sustainability

MDPI

Article

# Blockchain-Based Cyber Threat Intelligence System Architecture for Sustainable Computing

**Jeonghun Cha [1], Sushil Kumar Singh [1], Yi Pan [2] and Jong Hyuk Park [1,*]**

[1] Department of Computer Science and Engineering, Seoul National University of Science and Technology (SeoulTech), Seoul 01811, Korea; ckwjdgns@seoultech.ac.kr (J.C.); sushil.sngh001007@seoultech.ac.kr (S.K.S.)
[2] Department of Computer Science, Georgia State University, Atlanta, GA 30302-5060, USA; yipan@gsu.edu
* Correspondence: jhpark1@seoultech.ac.kr

Received: 30 June 2020; Accepted: 7 August 2020; Published: 9 August 2020


**Abstract:** Nowadays, the designing of cyber-physical systems has a significant role and plays a substantial part in developing a sustainable computing ecosystem for secure and scalable network architecture. The introduction of Cyber Threat Intelligence (CTI) has emerged as a new security system to mitigate existing cyber terrorism for advanced applications. CTI demands a lot of requirements at every step. In particular, data collection is a critical source of information for analysis and sharing; it is highly dependent on the reliability of the data. Although many feeds provide information on threats recently, it is essential to collect reliable data, as the data may be of unknown origin and provide information on unverified threats. Additionally, effective resource management needs to be put in place due to the large volume and diversity of the data. In this paper, we propose a blockchain-based cyber threat intelligence system architecture for sustainable computing in order to address issues such as reliability, privacy, scalability, and sustainability. The proposed system model can cooperate with multiple feeds that collect CTI data, create a reliable dataset, reduce network load, and measure organizations' contributions to motivate participation. To assess the proposed model's effectiveness, we perform the experimental analysis, taking into account various measures, including reliability, privacy, scalability, and sustainability. Experimental results of evaluation using the IP of 10 open source intelligence (OSINT) CTI feeds show that the proposed model saves about 15% of storage space compared to total network resources in a limited test environment.

**Keywords:** cyber threat intelligence; blockchain; sustainable computing; security

## 1. Introduction

The existing security system focused on restoring, minimizing, and recovering the damaged system after a cyber-attack. However, with the introduction of the concept of cyber threat intelligence (CTI), it is possible to respond to various attacks preemptively [1]. These attacks include the DDoS (Distributed Denial of Service) attack that paralyzes the system, Ransomware that extorts financial profits against companies, and the APT (Advanced Persistent Threat) attack on organizations in areas of high information value, such as government agencies and the financial industry [2]. CTI is a concept that protects against various cyber threats in multiple organizations so that data collection, analysis, and sharing can analyze previously known attacks and prepare for them. Most small and medium-sized companies, which are difficult to consume a lot of manpower and resources for security, are difficult to defend against numerous cyber-attacks, such as Ransomware, APT, and DDoS. For companies with limited security expertise and limited security budgets, it is useful to use CTI information sharing to help protect against cyber threats. According to a Sans study, about 81 percent of security professionals said the use of CTI improved their organization's security [3]. Because of this positive CTI effect,

standard technologies for research in the information security industry such as FBI and US-CERT were studied. These studies include Structured Threat Information eXpression (STIX) for the structured representation of cyber threat information, Trusted Automated eXchange of Indicator Information (TAXII), a protocol for sharing CTI information, and Cyber Observable eXpression (CybOX), integrated security information for expressing observation information about cyber threat status or events such as logs [4–6]. This structured cyber threat information enables consistent analysis and automated interpretation. Threat information-sharing methods such as TAXII are encouraged to be shared globally by various legal procedures and governments [7].

Today, CTI ultimately aims to automate all processes to minimize the manpower and resources required for security systems and preemptively respond to various attacks based on threat information [8,9]. However, many requirements have not yet been met to achieve this ideal purpose. The CTI data collection phase is the first important step and requires a reliable, sufficient amount of data [1,8–10]. Recently, there has been a growing number of data feeds that provide threat information. The amount of threat information collected has increased countless times, and enough data has been met for analysis [10,11]. However, the verification method of whether a large number of data from unknown sources is reliable data has emerged as a big issue [10,12,13]. In addition to reliability issues, these are referred to as the need for various stakeholder organizations to be able to respond flexibly to different systems and requirements [9]. If the wrong information is collected, the wrong information can be analyzed to not recognize the cyber threat, prevent the company from defending from the cyber threat, and waste resources through unnecessary analysis. Additionally, an insufficient amount of threat information may increase the false-positive rate of security solutions based on information analysis. It may not recognize attacks in case of infringement due to a lack of information. Besides, this information can lead to privacy issues [2]. Information shared by an organization can include personal information of specific users who have interests in the organization, or harm the organization's reputation, or competitors can use it [8–10]. In particular, the collection of network data is not only crucial for specifying attack groups, attack paths, and attack patterns, but also represents the movement of all data except for physically executed attacks, so that malware detection and analysis are also possible [2]. Additionally, because the amount of data collected in this way is too large, feed organizations that collect data may experience data storage problems. The average medium-sized organization reported 10 to 500 million events per day on their systems [13], while [12] noted 250 million events per day which create data saturation's problem. Thus, they need efficient and sustainable resource management. A method of using an external storage cloud server has been proposed to solve this problem [14]. However, using external storage cloud servers means a simple lease of storage space and is not a solution for data reduction. Using external cloud storage can cause problems, such as data integrity and privacy, because the CSP (Cloud Service Provider) manages physical access to the data [15]. Therefore, efficient management of resources requires the application of solutions that can reduce resources and lower communication overhead that occurs to take advantage of cloud storage.

Blockchain technology, based on Satoshi Nakamoto et al. [16] proposed an electronic currency that could be implemented in a P2P manner against a centralized financial system that had to be traded through a third party. This research can compensate for security vulnerabilities in a centralized system because multiple nodes have the same books distributed and can verify the data integrity between distributed nodes without a third party for verification. These advantages are being studied as decentralized system models using blockchain technology are not only financial systems, but also important industrial control systems (SCADA) and various Internet of Things (IoT) applications. In this paper, the architecture uses blockchain technology to efficiently process large data and providing security and privacy in a distributed way. The cloud server is used as a third party, which receives various distributed feeds from the feed layer. The cloud server's nodes and data feed node are utilized by blockchain. The cloud server nodes have two functions, block generation and verification data. The data feed node has only one function that verifies the data. So, both the cloud server and data feed node can verify the data in the blockchain network.

This paper proposes a blockchain-based cyber threat intelligence system architecture to address issues such as reliability, privacy, scalability, and sustainability. The scientific contribution of this research work is summarized as follows:

- We identify data collection issues and requirements that may arise during the data collection process for sharing threat information.
- We propose a blockchain-based cyber threat intelligence system architecture to address the limitations in the legacy system. The proposed model can obtain data in real-world scenarios from the information consumer perspective.
- We conduct a comparative analysis of the proposed model with the existing system based on the requirements of the data collection process.
- To assess the effectiveness of the proposed model, we perform the experimental analysis taking into account various measures. Experimental results show that the proposed model saves about 15% of storage space compared to total network resources in a limited test environment.

This paper is organized as follows. Section 2 describes the existing data collection methods at CTI and derives the requirements. We then look at studies conducted to (partly) alleviate these requirements. Section 3 proposes a blockchain-based sustainable system architecture to meet the various requirements of the data collection process. Section 4 analyzes and compares the existing research, and Section 5 concludes the paper.

## 2. Related Works

This section explains the concept of CTI, a new security system through information sharing, and presents the problems and requirements for mitigating them. It also describes various related studies conducted to mitigate these requirements.

### *2.1. Cyber Threat Intelligence*

CTI analyzes threat information and shares countermeasures to defend against cyber-attacks. Efficient sharing will not be possible when any organizations have different types of systems and data. Therefore, various studies are being conducted in CTI to standardize the sharing system of threat information that can occur in multiple organizational environments. STIX and TAXII are adopted and used as international standards by various information-sharing communities and organizations. Structured Threat Information eXpression (STIX) is a standardized language developed by MITRE and the OASIS CTI Technical Committee to describe cyber threat information [4,5]. STIX has been adopted as an international standard by various information-sharing communities and organizations. STIX enables the structure and automated analysis of threat information to automate the collection, analysis, and sharing of threat and CTI information. Information shared by CTI may include suspicious domain information, malicious file hash values, IP (Internet Protocol) addresses to identify attackers, URLs, domains, Network Artifacts, and attack tools used by attackers [8]. Trusted Automated eXchange of Intelligence Information (TAXII) is a threat information sharing protocol that shares cyber threat information through information exchange [6,7]. TAXII is designed to share STIX information and has three information-sharing models for sharing information. TAXII's information sharing is divided into three types of models: The Source-Subscriber method, which shares a single central information source, and the Hub and Spoke, Peer-to-Peer methods, which share information in a peer-to-peer manner. Hub and Spoke is a method of transmitting information to a single centralized organization in a Source-Subscriber format, which is a centralized single information source information-sharing model. Eventually, CTI improves security between organizations through a collection and analysis sharing cycle [17]. For example, a non-commercial or commercial organization collects information as malicious code itself, such as a domain or IP that can specify a file hash or attacker that converts raw data into information. Then, it shares the type of attack and the attacker's analysis to defend the attack in a specific manner (applying a security solution to the network architecture) with other users.

There are many ways to collect data to analyze threat information in the CTI concept. Data feeds can be shared outside the organization or collected within the organization to collect data [18]. Data obtained from outside the organization can be obtained from open-source intelligence (OSINT), security reports, or external feeds. Mohammed et al. [19] analyzed the hacker forum, which is easily accessible through OSINT, and predicted the frequency of cyber-attacks. Within an organization, it can be collected from SIEM (Security Information and Event Management) in the organization's network or from a core network that manages a centralized network such as Software Defined Networking (SDN) [20,21]. After this process, Feed analyzes, shares, and distributes threat information to information consumers based on the data it collects. However, stakeholder organizations that can benefit from commercial gains can make data appear large by creating false data because the sharing threat information is associated with the interests of the organization. Organizations that supply intelligence data are competing using threat data [8,10,17,22]. Generating random packet data to increase the amount of information shared is easily manipulated and difficult to recognize from outside. Similarly, data manipulation can be performed not only with packet data generation, but also with non-existent packet information to emphasize the specificity of information supplied by the organization. If an organization is attacked or provides inaccurate information with malicious intent, the systems of the organizations that trust the information may not function properly. For example, legitimate normal users can be distributed to attacker IPs from specific CTI feeds. This may be unreliable information or malicious purpose. Restricting IP using the CTI information received from the corresponding CTI Feed by network security policies may result in normal users not acquiring standard service and may have the same effect as DDoS [12]. In particular, this problem can increase the extent of the damage if the wrong information is applied on a large scale when the system is automated.

### *2.2. Requirements of Data Collection in CTI*

In the data collection phase, the requirements that must be met to protect against cyber threats can be summarized as follows:

Reliability: To prepare for cyber-attacks, CTI creates countermeasures against attacks by collecting and analyzing data. The collection of reliable data is used as a material to produce accurate analysis and countermeasures against cyber-attacks. However, false data can waste not only misanalysis, but also the computing resources needed for analysis. Many of the data collected are from unclear sources and difficult to verify if it is fabricated [1,9,10].

Privacy: Privacy issues arise when sharing data collected by an organization. It may be the user's personal information that should not be leaked outside the organization because it is considered an asset within the organization that shares the data [2]. For this reason, information sharing needs to be considered as a way of not directly exposing the data managed by the organization.

Scalability: There is a problem that the information provided may be different because the threat information required by the organization or any system is different [9]. Therefore, CTI must not only be able to respond to heterogeneous data, but also be flexible in terms of the type of attack and new data that changes as IT (Information Technology) evolves. It must be implemented efficiently as an integrated service rather than as many complex structures, such as a data analytics network, external cloud services to address massive data storage issues, and services for data sharing. This collaboration can help you achieve scalability as a base model for collaborative architectures that can benefit from improved data reliability, quality, and organizational incentive contributions.

Sustainability: A lot of data is required to create a countermeasure against cyber threats. In particular, the scale of network resources is vast, significantly affecting the data storage space of feeds collecting data. If there is a problem with storage due to inefficient resource management, it may not be possible to analyze relevant information. Efficient resource management of individual organization feeds can be a more pressing issue from the overall network perspective. With a sustainable environment such as collective actions, smart allocation, priority efficiency, and analytic opportunities with relevant data, each organization must manage its resources according to its purpose

effectively from the network perspective—the optimal management of sustainable data with improved sustainability performance in the network provided by the proposed architecture.

### *2.3. Existing Research*

CTI must be reliable and require large amounts of data to generate and analyze correct threat information. As data feeds are collected and shared recently, vast amounts of data for generating threat information can be satisfied. However, data shared by many data feeds can be easily manipulated.

Meier and Roland proposed FeedRank, a new ranking approach that aggregates reliability by comparing data from multiple feed organizations [22]. Poor-quality feeds can cause inaccurate analysis, and too few feeds will make it difficult to analyze threat information due to a lack of threat information. Feeds can easily manipulate the collected data for the performance of an organization, and it is difficult for shared users to know whether a feed is manipulated or not. Therefore, rather than verifying the reliability of feeds that collect data independently, it is necessary to verify the data shared by users.

Gong et al. [18] proposed a model that can analyze the reliability and validity of data using comparative analysis between CTI data to verify the reliability of OSINT (Open Source Intelligence) information that Feed collects externally. The study helps to find relatively reliable data feeds from an informed consumer perspective, and suggests a highly applicable model based on data comparison analysis.

Le et al. [23] propose a framework for collecting data from Twitter, a social media platform, to obtain appropriate information to counter cyber-attacks. The proposed framework detects threats by linking data collected from Twitter with Common Vulnerabilities and Exposure (CVE) identifiers. Koloveas et al. [24] proposed architecture for gathering cybersecurity information by crawling data at the Hackers Forum on the Dark Web using open source tools. These attempts have actively targeted and used frequency of data to collect information from specific organizations to enable reasonable inference against cyberattacks. In this case, it can be used to analyze different types of attacks in CTI Feeds that generate data rather than from an informed consumer perspective.

Threat information-sharing methods such as TAXII are encouraged to be shared globally by various legal procedures and governments [7]. For example, in 2013, the UK government launched a joint project to share cybersecurity, and in 2015, the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) established international standards to guide the exchange of sensitive security information. This standard is also used to implement information security management within the information-sharing community.

Malware Information Sharing Platform (MISP) is an open-source threat intelligence platform used by more than 6000 organizations worldwide. MISP serves as a threat sharing platform and community that automatically shares IoC (Indicators of Compromise) data collected by an organization's IDS (Intrusion Detection System) or SIEM to the platform [25].

According to a computer security collaborative management study by Jean-Marc Seigneur and Adam Slagell, cyber threat intelligence collaborative security is defined, "Instead of centrally managed policies, organizations or nodes share and collect cybersecurity knowledge with other organizations or nodes, and security decisions can be made" [26]. As a result, Daire Homan et al. [27] proposed a model for sharing data securely without sharing personal information when sharing sensitive information using blockchain.

Zhou et al. [28] have compiled several studies to collect network data. This study describes the data collection method by systematically classifying the data collection method generated in the network, and segmenting it by hardware and software. Networks need to consider the efficient processing of big data to design network architectures because massive data communication occurs [29–31].

Through the relevant research analysis, the data collection process focused on identifying targets and using the frequency of data to deduce cyberattacks [23,24], or on reliability verification for most organizations [18,22]. However, it is difficult to verify reliability for organizations that manage their data. Liu et al. [10] said that communication between different communities that currently use different

sharing standards is complex and requires an integrated system in which collaborative information sharing is efficient. From an information consumer perspective, it is important to have a reliable dataset to improve the security of your organization. Thus, the approach proposed in this paper proposes a model that must be able to secure direct reliability of data, not organization, and that can benefit the Data Feeds that provide data through an efficient process to obtain data from information consumers.

## 3. Proposed Blockchain-Based CTI System Architecture

This chapter proposes a model for preventing false information that has been tampered with for the benefit of Feed, which provides threat information to organizations that share CTI threat information and follow response strategies. In the process of collecting threat information on the network, we describe the proposed model to confirm that the information is actually communicated. False threat information shared by data feed organizations can share false security response strategies with information-sharing security organizations, making their security policies vulnerable and inefficient. To cope with this, we propose a blockchain-based CTI network architecture. The advantages of the proposed model are as follows: (1) Even if a reliable data feed sends wrong information, it can prevent the wrong information because it verifies the received information. (2) It can efficiently manage a large amount of information and provide a secure environment with the help of a distributed way and decentralized system configuration. (3) It manages distributed ledger using blockchain, it can guarantee the integrity of feed sharing information, and reward according to the organization's contribution.

### *3.1. Design Overview*

A large number of specific data collected can mean that it is highly likely that the data originated in the actual network. Furthermore, it can be determined that it is essential data with a high frequency of data generation. The frequency of these data can be used to prevent any particular data feed from generating fake data by reducing data's importance. In this paper, in order to judge the credibility of the collected data, several nodes participate and hand over the data of organizations to third party organizations without invading personal information. The third party's organization checks the number of data transmitted by the Feeds and determines that the accumulated data can be trusted.

Figure 1 shows the architecture proposed in this paper to collect network information from the network reliably. Device Layer is a layer where network data is generated, and feed is an organization that collects and shares data. Finally, Cloud Server (CS) receives data from the feed and indirectly secured the data (meaning to keep the abbreviated data set?), and service it to the information consumer.

The network architecture begins at the device layer where network packet data occurs. In the Device Layer, network packets, such as IoT devices for smart homes and cloud services, as well as computers and mobile equipment, are increasing at a large scale. Therefore, the network packet data collection method can be divided into two types. It can be divided by reducing the number of data collected by randomly extracting packets from all traffic, or by grouping packets into groups. Network packets contain a lot of information that can generate threat information. Indicator information (IP, Domain, URL, or C & C server information) and malicious file hash values are collected as relevant threat data [8,18]. In this case, the IP information may be used as indicator information afterward, such as Domain, URL, and Network Artifacts information. Therefore, in the architecture proposed in this paper, file information and IP information are grouped and collected.

Feed Layer processes the collected network information by grouping. The grouping of data refers to packet data, including file information and identifying indicator information. At this time, IP information, a network identification indicator, can specify attackers' identification, along with various additional identification information, such as domains, URLs, and network artifacts. It is because the scalable data variability can be exploited through the cumulative aggregation of identification information to create a reliable dataset. Collected data goes through processing to be processed like other information. The processing extracts the File Hash Value and the IP address, which is the identifier data, to check the integrity of the information indirectly. Then, this data is sent to Cloud

Server. The Cloud Server cannot know the original data by using the information sent by the feed, but can determine whether the original data sent from different Feeds is the same. At this time, it is determined that the more information transmitted from the feed, the more reliable the data is.

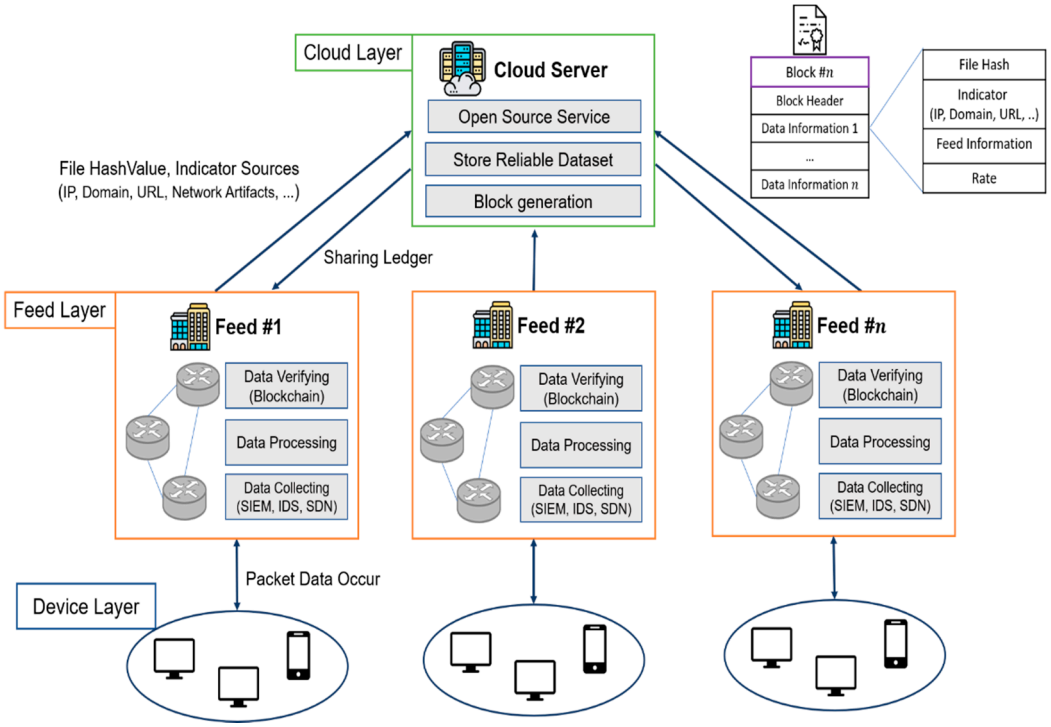


**Figure 1.** Proposed Cyber Threat Intelligence (CTI) system architecture for reliable data collection.

Cloud Server's Cloud Server (CS) is a third-party independent system that processes data received from distributed feeds. CS can secure its transparency through open-source, but there is a problem of network load and data storage space because a lot of data is concentrated. To solve this problem, CS reduces network load through low data transmission rate and minimum communication frequency, and stores information only for indicating data integrity and the importance of data, not original data. The cloud server's node and data feed node used in the Blockchain have two functions that block generation and verification of data in the blockchain network. However, the data feed node has only one function, that of verification of data. Blockchain provides a secure environment with the data integrity of the feed and blocking possibility being tempered via cyber-attacks in the feed. Blockchain technology has a variety of properties, including data integrity, decentralization, transparency. Additionally, through the blockchain, the contributions of feeds that contributed to the service (that is, sharing a lot of reliable data) can be checked to reward the feeds in the service policy and to derive the motivation for the feeds to cooperate. Feeds can contribute in two ways: By providing exclusive information and by providing contributions to information that already exists.

Issuing blocks in CS rather than feed is because of the efficiency to reduce the network load caused by the generation of a lot of collected data. This private blockchain is connected between CS and participating feeds, and the block information is designed so that private information is not stored. We can also consider the encryption method using the public key and private key so that only the feed that owns the data in the blockchain can access the data, but this may cause unnecessary network overhead. In particular, the service provider of CS can be solved by providing service response to the corresponding feed, the information owner.

### 3.2. The Methodological Flow of the Proposed Architecture

The methodology for the proposed model is a process that describes the process of generating blocks from the point of time when data communication occurs in the network as reliable data in the Cloud Layer.

According to Figure 2, the operation process can be explained in five steps as follows.

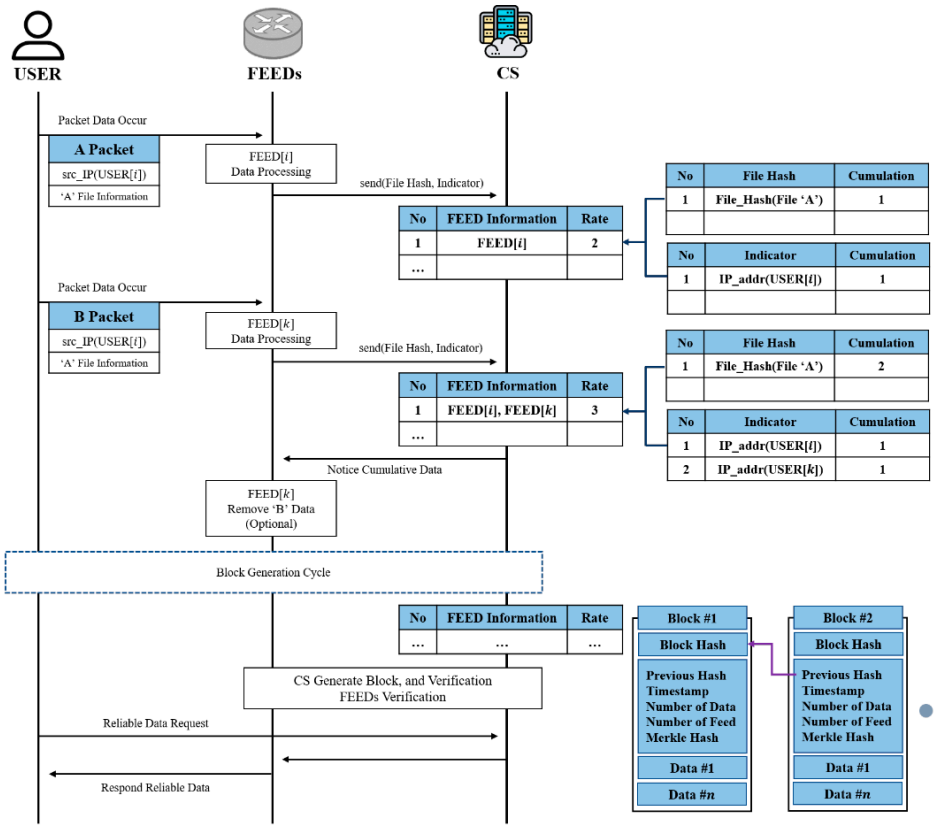


**Figure 2.** The methodological flow of the proposed architecture.

(1) Gathering Data: Data is generated from the network from the gathering data device layer. Collecting all data during the data collection process is inefficient and eliminates the need to collect unnecessary and repetitive data that operates on its system that occurs in large-scale IoT systems. Therefore, collect packet with file hash and IP address that can be threat information. At this time, collect Network Resource from Core Network (SIEM, IDS, SDN) from feed collecting data. Unnecessary traffic is generated, so packet data is collected from the organization's core network.

(2) Processing: File Hash Value and IP Address are extracted by pre-processing the data collected by the Processing Feed. The data preprocessing process transmits only the file information and indicator information, which is important information that can be identified as threat information, to the CS, except the unnecessary information included in the packet. File Hash Value uses the SHA-256 function because it needs to use the hash function of the same type as different feeds.

(3) Create Dataset: By sharing the File Hash value to the CS, it is possible to transmit information data without exposing the data of its original packet. Since CS has only Hash Value, original data cannot be obtained. It checks whether the same data is compared with Hash Value sent from other feeds. CS manages a table that stores File Hash Value and Indicator, and ranks reliability by cumulation of File Hash Value and Indicator. This network Indicator can be applied by aggregating IP information as well as information such as domains, URLs, and network artifacts. Formula (1) is the aggregation priority of data stored in CS.

$$Rate(D_i) = \alpha(D_i) + \beta(D_i) \quad (1)$$

The data priority for a specific data $D_i$ is the sum of the maximum value of the cumulation of the File Hash Table and the maximum value of the Indicator Table. The higher the sum of the rates, the more important data is judged to be important data. These formulas can be found in Formulas (2) and (3).

$$\alpha(D_i) = \sum_{k=1}^{n} \mathrm{F_k(D_f)} \tag{2}$$

$$\beta(D_i) = \sum_{k=1}^{m} \mathrm{F_k(D_{id})} \tag{3}$$

CS determines the number of feeds that verify specific data to determine the reliability of specific data. By verifying the files and verifying the indicators, the data consumers are informed of the data used for the most communication, and the information consumers can obtain the information in the order of the data with the highest importance and reliability verification. Table 1 describes the symbols used in the formula.

**Table 1.** Symbols used in formula.

| Symbol | Explanation |
|---|---|
| $D_i$ | The specific *i*-th data |
| $\alpha$ | Maximum value of File Hash Table Cumulation |
| $\beta$ | Maximum value of Indicator Table Cumulation |
| $D_{id}$ | Indicator Data |
| $D_f$ | File hash Data |
| $F_n$ | The specific *i*-th Feed |

(4) Remove Cumulative Data: The CS informs the feed that sends the same information in the process of creating its own data set as duplicate data. This process is to solve the problem of repeatedly storing a large number of data generated from several distributed feeds. Feed collects data after packet data is generated from Device Layer, and this data is transmitted to contribute to CS. At this time, if the specific information $D_i$ transmitted by the feed is the first information not presented in the CS data set, the CS stores the corresponding data in the data set. If the CS stores $D_i$, the CS informs the feed transmitting the data that the data is duplicated. Feeds know that their own data is owned by other feeds, so they don't need to store redundant data without discarding it. Through this process, each of the feeds distributed has only non-overlapping data, so that the larger the number of participating feeds that provide information, the more computing resources for data storage can be saved.

(5) Block Generation: Blockchain is used to protect centralized CS data and to reward organizations for contributing to creating reliable data sets. Blockchain creation, unlike general public blockchains, creates blocks in CS, a centralized institution for efficiency in the network architecture environment where big data communication takes place. The CS generates blocks every 10 min with accumulated data collected by the CS, that is, a reliable data set. The CS checks the contributions of the feeds every 10 min, the block generation cycle. At this time, the contribution rate gives 3 points to the feed that provided the first information and 1 point to the feed that provided the second information. Therefore, if the contribution that provided the first information is x and the y after the second, the organization's contribution is given by Formula (4).

$$Contribution = \frac{3\mathrm{x} + \mathrm{y}}{Total\_Contribution} * 100\% \tag{4}$$

Table 2 shows the components that the block stores. Blockchain ensures data integrity by connecting the *n*-th previous hash value with the *n*-th Block Hash value. Block Hash is a hash of all information in a block as input. Data Num represents the number of data sets according to the CS service policy, and Feed Num represents the number of feeds participating in the block generation.

Data of the block stores File Hash, Indicator, Data Owner Feed information, and Reliability Score, which are data set information that CS provides to information consumers. At this time, the feed information is represented by a unique identification number issued while participating in the private blockchain. These blockchains are distributed and stored among participating organizations.

**Table 2.** Block composition data.

| **Block Num** | **13** | **Block Hash** | **bc1abd0b888e6636aad892451** |
|---|---|---|---|
| Previous Hash | 382bc3dce38f86d52c34af93 . . . | Timestamp | Mon 01 06 2019 16:34:52 (KST) |
| Data Num | 500 | Feed Num | 78 |
| Data | No:1/Feed:7/01e0d128e0651e5e512103dea55690409edc4942210f946bceccc61 . . . | | |
| | No:2/Feed:52/92e7202e4703a048acf3ddebe66c3d8a49fcc9af5c52a24bf1dca6 . . . | | |

Information consumers can check the reliability of the collected data, not the reliability of the company, so that all the data of the relatively high-reliable company are not shared, but the data that has been secured can be partially imported by the organization. Acquiring this partial data was difficult in the feed organization that shared the data, because the process was complicated and there was no compensation system. However, the architecture presented in this paper contributed to the reliable collection of data from the blocks of feed and CS. The company's information is recorded so that it can reward the company, and this reward system can form a cooperative system of various organizations.

### *3.3. Service Scenario*

This sub-section describes a service scenario where an information consumer requiring CTI information receives reliable data. An information consumer is a security expert for improving the security of an individual or organization.

Figure 3 illustrates the service scenario from an information consumer perspective. Information Consumers (User) receive reliable data in two ways. The first method is to collect network data (IP, Domain URL, File Hash) that has not been analyzed. The second method is to collect threat information analyzed by external data analysis feed.

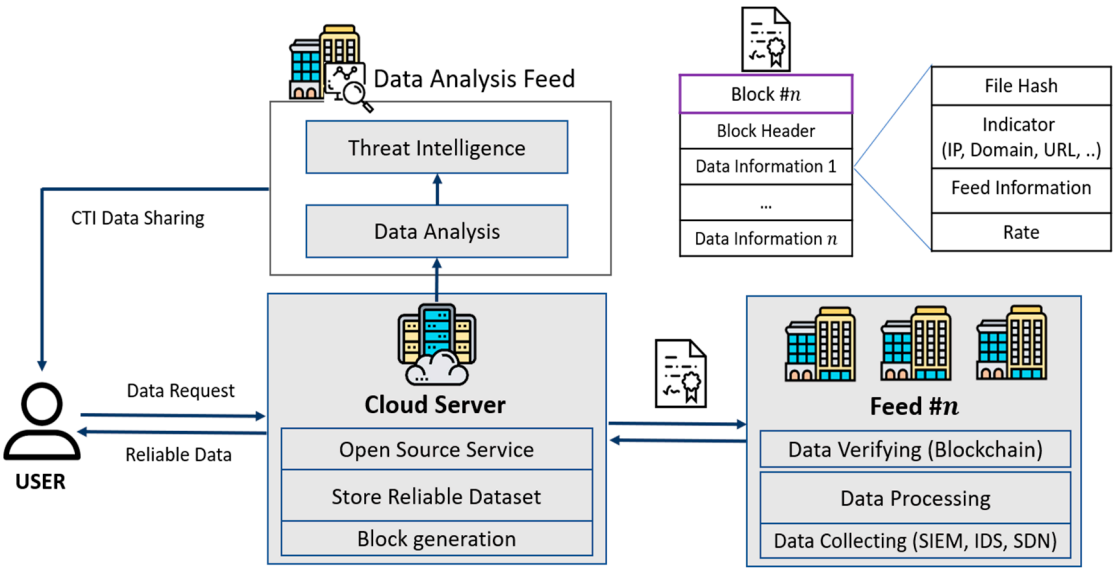


**Figure 3.** An information consumer perspective's service scenario.

Information consumers can receive reliable network data sets to analyze data on their own or to verify threat information from outsiders. In this case, the information consumer requests data from the Data Management Center. The Data Management Center stores only the data information of several

feeds, and requests the original data from the feeds in the order of reliable data. Then, it delivers the received original data to the information consumer.

Information consumers may collect only the threat information analyzed based on the data whose reliability has been verified instead of analyzing itself. Data Analysis Feed receives reliable data from the Data Management Center and derives threat information by analyzing the data. Data Analysis delivers verified threat information to information consumers.

## 4. Analysis and Discussion

This section compares the previous research to analyze the model presented in this paper, and explains the requirements from the data collection stage. Finally, we consider various cyber threats from the model point of view.

### *4.1. Comparative Analysis*

This sub-section explains the differences of the model presented in this paper by comparing it with related studies to secure the reliability of existing CTI data. Existing research studies such as OSINT (Open Source Intelligence) Method, Feed Rank Method, and Gathering CTI Method from Twitter are available. However, these methods have some limitations, including Reliability, Privacy, Scalability, and Sustainability. To mitigate these limitations, we proposed Blockchain-based Cyber Threat Intelligence System Architecture for Sustainable Computing.

- Reliability: There are two kinds of reliability of data handled in related studies. There are two methods, (1) verifying the data provided by a trusted organization, and (2) verifying the reliability of data itself. The problem in the first case is that the reliability of the institution is relatively judged. If logical reliability is assured, then the information the organization distributes will be trusted. However, a relatively reliable institution can be interpreted to mean that it can be manipulated at any time. Therefore, direct verification of the data is needed rather than the reliability of the institution. In this paper, it is possible to verify the data itself because distributed feeds judge reliability by using only information about data transmitted through cooperation.
- Privacy: The privacy issue is that data collected inside the organization is not leaked to the outside. This may be the leakage of data about internal users, or to prevent damage to the organization by exposing the organization's resources collected outside. Therefore, basically, the collected data is stored and managed only in the organization so that the information inside the organization is not leaked to the outside by using a Hash Function that verifies the data and increases the organization's contribution, but does not restore the original data. Even if CS does not leak information to other competitive feeds and CS leaks data through cyber-attacks, CS information alone does not risk leaking information inside the organization.
- Scalability: In the CTI concept, file hashes are used as direct evidence of an attacker's attack behavior and are very important information. Similarly, indicator information for identifying an attacker can be obtained through network information such as an IP address. However, the attacker's indicator can be modified in various ways. For example, IP tampering with IP spoofing can recognize an attacker as if it were a normal user. For this reason, using an IP address with other information rather than using it as independent information can improve identification. The indicator mentioned in this paper uses the only IP address, but the Indicator Table of CS can be flexibly changed by adding various identification elements such as Domain, URL, and Network Artifacts in packet information collected by feed.
- Sustainability: For efficient resource management of any organization feeds, there are various requirements such as collective actions, smart allocation, and analytic opportunities with relevant data sources. So, the proposed architecture provides a sustainable environment and optimizes data at the cloud layer with improved sustainability performance in the network. Both cloud server and data feed node are used in the blockchain network and have different functionalities

for providing sustainable infrastructure with smart allocation using collective actions properly. Comparison between existing research studied with proposed architecture is shown in Table 3.

**Table 3.** Comparison between existing research studies with proposed architecture.

| Research | Reliability | Privacy | Scalability | Sustainability |
| --- | --- | --- | --- | --- |
| Open source intelligence (OSINT) Method [18] | It ensures reliability in an open source-based data collection stage and verifies the feed organization's trustworthiness and the trustworthiness of the shared data itself. | It is not a matter of leaking personal information or violating information resources between feeds with this method. | This study has a high dependency on the data collected by the feed to obtain the data's reliability. | This method didn't use sustainable environment |
| Feed Rank Method [22] | It evaluates the credibility of the organization by evaluating and ranking the reliability of feed organizations, but did not secure the reliability of data. | This study does not cover privacy issues within the feed organization because it only uses information from feeds. | The relative comparison is made based on the feeds' data, so it can flexibly cope with the necessary change in different information. | It is use-only feed rank, not sustainability. |
| Gathering CTI Method from Twitter [23] | This method collects CTI data using Common Vulnerabilities and Exposure (CVE) identifiers, and the CTI data source must be premised, so it isn't easy to verify the reliability. | It does not follow a privacy issue because it collects information from social network platforms associated with CTI information from a consumer perspective. | It is a framework that extracts only specific external data through data crawling. Therefore, it has low scalability to collect various extended data. | It didn't consider sustainability |
| Proposed Model | The architecture presented in this paper has high reliability because it verifies data itself, not organization through the cooperation of feeds. | The feed that collects data is powerful in privacy problems because it uses an only hash value that can indirectly verify data without leaking original data. | It manages important information extracted from packet data and indicator table, which enables higher-level data collection and flexible coping. | It provides optimal management of sustainable data with improved sustainability in the network. |

### *4.2. Experimental Evaluation*

This sub-section collects OSINT, and analyzes and evaluates it by applying the proposed model. The three elements for evaluating the architecture proposed in this paper are as follows:

- Create a reliable data set through the cooperation of feeds that provide data.
- Efficient resource management is available in a vast network environment.
- According to CS policy, feeds participating in cooperation can be rewarded by measuring contribution.

In this paper, threat information is applied to the proposed architecture for the research question on these three factors and experimented with. For the experiment, we crawled and collected OSINT data from the following feeds. Feed: Abuse, Bambenek consulting, Blocklist.de, Emerging Threat, FireHOL, GreenSnow, IPsum, MalSilo, Mirai Tracker, Snort. We tried to collect the identifier IP and FileHash (SHA-256), an important malware resource, to identify the target on the network. However, FileHash data was not provided selectively, but was provided in a passive way to search for malicious code in files through input. In the process of establishing a cooperative environment using multiple feeds, it was difficult to obtain a hash value of a large amount of SHA-256 files, so this experiment uses only IPv4 Data. Table 4 is a list of data sets provided from CTI feeds.

**Table 4.** OSINT CTI data resources for experiment.

| CTI Feed | IPv4 Resources | SHA-256 Resources | Remarks |
|---|---|---|---|
| Abuse | 1453 | X | F1 |
| Bambenekconsulting | 162 | X | F2 |
| Blocklist.de | 766 | X | F3 |
| Emerging Threat | 778 | X | F4 |
| FireHOL | 404 | X | F5 |
| GreenSnow | 3902 | X | F6 |
| IPsum | 3077 | X | F7 |
| MalSilo | 585 | O | F8 |
| Mirai tracker | 1000 | X | F9 |
| Snort | 1115 | X | F10 |
| Total | 13242 | | |

We created a reliable data set using 10 feeds (F1, F2, . . . , F10) and a total of 13,242 IPv4 information. We believe that the more feeds we verify, the more reliable data it is. That is, CS stores the cumulative index of the same information that feeds send for specific data. When CS collects information on specific data, it can be divided into two cases. The first is when the CS receives independent data that the CS does not have, and the second is when the CS receives duplicate data for the data it already has. The first independent data received is called 'First Information', and three points are provided to the feed that provided the data. The data that is not is called 'Second Information', and one point is provided to the feed that provided the data. This scoring system is only applicable when recorded in a dataset created by CS. That is, if the data contributed by the feed is highly reliable, the contribution is provided.

Table 5 is a CS data set produced using the OSINT data set. To create a reliable dataset of the indicator, the Cumulation of $D_i$ of certain data can be calculated by the same number of indicators out of the $D_n$ of the 10 CTI feeds. The first feed of Feed Information is the First Information provider, and the other feeds become the Second Information provider. In the case of No1 data, F3 provided the First Information, and the collected Indicator was 59.10.5.156. As the F4, F6, and F7 feeds provided the Second Information, the Cumulation value became 4, which is the most reliable. It is difficult to know the first provided Information because there is information that does not include the timestamp in the OSINT resource obtained for this experiment. Therefore, it is assumed that a feed with a low sequence number is a First Information provider. In this experiment, the number of cooperating feeds and data is very small, so the reliable Cumulation Value is low. Only No1 data had a Cumulation Value of 4, and others had Cumulation Value 3 of 118, 2 of 1790, and non-overlapping data of 9304.

**Table 5.** Dataset created through experiments.

| No. | Feed Information | RATE | Indicator (IPv4) | Cumulation |
|---|---|---|---|---|
| 1 | F3, F4, F6, F7 | 4 | 59.10.5.156 | 4 |
| 2 | F3, F6, F7 | 3 | 103.27.238.202 | 3 |
| 3 | F3, F4, F7 | 3 | 104.236.72.187 | 3 |

Since the network generates a large amount of data, it is necessary to efficiently manage CTI resources. The Second Information provision feed generated while creating a reliable dataset does not need to be stored because the First Information provision feed already stores data and serves through CS. That is, Second Information does not store data. This can effectively manage the resources of the feeds by reducing the Second Information of each feed, in terms of the overall network.

Figure 4 shows that network resources decrease as the number of cooperating feeds increases. The Number of Feed on the $x$-axis means 1 for F1 feed and 2 for F1 + F2 feed. As the number

of cooperative feeds is less, the occurrence of Second Information is less, so the data reduction is insignificant, but the more the cooperation occurs, the more the resource reduction can be expected.

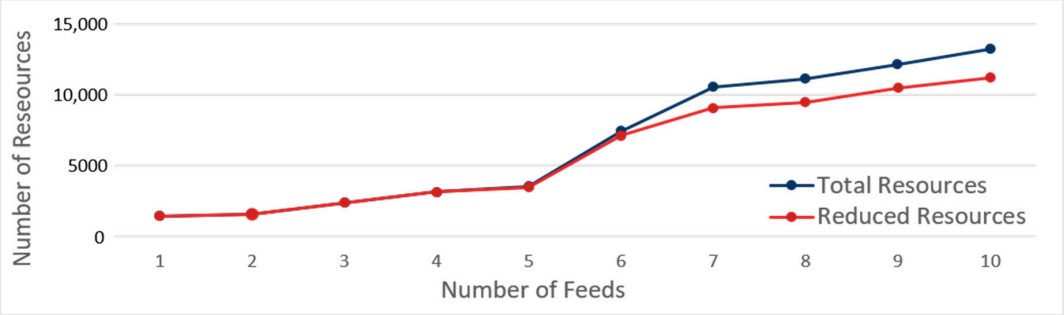


**Figure 4.** OSINT CTI data resources for experiment.

Figure 5 shows the resource reduction when the number of cooperating feeds is more than six. As the number of feeds increases, the total number of network resources increases, and data reduction can be seen by comparing the number of feeds to 6 and 10. These results show that the larger the system, the higher the resource reduction. As a result, when all 10 feeds that the experimental dataset is applied, the data resources are reduced by about 15.322% from a total of 13,242 data to 11,213 from a total network perspective.

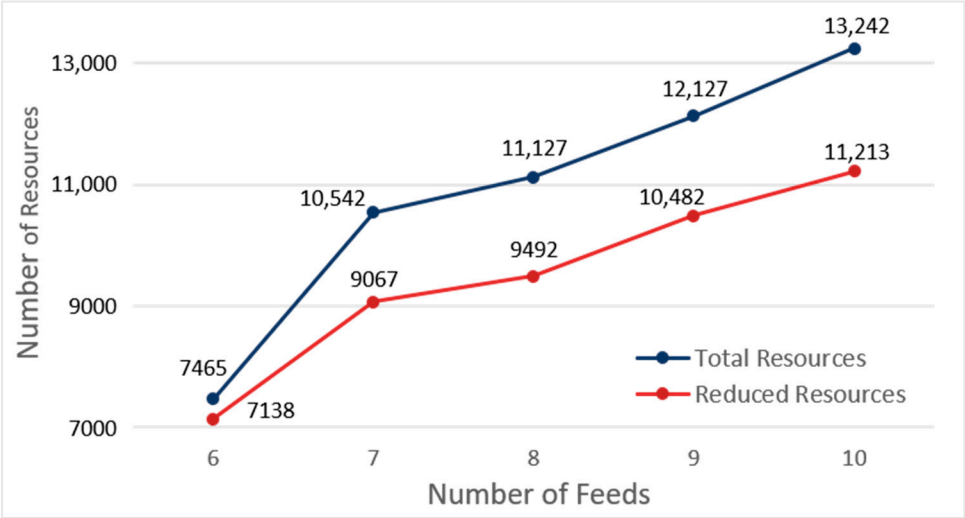


**Figure 5.** OSINT CTI data resources for experiment.

Figure 6 shows the experimental results for each feed that provided the OSINT data. The indigo bar means the total number of data collected from individual feeds, and the red bar means the number of data reduced when the model of this paper is applied. Finally, the green line represents the contribution score contributed by the cooperative system. Contributions can be used as an indicator to distribute rewards in future cooperative systems. In the case of F1 feeds, there are no reduced resources compared to the number of resources collected, but because of the large number of first information, more contributions are scored compared to the amount of data contributed. On the other hand, F9 feeds have very low resource reductions, as in F1, but they have not contributed to improving data quality with other feeds, so they have very low contributions.

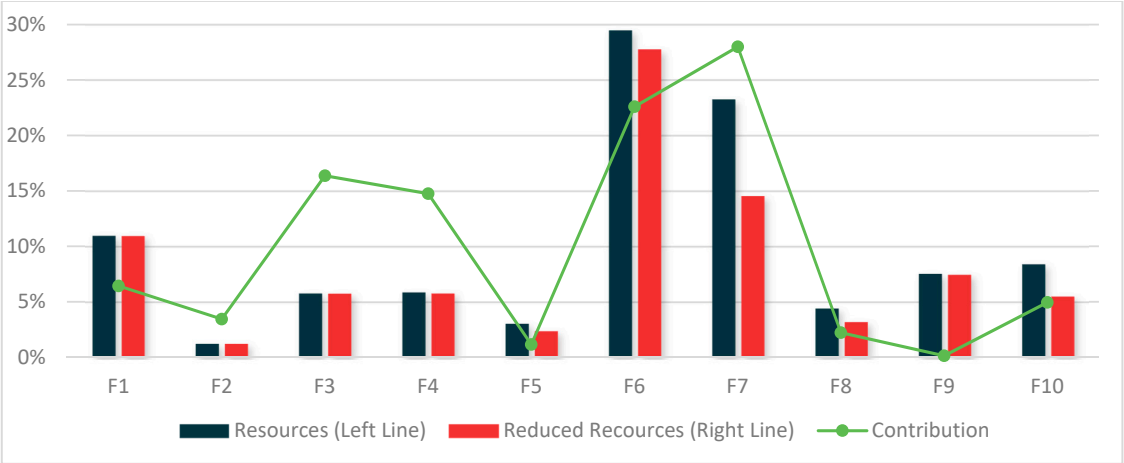


**Figure 6.** Summary information of feeds through experiment.

Table 6 shows the experimental values for all feeds from F1 to F10. Feeds at the beginning have little or no declining resources, as they become First Information providers in ascending order according to the order of the feeds. In other words, the feeds at the beginning have high First Information and low Second Information, so the resource efficiency is low, but the contribution is high. Conversely, high-order feeds tend to have low First Information and high Second Information, so they tend to have high resource efficiency instead of relatively low contribution. Although the experiment in this paper required large-scale data through cooperation of large-scale feeds, the amount of data used in the experiment was extremely small because OSINT data was collected and tested. According to the experimental results, it is expected that the reliability of the data and the resource efficiency of the CTI feed will be further increased in a large-scale data environment in which data is highly duplicated.

**Table 6.** Detail information of feeds through experiment.

| CTI Feed | Number of Resources | Reduced Resources | Second Resources | First Resources | Contribution (%) |
|---|---|---|---|---|---|
| F1 | 1453 | 1453 | 0 | 166 | 6.42 |
| F2 | 162 | 162 | 0 | 89 | 3.44 |
| F3 | 766 | 766 | 0 | 423 | 16.36 |
| F4 | 778 | 768 | 10 | 378 | 14.75 |
| F5 | 404 | 315 | 89 | 0 | 1.15 |
| F6 | 3902 | 3674 | 228 | 508 | 22.59 |
| F7 | 3077 | 1929 | 1148 | 341 | 27.99 |
| F8 | 585 | 425 | 160 | 4 | 2.22 |
| F9 | 1000 | 990 | 10 | 0 | 0.13 |
| F10 | 1115 | 731 | 384 | 0 | 4.95 |

### *4.3. Discussion*

Data Copy: By copying data from other feeds, you can consider vulnerabilities that increase the reliability of your data and maximize your organization's contribution. However, copying data is impossible because it does not expose the data of other feeds. In addition, the cloud that serves the data is made public after the data block is created, so it is impossible to copy the data [32–34].

Centralization: CS is logically centralized to receiving data from feeds, which can lead to security vulnerabilities. First of all, it is a feature of CS that operates in an open-source manner, so it is impossible to manipulate CS and feed. Likewise, if a CS is cyber-attacked and an attacker attempts to manipulate any feed organization in the CS for economic benefit, the information possessed by the CS cannot provide reliable data because only the hash value of the original data is present [35–37]. However,

even if the feed belongs to reliable data, the information consumer who accesses the data cannot obtain the information from the feed that does not have the data [38].

Hash Collision: The model presented in this paper uses a hash function to communicate efficiently in a network environment where a lot of data is generated and communicated. Since the hash function has a problem of hash collision, which is an inherent problem, this problem may occur in the proposed model. First, time is not taken into account in the data collection stage for sharing threat information. Time to provide threat information is not an important consideration because it is intended to use threat information to change an organization's security policy and protect against future cyber threats [39–42]. Therefore, the model proposed in this paper provides service after creating a dataset every 10 min, which is a block generation cycle. The purpose of the data collected for 10 min is to store information that can be the most threat from the information consumer's perspective. Therefore, although there is a possibility of a hash collision among the recent data occurring for 10 min, it is very unlikely that both of the two files with the conflict will be highly reliable. However, when sending a hash value to solve the hash collision, if the transmission includes additional information such as the size of the file data, the possibility of a hash collision can be much lower. However, most feeds are rarely collecting both files that accidentally crashed as important. For example, when data B, having a rate of 1, collides with data A, having a rate of 50, the CS recognizes the same hash and determines it to be 51. In this case, we considered the efficiency of data communication because it does not mean whether A has a rate of 50 or 51 as a big change.

## 5. Conclusions

In this paper, we analyzed the requirements for meeting the ideal purpose of CTI and studied the limitations that are currently faced. While the existing feed collects, analyzes, and shares data independently and internally to share threat information, the proposed architecture works with multiple feeds to verify the reliability of the data shared during data collection. Because of the credibility of individual data, not institutional, information consumers can gather data in the order of credibility, regardless of institution. This cooperation is achieved through the use of blockchain to confirm their contributions and to be rewarded in the service policy. The model proposed in this paper is a collaborative architecture that mitigates problems such as data collection reliability, privacy, efficiency, and scalability. Our experiments have shown that we can reduce network resources in terms of collaborative networks. The scalability of this architecture can be applied as a basic model at the CTI data collection stage. The proposed model saves about 15% of storage space compared to total network resources in a limited test environment. In future research, we will study an extensible framework that can efficiently and automatically apply security policies of information consumer organizations through data analysis as well as data collection.

**Author Contributions:** Investigation, J.C. and S.K.S.; writing—review & editing, J.C.; writing—original draft, J.C.; methodology, J.C.; validation, J.C. and S.K.S.; resources, J.C.; visualization, J.C.; formal analysis, J.C.; supervision, J.H.P.; project administration, J.C.; funding acquisition, Y.P. and J.H.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the Advanced Research Project funded by the SeoulTech (Seoul National University of Science and Technology).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Tounsi, W.; Helmi, R. A survey on technical threat intelligence in the age of sophisticated cyber attacks. *Comput. Secur.* **2018**, *72*, 212–233. [CrossRef]
2. Shin, B.; Lowry, P.B. A review and theoretical explanation of the 'Cyberthreat-Intelligence (CTI) capability'that needs to be fostered in information security practitioners and how this can be accomplished. *Comput. Secur.* **2020**, *92*, 101761. [CrossRef]

3. Brown, R.; Robert, M.L. *The Evolution of Cyber Threat Intelligence (CTI): 2019 SANS CTI Survey*; SANS Institute: Singapore, 2019.
4. Barnum, S. Standardizing cyber threat intelligence information with the Structured Threat Information eXpression (STIX). *Mitre Corp.* **2012**, *11*, 1–22.
5. Casey, E.; Back, G.; Barnum, S. Leveraging CybOX™ to standardize representation and exchange of digital forensic information. *Digit. Investig.* **2015**, *12*, S102–S110. [CrossRef]
6. Connolly, J. Davidson, M.Schmidt, C., Ed.; The Trusted Automated Exchange of Indicator Information (Taxii)The MITRE Corporation: Bedford, MA, USA; McLean, VA, USA, 2014; pp. 1–20.
7. Skopik, F.; Settanni, G.; Fiedler, R. A problem shared is a problem halved: A survey on the dimensions of collective cyber defense through security information sharing. *Comput. Secur.* **2016**, *60*, 154–176. [CrossRef]
8. Schaberreiter, T.; Kupfersberger, V.; Rantos, K.; Spyros, A.; Papanikolaou, A.; Ilioudis, C.A.; Quirchmayr, G. A quantitative evaluation of trust in the quality of cyber threat intelligence sources. In Proceedings of the 14th International Conference on Availability, Reliability and Security, Canterbury, UK, 26–29 August 2019.
9. Wagner, T.D.; Mahbub, K.; Palomar, E.; Abdallah, A.E. Cyber threat intelligence sharing: Survey and research directions. *Comput. Secur.* **2019**, *87*, 101589. [CrossRef]
10. Liu, M.; He, X.; Chen, J. Cyberthreat-intelligence information sharing: Enhancing collaborative security. *IEEE Consum. Electron. Mag.* **2019**, *8*, 17–22. [CrossRef]
11. Khan, T.; Alan, M.; Akhunzada, A.; Hur, A.; Asif, M.; Khan, M.K. Towards augmented proactive cyberthreat intelligence. *J. Parallel Distrib. Comput.* **2019**, *124*, 47–59. [CrossRef]
12. Griffioen, H.; Tim, M.B.; Christian, D. Quality Evaluation of Cyber Threat Intelligence Feeds. In Proceedings of the 19th International Conference on Applied Cryptography and Network Security, Kanagawa, Japan, 21–24 June 2021.
13. Afzaliseresht, N.; Miao, Y.; Michalska, S.; Liu, Q.; Wang, H. From logs to Stories: Human-Centred Data Mining for Cyber Threat Intelligence. *IEEE Access* **2020**, *8*, 19089–19099. [CrossRef]
14. Serckumecka, A.; Medeiros, I.; Bernardo, F. A Cost-Effective Cloud Event Archival for SIEMs. In Proceedings of the 38th International Symposium on Reliable Distributed Systems Workshops (SRDSW), Lyon, France, 1–4 October 2019.
15. Sookhak, M.; Tang, H.; He, Y.; Yu, F.R. Security and privacy of smart cities: A survey, research issues and challenges. *IEEE Commun. Surv. Tutor.* **2018**, *21*, 1718–1743. [CrossRef]
16. Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System. 2008. Available online: https://bitcoin.org/bitcoin.pdf (accessed on 20 April 2020).
17. Groš, S. Research Directions in Cyber Threat Intelligence. *arXiv* **2020**, arXiv:2001.06616.
18. Gong, S.; Cho, J.; Lee, C. A Reliability Comparison Method for OSINT Validity Analysis. *IEEE Trans. Ind. Inform.* **2018**, *14*, 5428–5435. [CrossRef]
19. Almukaynizi, M.; Grimm, A.; Nunes, E.; Shakarian, J.; Shakarian, P. Predicting cyber threats through the dynamics of user connectivity in darkweb and deepweb forums. *ACM Comput. Soc. Sci.* 2017. Available online: https://usc-isi-i2.github.io/papers/kristina02.pdf (accessed on 4 August 2020).
20. Cinque, M.; Cotroneo, D.; Pecchia, A. Challenges and directions in security information and event management (SIEM). In Proceedings of the 2018 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW), Memphis, TN, USA, 15–18 October 2018.
21. Mokalled, H.; Catelli, R.; Casola, V.; Debertol, D.; Meda, E.; Zunino, R. The Applicability of a SIEM Solution: Requirements and Evaluation. In Proceedings of the IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), Napoli, Italy, 12–14 June 2019.
22. Meier, R.; Scherrer, C.; Gugelmann, D.; Lenders, V.; Vanbever, L. FeedRank: A tamper-resistant method for the ranking of cyber threat intelligence feeds. In Proceedings of the 10th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 29 May–1 June 2018.
23. Le, B.; Wang, G.; Nasim, M.; Babar, A. Gathering Cyber Threat Intelligence from Twitter Using Novelty Classification. In Proceedings of the International Conference on Cyberworlds (CW), Kyoto, Japan, 2–4 October 2019.
24. Koloveas, P.; Chantzios, T.; Tryfonopoulos, C.; Skiadopoulos, S. A 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.

25. Wagner, C.; Dulaunoy, A.; Wagener, G.; Iklody, A. Misp: The design and implementation of a collaborative threat intelligence sharing platform. In Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security, Vienna, Austria, 24 October 2016.
26. Seigneur, J.; Seigneur, J.-M.; Slagell, A. *Collaborative Computer Security and Trust Management*; Adam, S., Ed.; Information Science Reference: UK, 2010.
27. Homan, D.; Shiel, I.; Thorpe, C. 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 June 2019.
28. Zhou, D.; Yan, Z.; Fu, Y.; Yao, Z. A survey on network data collection. *J. Netw. Comput. Appl.* **2018**, *116*, 9–23. [CrossRef]
29. Lin, H.; Yan, Z.; Chen, Y.; Zhang, L. A Survey on Network Security-Related Data Collection Technologies. *IEEE* **2018**, *6*, 18345–18365. [CrossRef]
30. Sharma, P.K.; Moon, S.Y.; Park, J.H. Block-VN: A distributed Blockchain based vehicular network architecture in smart city. *J. Inf. Process. Syst.* **2017**, *13*, 184–195. [CrossRef]
31. Park, J.; Park, J.H.; Salim, M.M.; Jo, J.H.; Sicato, J.C.S.; Rathore, S.; Park, J.H. CIoT-Net: A scalable cognitive IoT based smart city network architecture. *Hum.-Cent. Comput. Inf. Sci.* **2019**, *9*, 29.
32. Singh, S.K.; Rathore, S.; Park, J.H. BlockIoTIntelligence: A Blockchain-enabled Intelligent IoT Architecture with Artificial Intelligence. *Future Gener. Comput. Syst.* **2019**. [CrossRef]
33. Lee, Y.; Rathore, S.; Park, J.H.; Park, J.H. A blockchain-based smart home gateway architecture for preventing data forgery. *Hum.-Cent. Comput. Inf. Sci.* **2020**, *10*, 1–14. [CrossRef]
34. Yin, C.; Zhou, B.; Yin, Z.; Wang, J. Local privacy protection classification based on human-centric computing. *Hum.-Cent. Comput. Inf. Sci.* **2019**, *9*, 33. [CrossRef]
35. Zhang, J.; Zhong, S.; Wang, T.; Chao, H.C.; Wang, J. Blockchain-based systems and applications: A survey. *J. Internet Technol.* **2020**, *21*, 1–14.
36. Gu, K.; Yang, L.; Yin, B. Location Data Record Privacy Protection based on Differential Privacy Mechanism. *Inf. Technol. Control* **2018**, *47*, 639–654. [CrossRef]
37. Singh, S.K.; Jeong, Y.S.; Park, J.H. A deep learning-based IoT-oriented infrastructure for secure smart City. *Sustain. Cities Soc.* **2020**, *60*, 10225. [CrossRef]
38. Singh, S.K.; Rastogi, N. Role of Cyber Cell to Handle Cyber Crime within the Public and Private Sector: An Indian Case Study. In Proceedings of the 2018 3rd International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), Bhimtal, India, 23–24 February 2018; pp. 1–6.
39. Yin, C.; Ding, S.; Wang, J. Mobile marketing recommendation method based on user location feedback. *Hum.-Cent. Comput. Inf. Sci.* **2019**, *9*, 14. [CrossRef]
40. Liu, M.; Cheng, L.; Qian, K.; Wang, J.; Wang, J.; Liu, Y. Indoor acoustic localization: A survey. *Hum.-Cent. Comput. Inf. Sci.* **2020**, *10*, 2. [CrossRef]
41. Wang, J.; Gu, X.; Liu, W.; Sangaiah, A.K.; Kim, H.J. An empower hamilton loop based data collection algorithm with mobile agent for WSNs. *Hum.-Cent. Comput. Inf. Sci.* **2019**, *9*, 1–14. [CrossRef]
42. Li, T.M.; Chao, H.C.; Zhang, J. Emotion classification based on brain wave: A survey. *Hum.-Cent. Comput. Inf. Sci.* **2019**, *9*, 42. [CrossRef]