Intrusion Detection on IoT Services using Event Clustering

Boseok Park , Sangwook Kim

School of Computer Science and Engineering
Kyungpook National University
e-mail : boseok4u@knu.ac.kr, kimsw@knu.ac.kr

이벤트 클러스터링을 이용한 IoT 서비스 침입탐지

박보석, 김상욱
컴퓨터학부
경북대학교

Abstraction

In this paper, we describes the analysis of security methods in the area of IoT and describes a mechanism that analyzes logs generated by IoT devices attacks. We models an event network based on a graph of interconnected logs between network devices and IoT gateways. Moreover, suggests an algorithm that correlate logs into single meaningful messages.

1. Introduction

A botnet, which is adding new bots every day, has already infected one million businesses. Therefore, we need to consider how to protect and control countless IoT devices. However, different types of security frameworks are used in IoT devices[1,2].

In this paper, we propose event network model that include log contexts between network devices and IoT devices based on closeness of relationships. Each node corresponds to a IoT service and two services are linked if there is work in cooperation that they can bind to each other in the system. And we also explain how to select suspicious event and correlate in large networks. Through the proposed method, users who are neighbors can be searched for emergency then notify the context and respond quickly.

2. Related Works

There are several approaches to the event correlation task: rule based reasoning(RBR), model based reasoning(MBR), state transition graphs(STG), and so on[1].

An RBR system consists of three basic parts including a working memory, a rule base, a reasoning algorithm[1]. Using an RBR system to develop an event correlator that covers the entire domain of the enterprise is not a good idea. The enterprise is large, dynamic, and generally hard to understand.

An MBR system represents each component in the enterprise as a model[2]. A model is either a representation of a physical entity(e.g., a firewall, router), or a logical entity(e.g., network session, suspicious process). And a model represents a physical entity is in direct communication with the entity. A description of a model includes three categories of information: attributes, relations to other models, and behaviors. Event correlation is a result of collaboration among models.

The key concepts in the STG approach are a token, a state, an arc, a movement of a token from one state to another state via an arc, and an action that is triggered when a token enters a state[1,3]. The measure of correlation is dependent on the knowledge that represents the attack scenarios.

3. Event Network Model

As events arrive at the logging server, we use the topology information part of the event to identify the source of the event.

We defines events e as tuple <s, d, w> in which s, p and m is source device, destination device and alert weight respectively. Therefore, $e_1e_2e_3$, $e_3e_2e_1$, $e_2e_1e_3$, etc., are the same pattern. Furthermore, the multiplicity of an event within a pattern is equivalent to a single occurrence. Σ^k denote the set of words with length k. Thus, $\Sigma^k = \{w|\ w \text{ is a word over } \Sigma \text{ and } (|w| = k)\}$. For example, if $\Sigma = \{e_1, e_2, e_3\}$, $\Sigma^2 = \{e^1e^2, e^1e^3, e^2e^3\}$.

We defines an event network EN as $\Sigma^+ = \Sigma^1 \cup \Sigma^2 \cup \Sigma^3 \cdots = \cup \Sigma^k$. Note that Σ^+ is the set of words that might be constructed from one or more events of Σ, and is the largest set of possible words we might observe from an entity being modeled.

A $P_{aned}(k)$, average degree of the first neighbors of events with degree k is as follow and evaluates using ALGORITHM 1.

$$P_{aned}(k_j) = \frac{1}{k_i}\sum_{j=1}^{N} A_{ij}k$$

ALGORITHM: Event Clustering Algorithm

Input: A event network Σ^+, degree k and expire time t

Output: A probabilities P that have $P_{aned}(k)$

```
1: set P(k) to w value in all event tuples
2: for the events i connected to EN do
3:    $P(k_i) = sum(w_i)/k$
4: end
5: sampling two links with probability P(k) from
   the network: $e_1(s_1,d_1)$ and $e_2(s_2,d_2)$
6: measure the degrees $j_1$, $k_1$, $j_2$, $k_2$ of nodes $s_1$,
   $d_1$, $s_2$, $d_2$. replace the two selected links
   with two new ones $(s_1,s_2)$ and $(d_1,d_2)$ with
   probability :
   if $(w_{j_1j_2}w_{k_1k_2} < w_{j_1k_1}w_{k_2j_2})$ $Pe_1e_2(k) = w_{j_1k_1}w_{k_2j_2}$
   else $Pe_1e_2(k) = w_{j_1j_2}w_{k_1k_2}$
7: repeat from step 5 until time t is not
   expired
8: Return the P(k)
```

The algorithm updates intrusion probabilities P(k) repeatedly and correlates events within limited time. An event will be active until a predefined value is reached when it is considered obsolete and should be removed from the queue. This behavior can be modeled as a step function. In addition, an event is assigned a weight based on its age in the queue.

5. Implementation

To evaluate the model and algorithm, IoTroop botnet is used as an attack scenario.

Fig. 1 shows an example of the correlation analysis of the event network in an attack scenario. The Path marked in red are suspected paths of intrusion. We implemented using R igraph package, and the modularity value obtained by using function is as follows.

```
modularity(cluster_edge_betweenness(EN))
[1] 0.8359884
```

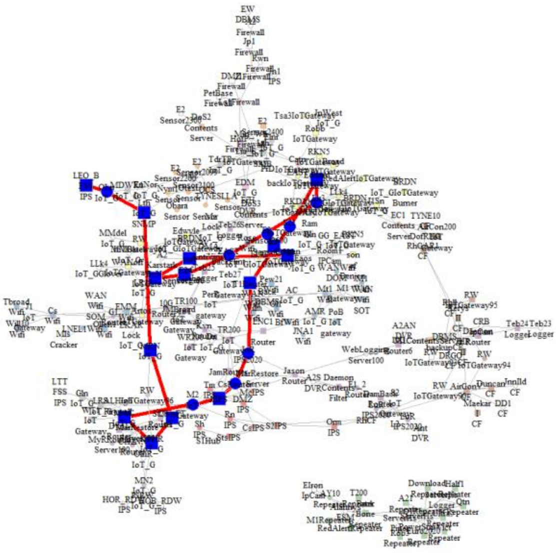

Fig. 1. Detected path in event network

6. Conclusion

In this paper, we introduced a framework and model for event correlation in security systems. We suggested algorithm that sample event based on probability within k neighbors in large network. Graph and network theories are used for correlation.

We introduced event network model based on graph. In addition, we suggested event network sampling and correlation algorithm.

ACKNOWLEDGMENTS

This work was supported by the BK21 Plus project (SW Human Resource Development Program for Supporting Smart Life) funded by the Ministry of Education, School of Computer Science and Engineering, Kyungpook National University, Korea (21A20131600005).

References

[1] A. Buczak and E. Guven E. 2015. Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. IEEE Communications Surveys & Tutorials, 18, 2, 1153-1176

[2] P. Kim and S. W. Kim. 2017. Detecting Community Structure in Complex Networks Using an Interaction Optimization Process. International Journal of Physica A, 46, 5, 525-542.

[3] S. Ryu and S. W. Kim. 2019. Neighbor Recognition by User Relationships in Internet of Things Graph. In Proceeding of the HCI Korea 2019, 36, 163-166.