A Critical Review on Detecting Structurally Anomalous Logins within Enterprise Networks

ELEC 535 Data Analysis and Pattern Recognition

COURSE PROJECT
Submitted by

Derrell D’Souza (V00901532)

University
of Victoria

Date: April 25th, 2018

A Critical Review on Detecting Structurally Anomalous Logins within Enterprise Networks

Derrell D'Souza

Department of Electrical and Computer Engineering
University of Victoria
3800, Finnerty Rd., Victoria, BC, V8P 5C2
derrelldsouza@uvic.ca

***Abstract* – This critique is aimed at evaluating the paper 'Detecting Structurally Anomalous Logins within Enterprise Networks'. In this paper, Siadati et al. has used the Network Login Structure to identify a specific subcategory of anomalous login within an enterprise network. In this critical review, discussions will be made on the materials and the algorithm used, the assumptions made and the conditions under which they are valid, the issues pertaining to the classification of the login patterns, and the ability of the system to perform in a real world environment.**

Index Terms – CLM, semi supervised, network login structure, pattern mining, classifier.

I. INTRODUCTION

Credential-based Lateral Movement (CLM) is a special type of network attack in which the attacker relies on stolen credentials to log into to a new computer to compromise it, append it to a chain of hacked computers and reach the target server that will hold data of attacker's interest [2]. The credentials can be stolen in various possible ways -social engineering, phishing, spam, reusing stolen passwords or shared credentials, brute force or due to security question reuse.

A. Problem

The CLM based attack can easily bypass Traditional Network Intrusion Detection Systems (NIDS) as opposed to remote exploitation because unlike the remote exploitation, where changes in network traffic content can identify the attack, the network traffic content is indistinguishable in CLM and is same as a benign login. Moreover, the access control tools and policies are relaxed in business environments to tackle worst case situations and hence fails to minimize the lateral movements within the network. This allows the attackers to freely roam within the network and spread to as many systems they can with a prime goal of achieving administrative- level credentials. [2]

B. Approach

In order to detect CLM, [1] proposed a Network Login Structure that specifies normal logins within a given network. The approach is to first model a Network Login Structure by automatically extracting a collection of login patterns which describes the normal logins within the network. This is achieved using the Pattern Mining Algorithm. Next, it uses a Semi-supervised Anomaly based Detection technique to detect the anomalous logins by identifying the patterns that are inconsistent with the Network Login Structure.

II. SUBJECT PAPER OVERVIEW

A. Contributions

The subject paper makes the following contributions

1. It explores the idea of detection of Credential-based Lateral Movement using the Network Login Structure based anomaly detection.
2. It proposes a method for building the network login structure using enterprise network using login patterns and provides an algorithm to automatically and efficiently extract login patterns from a large dataset.
3. It evaluates the method using a real data dataset of logins and based on labelling by security analysts.

B. System Architecture Overview

Siadati et al. uses an architecture that consisted of two main components – the pattern miner and the classifier. [1] The pattern miner models the normal logins which specifies normal logins using login patterns. Login patterns consists of the attributes of three components - compromised user/ stolen credentials, the compromised source computer and uncompromised destination computer. The pattern miner is fed with collection of past logins i.e. login history spanning few months and the attributes of the users and both the computers during the same time interval. It uses these to mine logins and extract login patterns. The logins are processed and the patterns are given out along with their confidence score to indicate their reliability. The classifier compares the attributes of new login patterns with those extracted from the pattern miner to decide whether the login is normal or malicious.

B. Overview of Pattern Mining Algorithm

The pattern mining algorithm in [1] uses a two-step procedure to mine patterns of network login. In the first step, it enumerates candidate login patterns from each login in the login history H. The enumeration is done by generating three power sets each based on login attributes of each element and then taking Cartesian product of power sets to generate all the possible candidate patterns. In the second step, this algorithm groups login patterns, counts the number of occurrences of each, computes their orientation scores and selects patterns with orientation scores above a specified threshold.

Orientation score is calculated w.r.t a patterns orientation to a source, destination or to a user.

Sadiati et al. proposes a new algorithm based on encoding and parallelization that can efficiently extract patterns. This is because time complexity associated with computing Cartesian products for power sets for a real dataset was very expensive due to the involvement of non-polynomial values. [1] This algorithm is a two-step procedure – First, encoding is done by using binary encoding to store Cartesian product of power set of attributes. Such an encoding takes less space than storing string values and memory required to represent the patterns is minimized. Second, parallelization is done to accelerate the extraction of patterns. This is achieved by splitting the binary masks generated by encoding, into several clusters and running the pattern processing algorithm on each cluster assigned to a CPU core for processing.

C. Overview of working of Login Classifier

The login classifier used in the subject paper is a hybrid two component and evaluates each login independently. The first component uses an exact matching approach wherein a login login is classified as benign if attributes of each element of the pattern matches with a pattern present in the login history otherwise the login is classified as malicious. The second component uses pattern matching which initially generates all possible combination of attributes related to a login by enumeration and the classify the login as benign if at least one of the combinations of login attributes matches a pattern of the set of network login patterns that describe the net- work structure. The algorithm will compute a confidence score which is the orientation score for each pattern that doesn't exactly match with the past benign login but matches the normal login pattern.

D. Evaluation

Siadati et al. evaluated the system using real data sets of a single enterprise. This data set contained all login entries with their attributes, each of which identifies a unique login event between two different computers. The evaluation was carried out by first exploring the network structure. The statistics of the data set shown in Table 1.1 confirmed that the connection between the computers are structured.

Number of usernames	25,450
Usernames belonging to computer account	12,550
Other usernames (admins, non-admins ,services)	12,900
Number of Unique computer names	33,150
Number of unique login connections per day	160,000
Number of unique login connectivities	633,657

Table 1.1 Statistics of the dataset used in [1] for network structure exploration

Upon studying the organizational structure and its correlation with logins it was also confirmed that there are frequent client-server interactions and infrequent client-client interactions. Next, the login dynamics was studied to identify whether there is any change in login connectivities with respect to the history and categorize them. If the login has already occurred in the past then there is no change and if it has not occurred in the past then the login was compared with the login history to see if either of source, destination or user or all of the details had changed to categorize them accordingly. Four months of data was used to create login history and the results related to login dynamics were observed after one month. Table 1.2 shows the statistics related to the login dynamics. It was concluded that majority of current logins have occurred sometime in the past and changes in network and organizational dynamics result in logins which have not been seen in the past.

No change Logins	85-95%
Changes due to non-admin users	70%
Changes due to admin users	20%
Changes due to services	10%
Destination change Logins	2-6%
Sources changes	1-2.5%
User changes and full changes	Negligible
Destination changes with server as destination	80%

Table 1.2 Statistics of the network dynamics in [1]

Siadati et al. carried out two experiments –one along with security analysts and other using synthetic attack traces.[1] In the first experiment, login data was taken from the security analysts of the same organization to evaluate the logins that the system detects as malicious as set of bad labelled logins were not available. The system was set up using four month of login history for exact matching. For pattern matching patterns were mined using pattern mining algorithm, with each login having eight attributes as shown in Table 1.3. Over 200,000 patterns were generated by the algorithm and orientation scores for each login pattern was also calculated. Entire logins of a day (about 177,000) was used and the system generated about 578 alerts. Only a subset of logins i.e.80 was selected due to resource constraints and were labelled by three security security analysts using an online labelling panel that showed only one login at a time to minimize the bias of showing more than one alert. They marked the alerts as either- most likely a good login, needs further investigation or most likely a malicious login. Consider the labelling as ground truth Sadiati et al. considered first case of labelling as a benign login and the other two cases were considered as suspicious logins. Out of 80, 11 alerts were labelled as suspicious, which suggests that the accuracy of the system is 13%.

Component	Attributes
User	Type, Business Unit
Source Computer	Type, Application Name, Location
Destination Computer	Type, Application Name, Location

Table 1.3. Attributes of login elements used in pattern mining

The second experiment was done using synthetic attack traces because the security analysts gave only the suspicious logins but not confirmed malicious ones. So without knowing the actual count of malicious logins, it was not possible to identify how many malicious logins that the algorithm misses. Five month of data was used out of which four month of data was used to train the classifier and extract the login structure and one month of data was used to calculate the false positive rate(number of benign logins classified as malicious). 150 Malicious login traces were generated and some workstations were randomly selected as a source of malicious login and five random target computers were selected as destination. These malicious login traces were injected into dataset of logins. Credentials were identified on the source workstations and were marked compromised as these would be possibly used by the attackers. For measuring performance of the experiment confidence scores, false positives, true positives and ROC curve were used. Patterns having higher orientation score were considered more reliable and hence had higher confidence. The subject paper states that a high threshold on orientation scores would easily detect malicious logins thereby resulting in an increased true positive rate while low threshold on orientation scores would easily detect new benign logins and decrease false positive rate. With increase in threshold on orientation score, both false positive and true positive rate increases and hence Sadiati et al. used Receiver Operating Characteristic (ROC) curve to achieve optimal threshold. The results were found to be 82% malicious logins and 0.3% false positive rate from ROC curve.

III. DISCUSSION

A. *Assumptions and their validity*

The subject paper assumes that the login history is free of any attacker's login patterns. In the real world scenario, it is possible that a stealthy attacker stays in the network for a long time even before the proposed system was implemented. He may create some logins with the goal of misleading the pattern miner module to include an illegitimate pattern in the set of login patterns. Alternatively, if the attacker is completely aware of the login structure, he can also use login credentials that are normally used between two users. Hence, if the login history has illegitimate patterns created by the attackers then it is possible for him to evade the detection.

In one of the experiments conducted along with the security analysts for statistical hypothesis testing, the subject paper assumes the labelling by the security analysts as the ground truth. This experiment resulted in an overall accuracy of 13%. This implies that there were a lot of false positives which resulted in a low accuracy. The security analysts were not able to make sure that the login is malicious due to resource constraints. Hence the difference between malicious and normal logins will be harder to recognize without a proper ground truth.

The subject paper does not explicitly explain the usage of four months of data for creating the login history and the usage of one month of data for testing the classifier. If the login history is found to be insufficient, then it can generate a lot of false positives and can degrade the system performance. Sadiati et al. found that the admin logins are infrequent and many did not pass the threshold to be included in the history. By studying the effect of longer period of logins and expanding the login history, it may be possible to model the admin logins and include them in the login history.

Several assumptions were made with respect to the nature of the attack while testing the system with synthetic attack traces. The following were assumed.

- The attacker has already infected and compromised a workstation within an enterprise network and tries to transmit to another computer.
- The attacker is able to steal the password of any user who is active on a compromised computer. This includes accounts that are used to login from or to the compromised computer.
- The transmission of an attacker is naturally constrained by standard access controls of a network.
- A stealthy attacker minimizes the number of malicious logins, as too many logins can raise an alert. Therefore, attackers that we simulated tried to login only to five other computers in the network

In the real world scenario, it is possible for an intelligent adversary to compromise the access control by gaining physical access, eavesdropping by observation, bypassing security, exploiting hardware and software, reusing or discarding media, electronic eavesdropping, intercepting communication etc. This may result in severe losses including disclosure of personal information, corruption of data, loss of business intelligence, danger to staff, facilities and systems, damage to equipment and a total failure of business and system processes. [3]

The subject paper implements a fast encoding and parallelization algorithm to extract mining patterns. It claims that it is easier to retrain this algorithm to adapt to the amount of changes in the login. However, the potential issues that can arise due to parallelization in real world scenarios such as synchronization, load balancing etc. has not been discussed.

The paper assumes an inherent ergodicity. It uses the ROC curve to tune the model complexity and to find the optimum threshold value for the system to perform with high true positive and low false positive. It doesn't discuss the changes in the ROC curve that can occur if applied to a different data set. This classifier is non-stationary to accommodate changes in login dynamics and if it is implemented in a developed environment where the login attributes change frequently than the algorithm won't be able to match the pace of its training data update with the dynamics of the logins.

B. *Implications*

The classifier used is a semi supervised anomaly based detection approach. Hence the training data used has the labelled instances for normal class. Intuitively, during training only probability of the normal data is modelled and during testing the logins which have low probability according to the match are classified as malicious logins. The bulk of login data stems from the same distribution are known as normal logins while those from a different distribution are malicious.

If the subject papers results are to hold, then the following conditions must exist.

- The network should be structured as this method is not useful for unstructured networks.
- The attacker should not be able compromise the access control. If he does, than he can corrupt the data, move laterally anywhere in the network without being detected.
- The login history should be free of any malicious logins to ensure proper training and no false negatives.
- The attacker must not combine CLM method with vulnerability based lateral movement, otherwise a separate software will be required along with the proposed system to detect the remote vulnerabilities.
- The network should be fairly static and not be largely non-stationary. Otherwise, it will be impossible for the system match its update of training data with that of login dynamics.
- There should not be any dramatic changes in the network, otherwise additional intervention may be required.

Some of the potential challenges that needs to be addressed before the proposed semi-supervised anomaly based detection technique is applied in a real world scenario include:

- Often it is may be the case that the normal region of the classifier won't be able to capture all the normal logins which results in the boundary between the normal and malicious logins to be blurred. Such a situation can generate a lot of false positives.
- Anomalies are as a result of malicious logins. The intelligent adversaries will always try to make anomalous observations seem normal.
- The normal login may change over time and then a current notion of normal behavior may not be applicable for future. Hence the login history should be updated periodically to match the network dynamics.
- Sufficient labelled data for login history is required for the anomaly based detection techniques to perform well.
- If the normal login data contains a lot of noise, then it is difficult to separate noisy instances from malicious logins. One form of noise would be missing attributes of elements of some login due to inability of the system to update the details of attributes of new users and computers quickly in the database.
- Cleaning up the mislabeled data in the validation set is necessary to reduce off-training error.

The system will only perform well in a network where the login dynamics do not change frequently. Siadati et al. has tested the system over only a single data set limited to a one type of enterprise network in a single company. As a result its performance is questionable in the real world scenario. To apply it to a real world scenario, it needs to be tested over more data sets and different type of enterprise networks as the stability of the login structure varies from enterprise to enterprise. The optimal window over which the algorithm should be retrained should also be studied for good generalization over other data sets. The pattern mining algorithm may encounter several issues during encoding and parallelization which should be properly dealt with. Lastly, emphasis should be laid on using a login history void of any attacker's login patterns.

There is a clearly ambiguity in the subject paper over the usage of cross-validation set and training set. It appears that Sadiati et al. uses the same set for testing and cross-validation. The learning algorithm works well over the validation/development set used, generating 82% true positives and 0.3% false positives. As already mentioned four month of data was used for training and one month of data was used for validation. Since the validation and training data set came from the same distribution, chances of overfitting is more because of which the system may not work well on data sets of other enterprises. In such a case, it is required that the algorithm should be tested over more data sets before applying this algorithm over same type of enterprise network in the real-world scenario. [4] However if the test set is entirely different from the validation set, which may occur in an environment where there is dramatic change in the login dynamics over time, the work to improve performance over validation set will be useless. Such a mismatched dataset will result in an added uncertainty about whether improving the cross validation set distribution also improves the test performance. This will in turn make it harder to figure out what is and isn't working, and thus makes it harder to prioritize what to work on. The cross validation set should be large enough for an accurate detection.

The subject paper does not discuss the false negatives which has a high possibility to be encountered in a real world scenario. If the login history is corrupted with the attacker's login pattern then it is possible for the system to have a lot of false negatives and the system will be unable to detect such attacks. In other words, the proposed system will become useless.

IV. CONCLUSION

Overall, the subject paper works well for the dataset it is trained on. It's generalization over other datasets is questionable as it is not trained over multiple datasets. Also the classifiers ability to handle huge amount of data is dubious as it was only trained over limited amount of data. The inability of the system to tackle large non-stationarity of login data makes it unsuitable for a developed environment where the login dynamics changes dramatically over time. The semi-supervised anomaly based approach is a good algorithm for separation of normal logins and malicious logins which aims to tackle the inherent problem of base-rate fallacy by using a hybrid two component classifier. However the potential issues related to classification such as the availability of sufficient labelled data, preprocessing to filter out the noise from the data, corruption of login history and the optimal window for retraining the algorithm needs to be studied and investigated further before applying the algorithm to a real world scenario. A lot of further work on error analysis, model selection and hyper parameter tuning also needs to be done before applying the algorithm to a real world problem.

ACKNOWLEDGMENT

I would like to thank Dr. Neville and all the Graduate Students for their valuable inputs and suggestions for my paper during the project presentation.

REFERENCES

[1] H.Siadati, N.Memon, "Detecting Structurally Anomalous Logins Within Enterprise Networks", *ACM CCS* , 2017

[2] J.Miller-Osborn, "Credential Based Attacks: Exposing the Ecosystem and Motives Behind Credential Phishing, Theft and Abuse", *Palo Alto Networks,* 2017.

[3] D.Kim, M.G. Solomon, *Fundamental of Information Security Systems,* 3rd Edition , ISSA, 2010

[4] A.Ng, *Machine Learning Yearning (Draft Version*), 2018