Towards Data Assurance and Resilience in IoT Using Blockchain

Xueping Liang[1,2,3], Juan Zhao[3], Sachin Shetty[4], Danyi Li[1]†
[1]State Key Laboratory of Information Security,
Institute of Information Engineering, Chinese Academy of Sciences, Beijing, 100093, China
[2]University of Chinese Academy of Sciences, Beijing, 100190, China
[3]College of Engineering, Tennessee State University, Nashville, TN 37209
[4]Virginia Modeling Analysis and Simulation Center, Old Dominion University, Norfolk, VA 23529

***Abstract*—Data assurance and resilience are crucial security issues in cloud-based IoT applications. With the widespread adoption of drones in IoT scenarios such as warfare, agriculture and delivery, effective solutions to protect data integrity and communications between drones and the control system have been in urgent demand to prevent potential vulnerabilities that may cause heavy losses. To secure drone communication during data collection and transmission, as well as preserve the integrity of collected data, we propose a distributed solution by utilizing blockchain technology along with the traditional cloud server. Instead of registering the drone itself to the blockchain, we anchor the hashed data records collected from drones to the blockchain network and generate a blockchain receipt for each data record stored in the cloud, reducing the burden of moving drones with the limit of battery and process capability while gaining enhanced security guarantee of the data. This paper presents the idea of securing drone data collection and communication in combination with a public blockchain for provisioning data integrity and cloud auditing. The evaluation shows that our system is a reliable and distributed system for drone data assurance and resilience with acceptable overhead and scalability for a large number of drones.**

***Keywords*-Reliability, Secure communication, Date Assurance, Blockchain, Drone, Auditing, Resilience, Integrity**

I. Introduction

The rapid development of Internet of Things (IoT) is starting to transform how we live [1]. As more physical devices such as mobile phones, wearable devices, and vehicles are connecting to the Internet through embedded systems and sensors, large amounts of data can be collected and sent to the cloud computing system to conduct data analysis for a better and faster decision making. Moreover, these devices can perform commissions and tasks that humans can not accomplish. For example, unmanned aerial vehicles, also known as drones, which is a microcosm of IoT, performing wide-ranging activities from delivering a package to tracking crop quality and finding farm anomalies [2].

However, as IoT grows, the connectivity is increasing, and the computing infrastructure will become more complex, opening up more vulnerabilities for the cyber attack. Some of the physical devices are located in unsecured environments and easily tampered by hackers. More of the data and the operation commands traveling over through wireless sensor network to the Internet, an untrusted communication channel, are likely to be modified. Therefore, device authorizations and data provenance [3] [4] would be a critical issue. Moreover, many existing IoT systems rely on centralized communication models to connect to servers or cloud computing that support processing and data storage. The problem is that the server will become a bottleneck and a new target for cyber-attack, as well as a point of failure that will disrupt the entire network and impact the data integrity. Therefore, how to build a truly trusted and integrated environments to support this connected devices and computing infrastructure to transfer data remains challenged.

To solve above challenges, we propose using BlockChain for the IoT architecture. BlockChain (BC) technology that originates from Bitcoin, the first crypto-currency system launched in 2008 [5], can provide an effective solution to IoT privacy and security, due to its three foundational tenets: 1) data in the blockchain is stored in a shared, distributed and fault-tolerant database that every participant in the network can share the ability to nullify adversaries by harnessing the computational capabilities of the honest nodes and information exchanged is resilient to manipulation, 2) Blockchain is a decentralized architecture to make the architectures robust against any failures and attacks and, 3) Blockchain relies on public key infrastructure which allows the contents to be encrypted in a way that is expensive to crack. With blockchain-based IoT architecture, all data operations are transparently and permanently recorded. Thus, the trust between devices and backend servers can be established.

The main contribution of this paper is that we propose a trusted and resilient architecture for IoT service based on Blockchain, which provides the ability for self-trust, data integrity audit and data resilience, as well as scalability. To exemplify our idea, we use the scenario of a cloud-based drone system in the rest of the paper. The reason is that drone is a very typical microcosm of IoT, where drones collect data from embedded sensors and cameras, and also receive the commands from remote control systems. However, the architecture is application-agnostic for diverse IoT use cases.

†Corresponding author.

978-1-5386-0595-0/17/$31.00 ©2017 IEEE

The proposed architecture fulfills the following objectives:

- **Trusted Data Origin.** By binding the device ID with the collected data from each drone and submitting to the blockchain network, the data origin is defined in a deterministic way, so that we can know which device sends the data, where the device is, and whether the traffic the device supposed to send is believable for that type of device. This is critical for data provenance and assurance.
- **Instant and Permanent Data Integrity.** The control system that receives the data collected by drones will submit a hash of each data record to the blockchain network instantly. The record will be included in a block as a transaction. The integrity of the record is guaranteed by the consensus mechanism used in the block mining process. Distributed node confirmation preserves the integrity and provides resistance against tampering.
- **Trusted Accountability.** Each control command from the control system or the cloud server is accountable by uploading the operation records to the blockchain network. This gives each operation a fingerprint, which makes every action traceable. Once anomaly is detected, the malicious entity can be identified for further investigation.
- **Resilient Backend.** Each distributed node in the blockchain network maintains a copy of the entire ledger, preserving the availability and persistent performance to make the architecture robust and resilient for any potential failures and attacks. Both collected data and control data are integrity protected and thus trusted. Moreover, the trust between blockchain nodes is removed, reducing the possible losses that a compromised node could cause. The distributed nature of blockchain nodes adds to the availability of both data and data validation, making it an on-demand service with no downtime. Meanwhile, the auditing service offered by the server makes the data collection system accountable in attack mitigation and cyber forensics.

The rest of the paper is organized as follows. Section II discusses the proposed architecture. Details of implementation are discussed in Section III. Analysis and evaluation is presented in Section IV. Section V presents some related work, while Section VI concludes the paper and talks about the future work.

II. System Architecture

A. System Overview

We present an architecture for blockchain-based drone system named as DroneChain, shown in Figure 1. DroneChain consists of four main components, which are drones, control system, cloud server, and blockchain network. One or more drones can form a drone cluster to perform complex commissions. Control system can be assigned and interact with a drone cluster for data collection and commission distribution. The cloud server provides the capacity of storage for the large mounts of data collected by drones and provide real time data processing and data analysis to facilitate further decision making. The blockchain, a decentralized network, is used for data validation and resilience.

B. Key Components

- **Drone.** Drones are capable of using sensors to collect various physical data such as soil composition and moisture content, and using embedded cameras to capture the image or video for the field. Sometimes, a number of drones comprise a drone cluster for delegation in an assigned area. Drones need to communicate with a control system, either directly, or through a representative drone, to send out the collected data or the flight status data and receive the commands.
- **Control System.** Control system is responsible for receiving collected data from drones or drones clusters and send out commands to adjust the flight movement of drones, as well as some other operations and physical behaviors. The control system serves like a indeterminate to aggregate the collected data from drones, hash the original data for integrity protection, and then send both original data and hashed data to both the blockchain network and the cloud. The same process will be launched on the command data from the control system itself, making a permanent record for monitoring the control system.
- **Blockchain Network.** The blockchain network may be used for three purposes. For data collected from drones and the commands from control system, each of the hashed data entry is uploaded to the blockchain network for integrity protection and could be stored in a distributed manner which ensures stability. Besides, each of the feedback from the cloud server and the database access activity will also be recorded on the blockchain for further auditing or investigation. Not only the data records are permanently stored, but also a blockchain receipt will be generated for data validation.
- **Cloud Database.** The cloud database stores the original data collected from drones, commands sent by the control system and data access from cloud server and auditor. Data access is made accountable by a fingerprint for each access, which is also stored in the database. A daemon process will look up each data entry in both the database and the blockchain network for consistency check periodically. Both object data and command are traceable. Once data leakage or intrusion is detected, the malicious entity can be found and identified.
- **Cloud Server.** Cloud server handles data from drones and data access records, which serves the same purpose as the blockchain network will do. To validate data record, the cloud server is responsible for requesting to the blockchain network for a blockchain receipt as a permanent proof of data integrity. Moreover, by adopting machine learning techniques, the back end system in the cloud can retrieve timely and explicit feedback for further manipulation of the drone, such as drone path navigation. Meanwhile, the analysis of the communication data be-

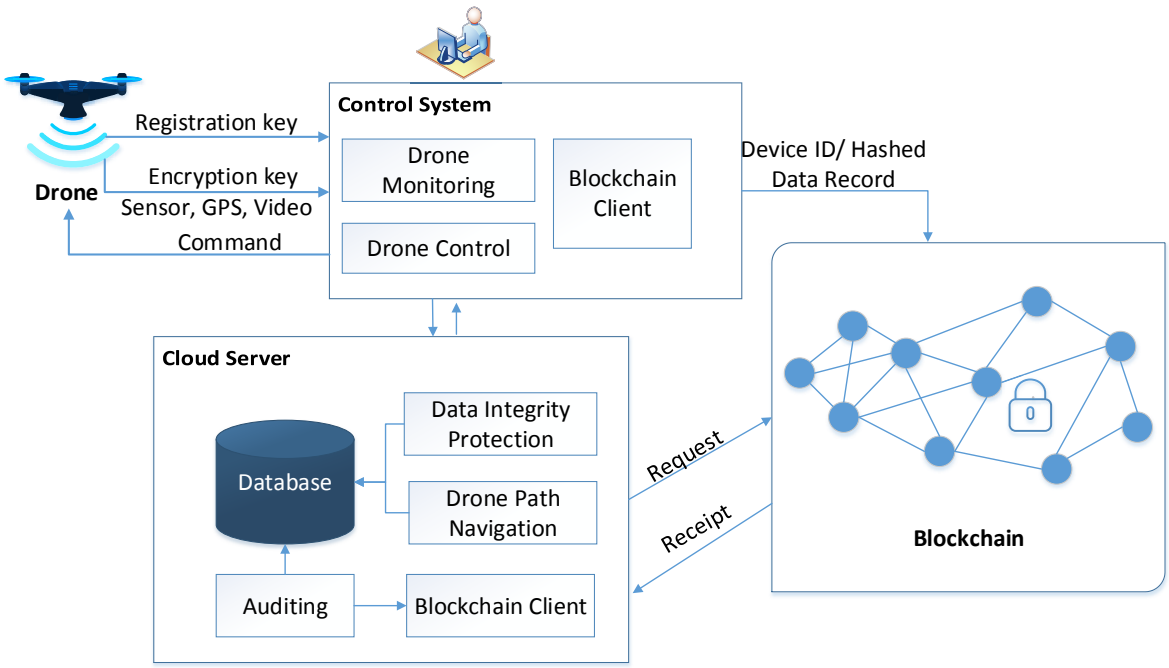

Fig. 1: DroneChain: Blockchain-based Drone Communication Architecture.

tween the drone and control system as well as the server can help detect the intrusions such as advanced persistent threat (APT) [6] or distributed denial of service (DDoS) attack [7] in early stages.

C. Threat Model

To build a secure-aware architecture for drone data collection and communication, we analyze the potential vulnerabilities in implementing DroneChain. The cloud server maintains a database to cache collected data from drones but cannot guarantee that data records will remain unchanged due to known vulnerabilities in cloud operating systems. Once DroneChain is enabled, the cloud server will be able to track the data, and the auditor will be allowed to access all the collected data and data operations, as well as control commands and drone monitoring data. However, the auditor cannot be completely trusted. The adversary can potentially access or modify collected data. Since DroneChains main objective is to protect the integrity of drone data, we assume that data is encrypted and stored, which is not accessible to anyone without the decryption key.

D. Key Establishment

In the drone data collection system, drones are required to register and commissioned before job assignment to be involved in the system. For cloud data storage, data encryption key pairs will be required to encrypt sensitive data for confidentiality. We describe each of key as follows.

- **Drone Registration Key** K_{DR}. Drone needs to be commissioned to the system to store data collected while moving in the cloud database. We denote the key as K_{DR}. Every time new data record generated, the registration key is needed. Similarly, the registration key for the control system is K_{CR}.
- **Data Encryption Key** K_{DE}. After registration, the drone generates an encryption key K_{DE}, for encrypting all the data. When a data entry is created, drone encrypts the data entry, which limits the data access only to valid key holders. Each time there is a data entry created, the hashed data entry will be recorded on the blockchain.
- **Data Access Public/Private Key Pair** (PK_{DM}, PR_{DM}). For data access, a public/private key pair will be generated, denoted as (PK_{DM}, PR_{DM}). For some cases that the data access activity is to be recorded on the blockchain, the private key is used to generate a fingerprint from the operator to indicate the data origin, while the public key is used by others to verify the stated origin.

III. DroneChain Implementation

In the system, there are seven phases of drone data collection and transmission among the five essential entities, including drone registration, data and command transmission, blockchain receipt generation, cloud data validation, auditing and decision making.

A. *Drone Registration*

In our system, drone needs to enroll as a node for storing data collected from a certain location. After registration, the data collection phase starts and a unique ID will be assigned

to each drone. Every data record will be associated with the device ID. The data type may contain some measurements, videos or images. For simplicity, we consider the data as an object and hash each data record for efficiency before uploading to the blockchain network. The original data is stored in a local database at the same time for future lookup.

B. *Data and Command Transmission*

Each time there is a data record collected from the drone, the data entry can be constructed as a tuple {*DeviceID, Time, Location, Data*}. After the tuple is sent to the controller, the controller will forward the data to blockchain network. At the same time, it will send back some commands based on the data and task. The commands will also be recorded on the blockchain, using the tuple {*ControllerID, Time, Location, Command*}.

C. *Blockchain Receipt Generation*

Once a collected data record from a drone is uploaded to the blockchain network via the controller, the event will be captured as a blockchain transaction. This provides the data management system with an ability for future validation, tracking and auditing. The record is hashed and eventually transformed into a Merkle tree node [8] using Tierion API [9]. The Merkle tree root node will be anchored in a blockchain transaction following the Chainpoint 2.0 protocol [10]. The use of Merkle tree offers the scalability which satisfies the vast throughput from large numbers of drones. A set of data records will be batched together as a transaction in the blockchain. A list of the transactions will be used to compose a new block, which will be confirmed by blockchain nodes. When the block is validated, it will be added to the existing blockchain, making it part of a tamper-resistant ledger.

The blockchain receipt contains information of the blockchain transaction and the Merkle proof used to validate the transaction. An example receipt is shown in Figure 2.


```
{
  "@context": "https://w3id.org/chainpoint/v2",
  "type": "ChainpointSHA256v2",
  "targetHash": "c74cb7b19193ee52b195f89e77c46e01fd4f79b6ab42b81524ab3c746246d559",
  "merkleRoot": "24b24cbf8d3ba0658605a8652e360bd15e8da916d35e3783abae37988ffb206a",
  "proof": [
    {
      "right": "cec1aacfcd506b6bcd964d5496b6f3b219bcd9c535bd1a6b2cd51d7a80749de8"
    },
    {
      "right": "8885b423dfe58b59f0ccc59c04e5b7d03e093f4de01cf42ef5164ebc2a411c98"
    },
    {
      "left": "243af84165465ca65e8d4df3cafef8b2ae5144414ca911af42ec37c0f12e444d"
    }
  ],
  "anchors": [
    {
      "type": "BTCOpReturn",
      "sourceId": "5f0b9d5eb087af62fa10a970553cd88d2e4a8d841cfae7e16b4bff4a9c04132f"
    }
  ]
}
```


Fig. 2: An example Blockchain Receipt for Data Record.

D. *Cloud Data Validation*

Since each record will be stored in the cloud instantly, the data integrity can be verified at any time. By periodically requesting the blockchain network for a blockchain receipt, every record will be validated by comparing the calculated hash with the targetHash from the generated blockchain receipt. If an inconsistency is detected, the record could be suspected for compromise. A daemon process in the cloud server is configured to make a request via Tierion API, as a proof of integrity, using the following URL.

`https://api.tierion.com/v1/records/<id>`

The request header should include *Content-Type: application/x-www-form-urlencoded* or *Content-Type: application/json* to set the data format and the requests to the Data API is protected by HTTPS. The data records are validated with the input including targetHash, merkleRoot and the proof from the blockchain receipt. The most important step is to reconstruct the Merkle tree from the blockchain receipt to compute the Merkle root. Each data record is stored together with other records in the blockchain network as one transaction. The proof part of the receipt indicates the relationship between each record from the same transaction. For example, the left node means its record is collected earlier than the record anchoring in the right node. The transaction attribute height represents the block index, and we can find the exact block information in Block Explorer [11]. To validate the format and contents of a blockchain receipt, and to confirm that the Merkle root of one record is stored in the blockchain, the following URL provided by Tierion API is used.

`https://api.tierion.com/v1/validatereceipt`

E. *Data Auditing and Decision Making*

With the cloud data available for validation, data auditing and decision making can be launched based on the trusted data set. The data records are stored in a time-based order and are accountable with a trusted data origin. Depending on the application scenarios of drones, either in a synchronized or asynchronized way. Data auditing is critical for detecting anomaly based on the command records from the control system and the cloud server. Based on the auditing results, effective decisions can be made to prevent and mitigate APT attacks or DDoS attacks.

IV. System Evaluation

A. *Security Analysis*

DroneChain integrates blockchain technology with drone-based IoT applications, offering a secure drone communication architecture and providing data assurance, resilience and accountability. The control system is an intermediate entity between the drone and the cloud server, and also between the drone and the blockchain network, responsible for forwarding and hashing the data collected from drones. The commands along with the drone data will be anchored to the blockchain network for integrity protection using blockchain receipts. By binding the device ID and location information, the data source is trusted by an unalterable fingerprint. The cloud server hosts a database for real-time processing and provides persistent data availability. Moreover, the server has the capability to integrate auditing module to inspect data and command records,

and drone path navigation module to dynamically adjust the aviation. The trusted data and command records contribute to the accountability of system components. By securing data process and distributing data process flow, high level of data assurance and resilience is preserved.

The completely decentralized architecture in public blockchain network helps to provide robustness and tamper resistance for data assurance, fully aligning with the IoT requirements, where nodes are equally distributed and participating. However, DroneChain is not designed for certain scenarios where nodes are assigned different roles and capabilities, in which case private blockchain is needed. DroneChain adopts proof of work as the consensus algorithm to generate the blockchain receipts while the drone and the control system are oblivious of this mining process. By maintaining the cloud database along with blockchain network in the back end, we balance the load with an acceptable latency, which will be evaluated in the following subsection.

B. Performance Evaluation

To test the performance of DroneChain, we build a prototype to simulate the data collection and data transmission process. The collected data is uploaded from the control system to the blockchain network. In this paper, we focus on the performance and overhead of the cloud server and the blockchain based data integrity protection. The evaluation environment setup includes the server, data collection application, and a benchmarking tool. The specifications of software and the version used are listed in Table I.

TABLE I: Evaluation Environment Specification

Software	Name	Version
Server Operating System	Ubuntu	14.04
Web server	Apache server	2.4.6
Database	MariaDB	5.5.44
Performance benchmarking	Apache JMeter	3.2

Apache Jmeter [12] is an open source and Java-based software designed to test server functions and behaviors at a large scale. The provided data analysis and visualization plugins allow great extensibility. We build a test plan to measure the performance of DroneChain. Our test plan aims to simulate the action of uploading collected data to the blockchain using hashing algorithms. The simulation also uses random numbers to represent data content collected by the drone. The test plan contains one controller to generate HTTP POST request to the server. A different number of drones and different size of data are simulated to test the scalability of DroneChain.

Figure 3 shows the average response time of DroneChain for data transmission application with a varying number of drones, with the data size of 64 Byte. It shows that the average response time increases linearly and thus provides better scalability.

Figure 4 shows the average response time with a varying size of data, with 100 drones. It shows that the average response time increases linearly and thus provides better process capability for different data size to be collected.

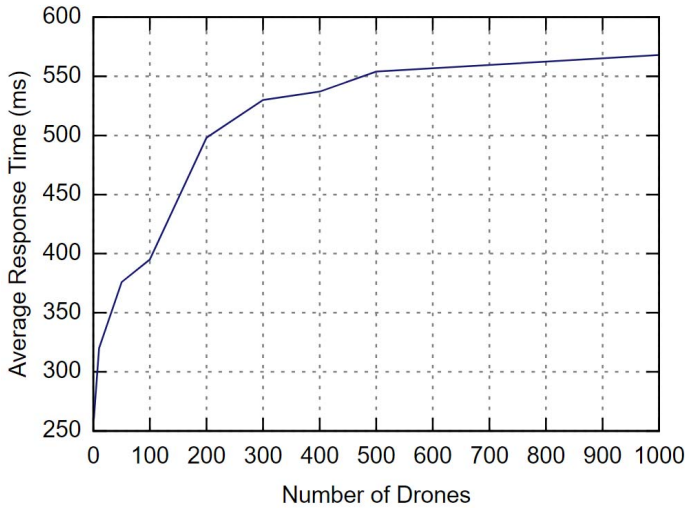

Fig. 3: Average Response Time of DroneChain with a Varying Number of Drones.

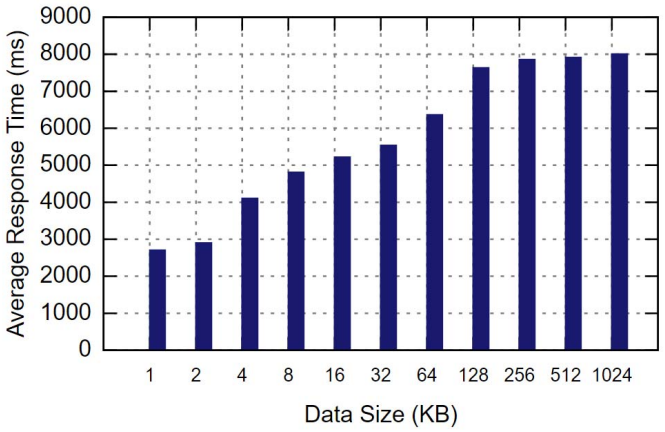

Fig. 4: Average Response Time of DroneChain with a Varying Size of Data.

Figure 5 shows the average latency of DroneChain for data transmission from 100 drones. It shows that the average response latency is relatively stable in the observed time range.

V. Related Work

A. IoT security

The Internet of Things (IoT) is growing rapidly, and the number of connected devices will foreseeably exceed several billion. One major problem is the embedded devices are inherently vulnerable to attacks on software and operating systems, such as buffer-over-flow attack. A compromised devices may leak the private information or send faked data to servers. Self-trust or data provenance could be an issue. Some study

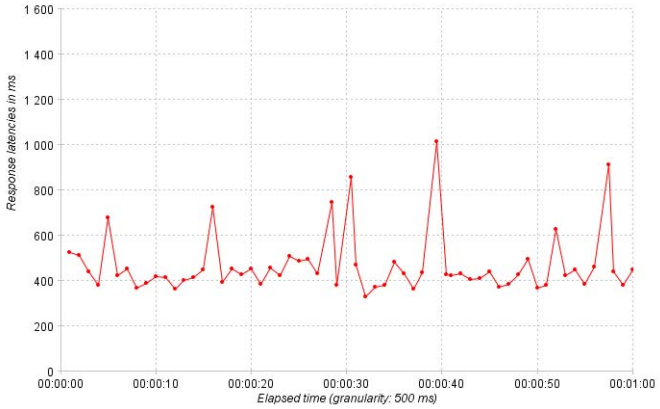

Fig. 5: Average Response Latency of DroneChain.

proposed using hardware security primitives and PUFs to solve the problem [13]. However, these techniques could make a high load in the devices, and the data integrity in the back ends are not guaranteed. Some research work focuses on inspecting network traffic load in IoT to detect attacks such as denial-of-service(DoS) [14], which did not fundamentally solve the data integrity and resilience problem. A study in [15] demonstrated that a wide variety of off-the-shelf IoT devices lack fundamental security considerations.

B. Blockchain and IoT

Blockchain has attracted peoples' interests due to immutable, shared, distributed ledger. BC is an attractive technology for addressing the mentioned security and privacy challenges as a result of its key features including decentralization, anonymity and security. However, there are still several challenges that need to be addressed for IoT devices [16]. One major challenge is the consensus algorithms in Blockchain, known to cause a delay as Proof of Work. Recently, there is some work in using blockchain in Cloud computing, and only few work focus on IoT devices. [17] proposed an optimized lightweight blockchain combining a private ledger in IoT local networks with the public ledger for smart home devices, but the integrity of the private local ledger is not preserved.

VI. Conclusion and Future Work

The drone has the potential to be widely adopted and leveraged in future IoT applications with its capability to sense and deliver in a less limited range of locations. In this paper, we propose a general architecture for drone data collection and control using blockchain, making it a step closer to such a vision that drone-based applications can collect sensor data and be controlled in a trusted and dependable way while reducing potential attacks and data losses. This system is capable of providing reliability and accountability, as well as data assurance for real-time data collection and drone control. We implement a prototype of the drone system, and the evaluation shows that the performance is acceptable. In the future, we will extend the system to Hyperledger Fabric platform [18], a private blockchain where nodes require permissions to participate in the blockchain mining, for a broader coverage of DroneChain in IoT applications.

VII. Acknowledgements

This work was supported by Office of the Assistant Secretary of Defense for Research and Engineering (OASD (R& E)) agreement FA8750-15-2-0120. The work was also supported by a grant from the National Natural Science Foundation of China (No.61402470) and the research project of Trusted Internet Identity Management (2016YFB0800505 and 2016YFB0800501).

References

[1] M. N. Aman, K. C. Chua, and B. Sikdar, "Secure data provenance for the internet of things," in *Proceedings of the 3rd ACM International Workshop on IoT Privacy, Trust, and Security*, ser. IoTPTS '17. New York, NY, USA: ACM, 2017, pp. 11–14.

[2] N. Joubert, "Looking forward to unmanned aerial systems and the internet of things," University Lecture, 2016.

[3] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat, and L. Njilla, "Provchain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability," in *International Symposium on Cluster, Cloud and Grid Computing*. IEEE/ACM, 2017.

[4] X. Liang, S. Shetty, L. Zhang, C. Kamhoua, and K. Kwiat, "Man in the cloud (mitc) defender: Sgx-based user credential protection for synchronization applications in cloud computing platform," in *The 10th IEEE International Conference on Cloud Computing (CLOUD 2017)*, 2017.

[5] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.

[6] Y.-H. Kim and W. H. Park, "A study on cyber threat prediction based on intrusion detection event for apt attack detection," *Multimedia tools and applications*, vol. 71, no. 2, pp. 685–698, 2014.

[7] K. Lee, J. Kim, K. H. Kwon, Y. Han, and S. Kim, "Ddos attack detection method using cluster analysis," *Expert Systems with Applications*, vol. 34, no. 3, pp. 1659–1665, 2008.

[8] R. C. Merkle, "Protocols for public key cryptosystems," in *Security and Privacy, 1980 IEEE Symposium on*, April 1980, pp. 122–122.

[9] "Tierion api," https://tierion.com/app/api.

[10] "Chainpoint: A scalable protocol for anchoring data in the blockchain and generating blockchain receipts," http://www.chainpoint.org/.

[11] "Bitcoin block explorer," https://btc.com/.

[12] "Apache jmeter," jmeter.apache.org/.

[13] A. Kanuparthi, R. Karri, and S. Addepalli, "Hardware and embedded security in the context of internet of things," in *Proceedings of the 2013 ACM Workshop on Security, Privacy & Dependability for Cyber Vehicles*, ser. CyCAR '13. New York, NY, USA: ACM, 2013, pp. 61–64. [Online]. Available: http://doi.acm.org/10.1145/2517968.2517976

[14] S. Misra, P. V. Krishna, H. Agarwal, A. Saxena, and M. S. Obaidat, "A learning automata based solution for preventing distributed denial of service in internet of things," in *Proceedings of the 2011 International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing*, ser. ITHINGSCPSCOM '11. Washington, DC, USA: IEEE Computer Society, 2011, pp. 114–122. [Online]. Available: http://dx.doi.org/10.1109/iThings/CPSCom.2011.84

[15] V. Sivaraman, H. H. Gharakheili, A. Vishwanath, R. Boreli, and O. Mehani, "Network-level security and privacy control for smart-home iot devices," in *2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct 2015, pp. 163–167.

[16] D. Tosh, S. Shetty, X. Liang, C. Kamhoua, K. Kwiat, and L. Njilla, "Security implications of blockchain cloud with analysis of block withholding attack," in *17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)*, 2017.

[17] A. Dorri, S. S. Kanhere, and R. Jurdak, "Towards an optimized blockchain for iot," in *Proceedings of the Second International Conference on Internet-of-Things Design and Implementation*, ser. IoTDI '17. New York, NY, USA: ACM, 2017, pp. 173–178. [Online]. Available: http://doi.acm.org/10.1145/3054977.3055003

[18] C. Cachin, "Architecture of the hyperledger blockchain fabric," in *Workshop on Distributed Cryptocurrencies and Consensus Ledgers*, 2016.