Poster: An Approach to Verifying Threat Intelligence Based on Graph Propagation

Xin Wang, Zhigang Lu, Zhengwei Jiang, Qiang Li

Institute of Information Engineering, Chinese Academy of Sciences

School of Cyber Security, University of Chinese Academy of Sciences

Beijing, China

{wangxin9032, luzhigang, jiangzhengwei, liqiang7}@iie.ac.cn

***Abstract*—Threat intelligence is evidence-based knowledge, including context, mechanisms, indicators, implications and actionable advice that helps identify security threats and make clear decisions. However, there are a large number of false positive indicators in the current threat intelligence database. This poster demonstrates our approach to verifying threat intelligence based on graph propagation. We represent the relationship of indicators as a complex network of directed graphs. The threat value of each node propagate among neighbors, and finally the true positive indicators are selected according to it.**

Keywords—threat intelligence; verifying; graph propagation

I. INTRODUCTION

Threat intelligence is a new concept that comes along with the emergence of massive data in the era of big data. Gartner first defines threat intelligence, which is evidence based knowledge and can provide clear decisions for analysts. According to the 2018 Cyber Threat Intelligence Survey[1] just released by the SANS™ Institute: 90% of respondents report that they consume CTI data or plan to use in the future.

There are many threat intelligence sources, but without verification. As data volume continues to rise, the quality of CTI data has dropped away. More than 83% of analysts said that they have received too many alerts and false positives according to a survey[2] by the Ponemon Institute. Because of the large amount of noise data in threat intelligence, it's difficult to understand and respond to new threats. For this reason, we concentrate on threat intelligence verification in this poster.

In the field of threat intelligence, there are some studies on threat intelligence assessment. Pawel Pawlinski[3] introduce the method of evaluating threat intelligence feeds, including five aspects: relevance, accuracy, completeness, timeliness and ingestibilty at first technical colloquium for threat intelligence. Omar[4] introduce the notion of quality of indicators (QoI) for the assessment of the level of contribution by participants in information sharing for threat intelligence. But they all focus on the overall assessment of threat intelligence. In practice we always need to verify and evaluate single indicator. Amine[5] calculate the PageRank of each indicator as an evaluation result, but it cannot solve the problem of false positive.

Contributions:

In this poster, we propose a novel approach to verify threat intelligence. Our poster makes two contributions as follows:

1) We design a framework for threat intelligence verification, which focus on how to distinguish between positive and negative indicators.

2) We implement a prototype system based on the proposed framework and calculate the threat value for each indicator.

II. DESIGN

The approach of verifying threat intelligence consists of steps including initial threat calculation, the construction of a relational graph and the propagation of threat value. The architecture of framework is shown in Fig 1.

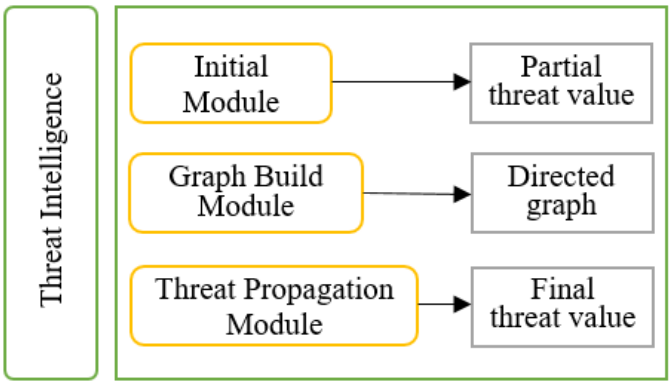

Fig. 1. The architecture of framework

A. Initial Module

First, we define threat value as the effective degree of an indicator. The lower the threat value, the more likely it is to be false positive. The initial module calculate the threat value of a small part of indicators, including expert evaluation and majority voting from open threat intelligence source. For example, the threat value of indicator A is 0.2, evaluated by experts. Indicator B is 0.6, calculated by majority voting from IBM X-Force, Threat Crowd, Virus Total, ThreatBook, 360 and so on.

B. Graph Build Module

The graph build module represent the relationship of indicators as a complex network of directed graphs. The nodes on the graph are indicators. We design the weights of edges, which affect threat propagation. The weight (see TABLE I) is calculated by product of propagation index and reliability index. The propagation index takes a different value based on the type of the edge node. The reliability index indicates the reliability of the edges established by the relationship. The types of relationships can be divided into three categories: parsing service based, sample evidence based and artificial analysis based.

TABLE I. THE WEIGHT OF EDGE

Type	Relationship	Propagation index	Reliability index
Parsing service	domain-[resolve]-ip	0.4	1
	domain-[resolve]-email	0.6	1
	……	……	……
Sample evidence	md5-[evidence]-ip	0.8	0.8
	md5-[evidence]-url	0.8	0.8
	……	……	……
Artificial analysis	url-[associate]-email	0.6	0.7
	ip-[associate]-domain	0.4	0.5
	……	……	……

C. Threat Propagation Module

Threat indicators are interdependent and correlated. The threat propagation module uses the constructed graph to propagate threat value. In the directed graph, the threat value of each node propagate among neighbors. Mathematically we calculate the threat values of each node by formula (1) :

$$T_v = I_v + \frac{1-I_v}{n} * \sum_{i=1}^{n}(T_v(i) * \Phi(i)) \quad (1)$$

The I_v represents the initial threat value of an indicator, T_v represents the final threat value, which is composed of initial value and propagation value. The propagation value is calculated by summing up the $T_v * \Phi$ of each neighbor node and the normalization has been applied. In addition, the Φ stands for the weight of propagation.

We calculate the threat value iteratively through the above formula, and get the final threat value of each node. Finally, we select the true indicators according to it.

III. PRELIMINARY RESULTS

We implement a prototype system based on the proposed framework. In order to test the effectiveness of the approach, we collect thousands of indicators about the "Xcode Ghost" incident, and build a directed graph with tens of thousands of edges. Some of experimental results are listed in Table II. We have calculated the threat value of some unknown indicators, the false positive indicators are low, while those of the true positive indicators are higher.

TABLE II. THREAT PROPAGATION RESULTS

Indicator	Initial Threat Value	Final Threat Value
saltsecond.net (C2)	0	0.76
208.73.211.183 (Phishing)	0.1	0.6
198.58.94.108 (Botnet)	0.1	0.81
1d80e359d448f2679b53aa67d6f1437a (Trojan)	0.8	0.91
www.eygwindows.co.uk (Phishing)	0	0.75
……	……	……

IV. CONCLUSION AND FUTURE WORK

In this poster, we propose an approach to verify threat intelligence and then we design a framework for it. Moreover, we implement a prototype. Experiments show that the proposed approach is feasible and can calculate threat values for indicators. The results make it possible to distinguish between positive and negative indicators. In the future, we will collect more indicators and do some detail experiments.

ACKNOWLEDGMENT

This work is supported by Key Laboratory of Network Assessment Technology, the Chinese Academy of Sciences and Beijing Key Laboratory of Network security and Protection Technology. Foundation item: National Key Research and Development Program of China (NO.2016QY06X1204).

REFERENCES

[1] SANS. CTI in Security Operations: SANS 2018 Cyber Threat Intelligence Survey[EB/OL].[2018-02-05].https://www.sans.org/reading-room/whitepapers/analyst/cti-security-operations-2018-cyber-threat-intelligence-survey-38285.

[2] LLC, P. I. The Importance of Cyber Threat Intelligence to a Strong Security Posture[EB/OL].[2015-03-01].https://webroot-cms-cdn.s3.amazonaws.com/9114/5445/5911/ponemon-importance-of-cyber-threat-intelligence.pdf.

[3] I Paweł Pawlinski, Evaluating Threat Intelligence Feeds[EB/OL].[2016-02-24].http://www.necoma-project.eu/m/filer_public/b9/da/b9dafadd-adf8-4875-afd5-2d188dd96449/pawel-pawlinski-evaluating-ti-feeds.pdf.

[4] Al-Ibrahim O, Mohaisen A, Kamhoua C, et al. Beyond Free Riding: Quality of Indicators for Assessing Participation in Information Sharing for Threat Intelligence[J]. arXiv preprint arXiv:1702.00552, 2017.

[5] Boukhtouta A, Mouheb D, Debbabi M, et al. Graph-theoretic characterization of cyber-threat infrastructures[J]. Digital Investigation, 2015, 14: S3-S15.