited time. Given the rapid influx of network monitoring data, analysts usually have to make quick decisions because earlier detection of cyber attacks is the premise of timely incident response. Besides, analysts have to rotate on shifts on a 24/7 schedule (24 hours a day, 7 days a week) in order to ensure uninterrupted monitoring. Especially for data triage, analysts need to make decisions within a very short period for filtering the incoming data to identify indicators for suspicious events, weeding out false alerts, generating hypotheses regarding malicious events, and investigating data from different sensors concerning the suspicious events.

Considering those challenges, several major companies and government organizations have adopted advanced cybersecurity analytics systems, such as Security Information and Event Management (SIEM) products, to support integrated cross-source data analysis and incident management. Although the SIEM systems take a big leap forward in cybersecurity analytics, these SIEM systems are extremely expensive not only for the high cost of its license and deployment but also for the large amount of time and expertise required in regular system management and customization. Due to SIEM's high cost, most organizations can't afford good protection of their networks, despite the potential perils and loss of the cyber attacks

to their networks.

1.2 My Approach

There is a big gap between the demands for enhancing the cybersecurity analytics capability of analysts and the limited resources of expert analysts and the extremely high cost of advanced cyber analytics systems. My research is motivated to bridge that gap. My dissertation work takes the first step to facilitate cybersecurity analytics by understanding and leveraging analysts' analytical reasoning processes. To hone in on a reasonable research scope, I mainly focus on data triage which is a fundamental but the most tedious and time-consuming stage in cybersecurity analytics.

My approach is built on three important insights. Firstly, it is possible to trace human analysts' cognitive processes in a less intrusive manner while they are performing analytics tasks. Secondly, useful information can be mined from the captured cognitive processes in a largely automated way. Thirdly, data triage automatons can be constructed to reduce analysts' workloads by utilizing the mining results.

1.2.1 Method for Tracing Analysts' Cognitive Processes

The first objective of my research is to gain understanding of human analysts' cognitive processes of data triage at a fine-grained level. A critical factor influencing the success of a SOC in tackling the cybersecurity analytics challenges is the effectiveness of the cybersecurity analysts' cognitive processes in data analysis tasks. However, the detailed cognitive process of data triage analysts is rather complicated and far from well-understood.

I choose to focus on capturing the fine-grained cognitive processes because it can better explain how the analysts deal with the challenges in data triage. One

main challenge for analysts exists in detecting the attack evidence among a large volume of massive data sources. It is likely to deduce the way how an analyst manages to make it on the ground of examining the data triage operations and the related hypotheses that triggered these operations.

Cognitive Task Analysis (CTA) is a traditional method for studying human working processes. Some researchers with the access to the cybersecurity analysts have conducted several CTA studies using various techniques, such as observation and interviews. Most of these studies focused on macro-level descriptions of cybersecurity analytics (e.g., the stages of data filtering and the roles of analysts), had have gained some valuable insights into the analysts' cognitive processes. However, few have attended on the fine-grained cognitive activities due to several real-world difficulties in conducting CTA studies in cybersecurity. For instance, CTA studies can be too time-consuming as analysts, who have to rotate through day shift and night shift on a 24/7 schedule (24 hours a day, 7 days a week) may have little time to participate in interviews. Besides, data triage tasks are memory-intensive and require vigorous concentration so that it is hard for analysts to give complete and accurate reports on their cognitive processes with a commonly-used think-aloud protocol.

To overcome the above disadvantages, I propose an integrated process tracing method for capturing analysts' fine-grained cognitive processes. First of all, *a trace representation* is proposed to specify a cognitive process at the fine-grained level. Drawn on the sensemaking theory in cognitive science, the trace representation identifies the key elements in an analyst's cognitive process of data triage, including actions of data triage, observations of suspicious network data, and hypotheses about the possible cyber attack activities. To capture the analysts' cognitive processes specified in the trace representation, this method integrates *automatic capture* with *situated self-reports*. On one hand, it automatically capture an analyst's data filtering and correlating actions and the resultant observations of suspicious

network events. On the other hand, the analyst spontaneously reports his/her hypotheses about attacks based on the current observations. Each hypothesis is linked automatically to the corresponding observation once generated. The automatically captured and self-reported information is mutually complementary.

As a specific measure of the tracing method, an interactive toolkit has been designed and developed to support the proposed tracing method. The biggest problem in design lies in that the toolkit should not detract the analyst from performing his/her task as is usual in think-aloud protocols. Therefore, I adopted a user-centered approach (i.e., scenario-based design) at the design stage and then to test it in the following empirical studies.

1.2.2 Empirical Study of Fine-Grained Analysts' Cognitive Process of Data Triage

A sophisticated understanding of the fine-grained cognitive process of an analyst can provide several critical benefits for enhancing the effectiveness of SOC. The fine-grained cognitive process of an expert also offers the opportunity to allow other junior analysts to leverage it to improve their own analysis efficiency. Moreover, such understanding can enhance the accountability of decision making, improving the effectiveness in analysts training, developing better cognitive aids and collaboration supports to address the three challenges described above. Furthermore, understanding the fine-grained cognitive process of an analyst is the basis for developing automation toolkits to facilitate data triage.

Several CTA studies have been conducted which provide valuable insights into the high-level processes of analysts such as their roles and the work-flows [18, 17], their cognitive demands [24], and their performance in Cyber SA data analysis [34]. However, it still remains unclear as for analysts' fine-grained cognitive activities in data triage [96]. With the proposed tracing method, it is possible to trace analysts' fine-grained cognitive processes of data triage. Therefore, an empirical

study was conducted to capture analysts' traces to gain deeper understanding of their cognitive processes. An experiment had been carefully designed to collect the traces of analysts' cognitive processes while they performed a data triage task. The duration of each experiment lasts about 100 minutes including four sessions: a pre-task questionnaire, a tutorial session (with a quiz), a data triage task and a post-task questionnaire. During the task session, each participant was asked to work with the toolkit to accomplish the assigned task. With the close collaboration with Army Research Lab (ARL), we were able to recruit thirty professional cyber-security analysts and to collect the traces of the analysts' cognitive processes in the experiment within about seven months. The experiment results verify the feasibility of the proposed method for capturing analysts' cognitive process. Moreover, several important assumptions have emerged as a result of this work:

- The captured traces contain the details of the critical data triage operations and hypotheses of analysts while they were performing the data triage task [100].
- Several common behavior patterns emerge in these traces [101] and various analytics strategies (e.g., data triage and correlation strategies) are implied by the traces [100].
- Traces help analysts with self-reflection so that they can learn from their previous experience [101], which is quite promising in training by utilizing traces.

The results of the experiment in return shows that the proposed method can be an effective way to conduct on-the-job CTA studies with less intrusion on the human performance being studied. Besides, the captured traces are a good source for further in-depth analysis to gain better understanding of how the subjects perform the assigned task.

1.2.3 Automated System for Generating Data Triage Rules from Cognitive Traces

Given the traces collected from the experiment, I proposed a trace analysis method for constructing useful data triage rules from traces which can help analysts identify and correlate the key evidence of malicious events from raw network monitoring data. Extensive research has been conducted on alert correlation, which is a milestone in automated data triage. Motivated by the benefits of cross-data-source analysis, SIEM systems have been focusing on security event management and correlation across multiple data sources. The complicated SIEM rules can be decomposed into a set of basic rules (for data filtering or correlation) and logic connectors ("AND" or "OR"). Rule generation requires large amount of time and expertise, which makes SIEM systems extremely expensive. Therefore, it is desirable to automate the rule generation process.

One important research question is how to distinguish the critical data triage operations recorded in the traces from the exploratory ones. We tackled the problem with a graph-based analysis method. A *graph model* was developed which represents the temporal and logical relationships among the data triage operations. The proposed approach has been evaluated through a human-in-the-loop study. The result of the study illustrates the rules mined from the traces are helpful for analysts to conduct data triage. The result validates the practical value of the captured traces. Although this method still can't generate the rules as complicated as those generated by experts in SIEM systems, the analysts can embark on with the from the automatically generated rules to construct more complicated rules instead of from scratch.

To the best of our knowledge, this is the first method that directly leverages traces of human analysts' cognitive processes to develop data triage automation. Relating this method to the rule-based SIEM systems, it can greatly reduce analysts' workloads in generating SIEM rules and thus render the cost of data triage