J. Sustain. Wireless Syst., vol. 02, no. 4, pp. 149-154
https://doi.org/10.36548/jsws.2020.4.002

Data Elimination on Repetition using a Blockchain based Cyber Threat Intelligence

Dr. S. Smys,
Professor,
Department of CSE,
RVS Technical Campus,
Coimbatore, India.
Email id: smys375@gmail.com

Dr. Wang Haoxiang,
Director and lead executive faculty member,
GoPerception Laboratory,
NY, USA.
Email id: hw496@goperception.com

Abstract- Cyber threat is a major issue that has been terrorizing the computing work. A typical cyber-physical system is crucial in ensuring a safe and secure architecture of a sustainable computing ecosystem. Cyber Threat Intelligence (CTI) is a new methodology that is used to address some of the existing cyber threats and ensure a more secure environment for communication. Data credibility and reliability plays a vital role in increasing the potential of a typical CTI and the data collected for this purpose is said to be highly reliable. In this paper, we have introduced a CTI system using blockchain to tackle the issues of sustainability, scalability, privacy and reliability. This novel approach is capable of measuring organizations contributions, reducing network load, creating a reliable dataset and collecting CTI data with multiple feeds. We have testing various parameters to determine the efficiency of the proposed methodology. Experimental results show that when compared to other methodologies, we can save upto 20% of storage space using the proposed methodology.

Keywords: Sustainable computing; blockchain; cyber threat intelligence; attacks; scalability and reliability;

1. Introduction

When a cyberattack occurs, the system is in a vulnerable state. However, a number of existing systems have addressed this issue, focusing on recovering, minimizing and restoring data. However, with the help of Cyber Threat Intelligence (CTI) [1], we can pre-emptively calculate the attack and act accordingly. Attacks on organizations with Advanced Persistent Threat (APT), extortion of financial profit using Ransomware and paralyzes of the system through Distributed Denial of Service (DDoS). Taking into consideration these attacks, CTI serves as a proper mechanism to securely transfer information. Moreover, small and medium scale business finds it difficult to sustain against the different attacks and the use CTI will enable secure transaction. The goal of using CTI is to reduce resources and manpower by automating all processes so that based on the threat, the system can respond pre-emptively [2].

The first step in CTI is collection of data which can be used as a reference to predict threats. Hence sufficient information is researched and stored. Recently, this information is found to be increasing at a fast pace as reliability of data has become a crucial issue and their credibility has also been questioned. Apart from credibility issues, a number of organisations face the complication of addressing the various system requirements and variation. When the gathered information is false or wrong, it might result in wrong analysis, which will further compromise the security of the system. Hence this will lead to cyber threats going unidentified and thereby prevention of such threats will also become impossible [3]. Data collection is not only important for attack identification, but it is also essential for reducing false-positive security rate. Moreover, since the amount of data gathered is high, it will be difficult to find an apt storage limit and will require a better way of managing it, by means of cloud network. The issue with using external cloud storage is that it will be using a CSP that oversees data access and will lead to issues like data privacy and integrity. Hence we require a system that will lower communication overhead and also reduce resources to use the cloud storage in an optimal manner [4-5]. In the proposed methodology, we use blockchain architecture to process data in a more efficient manner and to give privacy and security in a distributed way. Data feed node and the cloud's server are used in blockchain in order to verify data. The paper can be organised such that the following section briefly outlines

ISSN: 2582-3167
Submitted: 10.12.2020
Accepted: 28.12.2020
Published: 05.01.2021

the existing methodologies while section 3 explains the proposed methodology. Experimental results are carried out in section 4 and based on the observations, a conclusion is drawn in section 5.

2. Literature Survey

The CTI defends the information against cyber-attacks by analyzing the threat and providing adequate counter measures. Hence a number of researches are going on to determine a standard for sharing threat information. TAXII and STIX are the two methods adopted internationally to address this issue. There are three methods of information sharing in Trusted Automated eXchange of Intelligent Information (TAXII). Similarly, Structured Threat Information eXpression (STIX) is used to automatically share, analyse and collect CTI information and threat. A framework on data collection from social media websites are used to gather information and detect the threats in [6]. A study has been conducted in [7] which is used to verify how OSINT (Open Source Intelligence) is reliable and performs a comparative study on the available CTI. It showed that it enabled a more reliable model with informed consumer perspective. A methodology for gathering information using crawling data is proposed and analyzed in [8] with the help of open source tools.

In [9], an analysis was made on the cyber attacks and it was used to predict the future attacks using the hacker forum. Similarly, data can be generated randomly to increase the total amount of information that can be manipulated with ease. It is also possible to manipulate non-existent information in order to infect the other existing information in an organization. Moreover, when the system becomes automated, wrong information might damage the entire system, on a large scale. In a number of countries like UK, legal procedures and governments encourage the data of information sharing to be available globally. MISP (Malware Information Sharing Platform) is one type of globally shared information that is being used by more than 6000 organisations. In [10], Daire Homan et al. introduced a methodology which could be used to share information without the need for disclosing personal information. A complete study on different ways to gather network data was analyzed by Zhou et al. in [11]. From the consumer point of view, it is necessary to identify a dataset that is reliable so that the security of the organisation is intact. In view of this, we propose a novel methodology which imposes a secure direct reliability of data that can be fed with the help of a more efficient process [12].

3. Proposed Methodology

This paper presents a model which is used to prevent false information from tampering the network as it provides a serious threat in CTI sharing. We have introduced a CTI network with blockchain to ensure that information sharing is intact and secure. The advantage of this method is its ability to maintain integrity of the data while sharing, safe environment and efficient management of data and verification of information as and when received.

3.1 System Model

We propose to use a third party organization which is used to judge the credibility of the data collected. Data that is obtained from a particular feed have possibility of fake data. However, when a third party is used, we can easily determine if the data is credible without the need for other personal information. The architecture proposed in this work is represented in Fig.1. where, based on reliability of the network, the network information is obtained. In the figure, feed represents the organisation that shares and collects information. At the device layer, network data is produced while the cloud server will receive feed’s data and secures it in order to make it serviceable to the customer.

ISSN: 2582-3167
Submitted: 10.12.2020
Accepted: 28.12.2020
Published: 05.01.2021

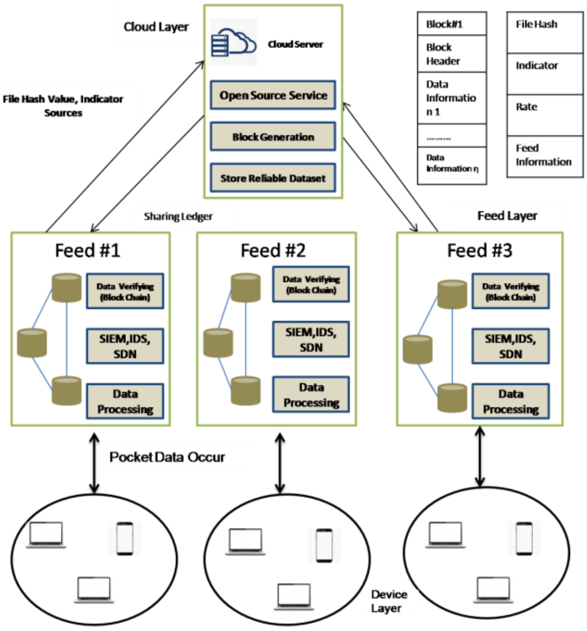

Fig.1.Architecture of Proposed Methodology

There are two ways in which network packet data is collected: by grouping packets or by randomly reducing the data collected. The Cloud Server used is that of an independent third party. It makes use of minimum communication frequency, low data transmission rate and storing of data integrity instead of the actual data to address prevalent issues like data storage space and network load.

3.2 Flow of Methodology

There are five crucial steps to be followed in the proposed methodology:

1. The first step involves gathering the data from device layer and it eliminates data that is repetitive and unnecessary, thereby ensuring that data collection is done in a more efficient manner. Any data that is identified as threat information is collected with IP address and file hash. Collection of data is done at the organisation's core network in order to avoid unnecessary traffic.
2. Using Processing feed, the data collected is pre-processed to extract IP Address and File Hash Value.
3. By sharing this File Hash value, it is possible to send information without having to enclose the important data. This Hash value is compared with other Hash Values received from similar organisations. On determining the indicators, the consumers are intimated about the most used data.
4. Similarly, if the same information is being sent as a part of building another data set, it is set as duplicate in order to prevent repetition of data and large storage area.
5. Blocks in the CS are created that are made up of data that is collected using blockchain. This allows the efficiency of the architecture to improve dramatically by enabling information storage in a more well-defined and reliable manner.

Based on consumer perspective, a service scenario is represented in Fig.2. Here information is requested from the consumer and data that is already analysed and approved is sent to the Feed.

ISSN: 2582-3167
Submitted: 10.12.2020
Accepted: 28.12.2020
Published: 05.01.2021

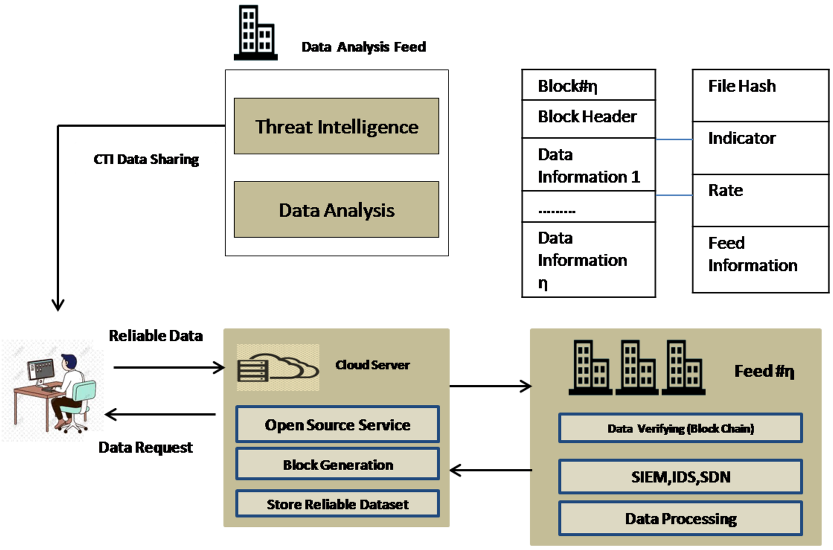

Fig.2. Service Scenario- Consumer Perspective

4. Results and Discussion

In Fig.3, it is observed that as the number of cooperating feeds increases, the network resources decreases. In the graph below, F1 is represented as 1 and F1+F2 is represented as 2 represented on the x-axis. There is not much data reduction when the second information occurrence and cooperative feeds are less, but when they are more in number, the reduction in resources will be more prominent.

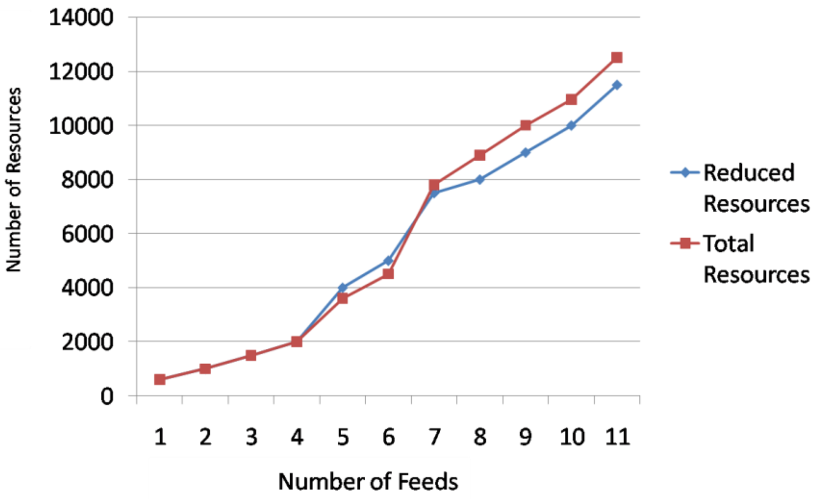

Fig.3. CTI data resources

When the number of feeds is greater than 6, the resource reduction can be observed in Fig.4. There is an increase in network resources as number of feeds increases and the change is more prominent in 10 feeds when compared to 6 feeds. The resources are reduced by about 20% with respect to the network. Fig.5. shows the feeds from 1 to 10 and their impact on resources and reduced resources. The Red bar shows the number of reduced data and the Blue bar indicates the data collected. The Green line shows the contribution score that can be used in cooperative systems in the future. In F1, more contribution is scored because of the large data collection in comparison to the data contributed while in F1 the contribution is comparatively low because of their inadequate contribute to improve data quality.

ISSN: 2582-3167
Submitted: 10.12.2020
Accepted: 28.12.2020
Published: 05.01.2021

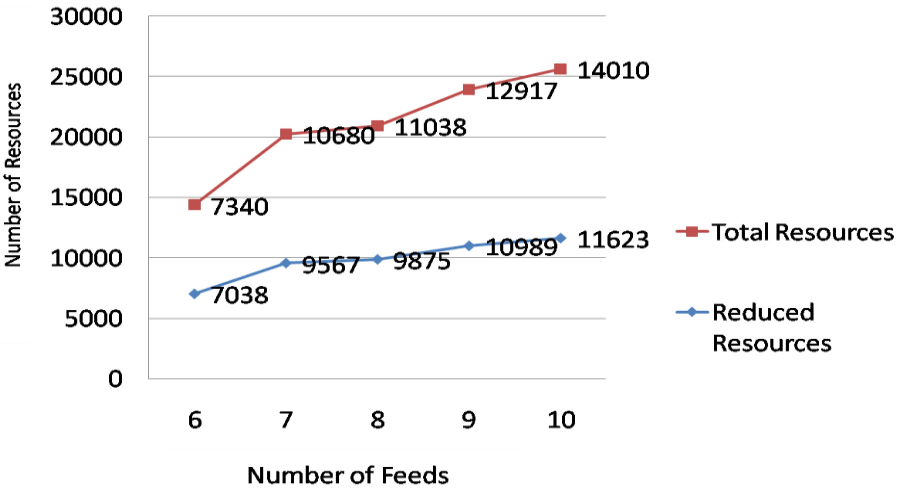

Fig.4. CTI data resources for 1 to 10 Feed

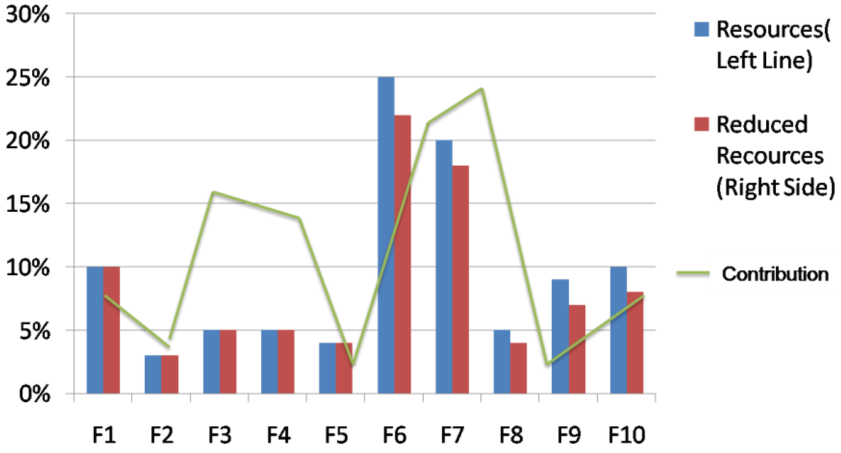

Fig.5.CTI data resources for 6 to 10 Feed

5. Conclusion

In this proposed work, we have introduced a blockchain based CTI that is used to share, analyze and collect data without the need for revealing threat information by means of using multiple feeds to check credibility of the data. Hence, independent of the organisation, we can gather data based on their reliability from the individuals. The implementation of blockchain plays a crucial part in achieving this status and a number of issues such as scalability, efficiency, privacy and reliability. The results indicate that when using third-party resources, it will reduce our network resources. Moreover, this will also result in saving space by 20% when compared with other network resources. As a future scope, we can further extend this methodology to analyse data automatically along with automated application of security policies.

References

[1] Macaulay, T. (2015). U.S. Patent No. 9,118,702. Washington, DC: U.S. Patent and Trademark Office.

[2] Burger, E. W., Goodman, M. D., Kampanakis, P., & Zhu, K. A. (2014, November). Taxonomy model for cyber threat intelligence information exchange technologies. In Proceedings of the 2014 ACM Workshop on Information Sharing & Collaborative Security (pp. 51-60).

[3] Samtani, S., Chinn, R., Chen, H., & Nunamaker Jr, J. F. (2017). Exploring emerging hacker assets and key hackers for proactive cyber threat intelligence. Journal of Management Information Systems, 34(4), 1023-1053.

[4] Tundis, A., Ruppert, S., & Mühlhäuser, M. (2020, June). On the Automated Assessment of Open-Source Cyber Threat Intelligence Sources. In International Conference on Computational Science (pp. 453-467). Springer, Cham.

[5] Gong, S., & Lee, C. (2020). BLOCIS: Blockchain-Based Cyber Threat Intelligence Sharing Framework for Sybil-Resistance. Electronics, 9(3), 521.

ISSN: 2582-3167
Submitted: 10.12.2020
Accepted: 28.12.2020
Published: 05.01.2021

[6] Haoxiang, W., & Smys, S. (2020). Secure and Optimized Cloud-Based Cyber-Physical Systems with Memory-Aware Scheduling Scheme. Journal of trends in Computer Science and Smart technology (TCSST), 2(03), 141-147.

[7] Gong, S.; Cho, J.; Lee, C. A (2018) Reliability Comparison Method for OSINT Validity Analysis. IEEE Trans. Ind. Inform., 14, 5428–5435.

[8] Koloveas, P.; Chantzios, T.; Tryfonopoulos, C.; Skiadopoulos, S. A (2019) crawler architecture for harvesting the clear, social, and dark web for IoT-related cyber-threat intelligence. In Proceedings of the 2019 IEEE World Congress on Services (SERVICES), Milan, Italy, 8–13 July 2019; Volume 2642.

[9] Shakya, S. (2020). Survey on Cloud Based Robotics Architecture, Challenges and Applications. Journal of Ubiquitous Computing and Communication Technologies (UCCT), 2(01), 10-18.

[10] Homan, D.; Shiel, I.; Thorpe, C. (2019) A New Network Model for Cyber Threat Intelligence Sharing using Blockchain Technology. In Proceedings of the 10th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Canary Islands, Spain, 24–26.

[11] Zhou, D.; Yan, Z.; Fu, Y.; Yao, Z. (2018)A survey on network data collection. J. Netw. Comput. Appl. 2018, 116, 9–23.

ISSN: 2582-3167
Submitted: 10.12.2020
Accepted: 28.12.2020
Published: 05.01.2021