




Article

# Enable Fair Proof-of-Work (PoW) Consensus for Blockchains in IoT by Miner Twins (MinT)

Qian Qu<sup>1</sup>, Ronghua Xu<sup>1</sup> , Yu Chen<sup>1,\*</sup> , Erik Blasch<sup>2</sup> , Alexander Aved<sup>2</sup>

<sup>1</sup> Dept. of Electrical and Computer Engineering, Binghamton University, Binghamton, NY 13902, USA; {qqu2, rxu22, ychen}@binghamton.edu

<sup>2</sup> The U.S. Air Force Research Laboratory, Rome, NY 13441, USA; {erik.blasch.1, alexander.aved}@us.af.mil

\* Correspondence: ychen@binghamton.edu

**Abstract:** Blockchain technology has been recognized as a promising solution to enhance the security and privacy of Internet of Things (IoT) and Edge Computing scenarios. Taking advantage of the Proof-of-Work (PoW) consensus protocol, which solves a computation intensive hashing puzzle, Blockchain assures the security of the system by establishing a digital ledger. However, the computation intensive PoW favors members possessing more computing power. In the IoT paradigm, fairness in the highly heterogeneous network edge environments must consider devices with various constraints on computation power. Inspired by the advanced features of Digital Twins (DT), an emerging concept that mirrors the lifespan and operational characteristics of physical objects, we propose a novel Miner-Twins (MinT) architecture to enable a fair PoW consensus mechanism for blockchains in IoT environments. MinT adopts an edge-fog-cloud hierarchy. All physical miners of the blockchain are deployed as microservices on distributed edge devices, while fog/cloud servers maintain digital twins that periodically update miners' running status. By timely monitoring miner's footage that is mirrored by twins, a lightweight Singular Spectrum Analysis (SSA) based detection achieves to identify individual misbehaved miners that violate fair mining. Moreover, we also design a novel Proof-of-Behavior (PoB) consensus algorithm to detect byzantine miners that collude to compromise a fair mining network. A preliminary study is conducted on a proof-of-concept prototype implementation, and experimental evaluation shows the feasibility and effectiveness of proposed MinT scheme under a distributed byzantine network environment.

**Keywords:** Digital Twin, Blockchain, Proof-of-Work, Microservices, Singular Spectrum Analysis (SSA), Byzantine Fault Tolerance.

## 1. Introduction

Advancement in Internet of Things (IoT), edge computing, Big Data (BD), and artificial intelligence (AI)/machine learning (ML) technologies makes the concept of Smart Cities realistic. However, widely adopting IoT-based applications and services in smart cities also brings new security and privacy concerns. Thanks to multiple attractive features including decentralization, auditability and traceability, blockchain has been widely recognized as a great potential to revolutionize the fundamentals of information and communication technology (ICT) [1]. Applying blockchain to smart cities is promising to bring efficiency, scalability and security properties to IoT-based applications, such as smart surveillance [2], privacy preservation [3], decentralized data marketplaces [4], time banking of community [5], identity authentication [6] and access control [7,8].

*Digital Twins (DT)* is being developed to optimize manufacturing and aviation processes [9]. By monitoring, simulating and mirroring the status of a physical object (PO), DT can build an intelligent and evolving system model based on the logic object (LO). Leveraging data fusion and AI/ML algorithms, DT can be used to predict the behavior of the PO given some specific situations or environments. Like DT, the Dynamic Data Driven Applications Systems (DDDAS) concept developed in the late 1990s seeks to use modeling to support predictive expectations based on the coordination with models and data [10]. Thus, DDDAS can determine optimized solutions or even failure preventive actions on POs to enable an intelligent and resilient system.



**Citation:** Title. *Preprints* 2021, 1, 0.  
<https://doi.org/>

Received:

Accepted:

Published:

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Research has been conducted to apply blockchain to enable many attractive features in DTs, including transparency, decentralization, data immutability and Peer-to-Peer (P2P) communication [11]. However, directly integrating existing blockchain technologies into the highly heterogeneous IoT environments presents critical challenges in terms of scalability, performance, security, and fairness [12]. Some permissioned blockchains use a Practical Byzantine Fault Tolerance (PBFT) [13] protocol, which demonstrates high throughput, and low latency but only allows a very limited network scalability in terms of the number of validators. Most of permissionless blockchain networks utilize a hashing-intensive proof-of-work (PoW) consensus protocol to achieve security and scalability guarantees. Due to the various computation capability of miners, mining centralization in PoW blockchain not only leads to inequity of rewarding among participants, but it also brings security issues, like majority (51%) attack [14].

Inspired by the essential features of DTs, mirroring and monitoring, this paper proposes a novel edge-fog-cloud Miner-Twins (MinT) architecture to enable a fair PoW consensus mechanism for blockchains in IoT environments. In the MinT architecture, the fog/cloud sever establishes and maintains digital twins for the miners of the blockchain, which are deployed as microservices in edge devices that participate in the blockchain network. Container technology is adopted to encapsulate PoW algorithm as microservices, and each containerized miner is dedicated to mining tasks using pre-configured computation power. Because each miner has the same constrained computation resources, it becomes affordable to optimize resource limited IoT devices.

The MinT architecture enables a fair-mining-as-a-service (FMaaS) framework that timely monitors the computing resources usage at miners and regularly applies anomaly detection to deter misbehaved nodes from unfairly overwhelming honest peers by using extra computing power. Our MinT adopts a lightweight SSA Singular Spectrum Analysis (SSA) based detection to identify individual misbehaved miners that violate fair mining policies. While a Proof-of-Behavior consensus algorithm is designed to detect multiple Byzantine miners that collude to compromise a fair mining network.

The remainder of this paper is organized as follows: Section 2 reviews background on blockchain and PoW consensus, then briefly discusses the state-of-the-art research on DT. Section 3 introduces the rationale and architecture of MinT. The miner twin enabled fair-mining mechanism including SSA and PoB based detection algorithms is explained in Section 4.1. Section 5 presents prototype implementation with numerical results. Section 6 concludes the paper with the future work.

## 2. Related Work

This section introduces blockchain and PoW consensus background knowledge. Following that, we describe digital twin technology and how DT can be used to guarantee the fair mining scheme in blockchain.

### 2.1. Blockchain and Nakamoto Consensus Protocol

As a form of distributed ledger technology (DLT), *Blockchain* was initially implemented as an enabling technology of Bitcoin [15], which aims to provide a cryptocurrency to record and verify commercial transactions among trustless entities in a decentralized manner. With the decentralized P2P network architecture and cryptographic mechanisms, participants in a blockchain system maintain the immutability and auditability of data and transactions recorded on the distributed ledger, instead of relying on a centralized third party trust authority.

As one of the most fundamental problems in a distributed/decentralized computing environment, *consensus* in a blockchain network can be defined as a fault-tolerant state-machine replication problem, which aims to maintain the globally distributed ledger state across the P2P network. Bitcoin adopts the Nakamoto consensus based on a Proof-of-Work (PoW) scheme to achieve pseudonymity, scalability and probabilistic finality in an asynchronous and open-access network environment. The goal of Nakamoto consensus

is to ensure all participants agree on a common network transaction log as a serialized blockchain [12].

PoW is essentially an incentive-based consensus algorithm, which requires all participants to compete for rewards through a cryptographic block discovery racing game. To be a winner in PoW block generation, every node has to solve a computing-intensive hash puzzle problem. In brief, a valid PoW solution requires exhaustively querying a cryptographic hash function for a partial preimage generated from a candidate block [16]. Finally, the hash code of a candidate block must satisfy a pre-defined difficulty condition parameter  $h$ , like having a fixed length of bits as zeros. The PoW puzzle problem can be formally defined as:

$$\text{hash\_block} = \mathcal{H}(\text{block\_data}|\text{nonce}) \leq D(h), \quad (1)$$

where for some fixed length of bits  $L$  and difficulty condition  $D(h) = 2^{L-h}$ .  $\mathcal{H}(\cdot)$  is a pre-defined collision-resistant cryptographic hash function that outputs a hash string  $L \in \{0, 1\}^\lambda$ , and  $\lambda$  is the length of a hash string.

The PoW process defined by Eq. (1) is essentially a verifiable process of a weighted random coin-tossing [12]. Thus, the probability of generating a valid block is in proportion to miners' computation resources. Higher computation power leads to higher hash string rate in PoW, which means more rewards and benefits. Such a mining centralization may discourage participants who have limited computation resources, like IoT devices; but it also lead to majority (51%) attack if an adversary has controlled more than 50% computation resource of of the whole network.

To reduce energy consumption in PoW consensus, Peercoin [17] proposed Proof-of-Stake (PoS), which requires a miner to use its coin stake to solve the puzzle solution. Unlike PoW protocols that relies on a brute-force hash calculation, PoS miners use a process of "virtual mining" manner that only consumes limited computational resources. However, PoS still has a mining centralization issue, because an attacker can amplify its power by simply accumulating the credit stake. As the first practical BFT consensus, Practical BFT (PBFT) [13] guarantees both liveness and safety in synchronous network environments given the assumption that at most of  $\lfloor \frac{n-1}{3} \rfloor$  out of total of  $n$  participants in consensus protocol are Byzantine faults. As PBFT requires that all nodes communicate synchronously to achieve consensus purposes, it has poor scalability due to high latency and communication overhead as more nodes join the consensus network.

## 2.2. Digital Twins

The concept of DT was proposed in 2002 and archived in a NASA white paper in 2014 [18]. Essentially, a DT is a digital representation of the components and dynamics of a physical system [19]. Based on the functionalities, DTs can be roughly categorized into three kinds: monitoring DTs, simulational DTs, and operational DTs [20]. As suggested by the names, monitoring twins allow system operators monitor the status of a physical system; simulation twins can predict the future status of the physical system in different scenarios by using various simulation tools and ML algorithms; operational twins is a *complex sensing and control system* that enabled human operators to interact with a cyber-physical system and perform different actions in addition to monitoring, analysis and prediction [21], which is similar to human-machine teaming [22].

Earlier studies of DT mainly focused on the area of manufacturing covering different key factors for smart manufacturing including simulation, optimization and the use of AI. For instance, an event-driven simulation for manufacturing and assembly tasks based on Digital Twin and human-robot collaboration is presented [23]. A DT based framework is proposed to achieve high precision and multidisciplinary coupling during the assembly process, which mainly focused on High precision products (HPPs) workshops [24]. HPP also establishes a predict and optimization model as well as a case study to verify the effectiveness and feasibility. A case study presents an ice cream machine as an application example of DT in food industry [25], which focused on the visualization and interaction

based on virtual reality (VR) and augmented reality (AR) technologies. Secure data transmission is also highlighted in the framework by employing a secure gate between machine and cloud.

Recently efforts are reported on variant aspects of smart cities including Smart Driving, Smart Grid and Smart Healthcare. For instance, the optimization issue in the electric propulsion drive systems (EPDS) of self-driving electric vehicles are discussed [26]. In the proposed DT-based framework, the connection between logical twin in the control software with the propulsion motor drive system enables EPDS performance estimation. However, there are no experimental results are presented after giving the concepts of the platform. A behaviors based algorithm is proposed to help the drivers to avoid potential risk [27]. Combining the ML techniques and DT relies on the connectivity of the system and faces challenges in optimization and accuracy [28]. A case study has been reported that tackles the management of wind farm using DT and cloud technologies combined with big data analysis to build remote control station [29].

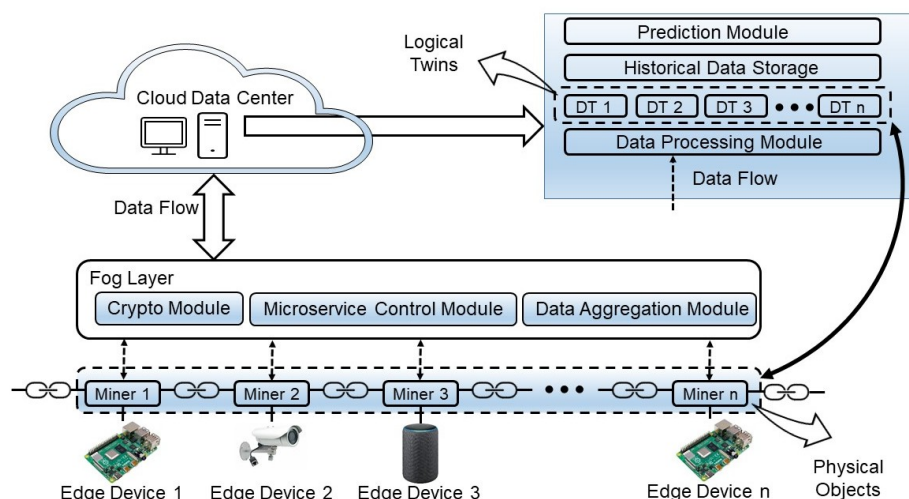
Recently, some healthcare applications redefined DT by including living objects [30]. A DT-based healthcare framework is proposed for monitoring and predicting the health condition of an individual using wearable devices [31]. A DT-based remote surgery prototype is introduced consisting of VR, 4G and AI to create a digital twin of a patient and to realize real-time surgery over mobile network [32]. Due to the fast development of telecommunication technologies, 5G and beyond networks will be very complicated as they are expected to support more emerging applications with more diverse requirements [33]. The community is considering DT as an efficient, cost-effective approach to accelerate the design, test, and implementation of 5G/6G networks [34].

Due to the foreseeable importance and popularity of DT in IoT, 5G/6G, and edge computing area, blockchain is adopted to enhance the security, trust, and reliability of DTs [35], [11]. Work reported in this paper, however, is the first in this area that leverages the DT to tackle the unfair mining problem in the PoW consensus protocol. Using digital twins, MinT monitors the computing resource utility of the miners and quickly detects abusers using Singular Spectrum Analysis (SSA) [36], one of the fastest change point detection algorithms [37]. Our MinT also uses a Proof-of-Behavior (PoB) consensus algorithm to guarantee byzantine tolerant anomaly detection.

### 3. MinT: Rationale and Architecture

Aiming at a secure-by-design fair PoW mining network in heterogeneous IoT environments, our MinT scheme leverages DT technology to continuously monitor the usage of containerized miners and discourages misbehaving nodes from unfairly overwhelming the peers by using extra computing power. Figure 1 illustrates the high-level system architecture of MinT, which adopts a hierarchical cloud-fog-edge computing paradigm. Such a hierarchical framework not only provides system scalability for large-scale fair mining tasks based on geographically distributed IoT devices; but it also supports flexible management and coordinated central and decentralized local decisions given heterogeneous networks and application domains. Moreover, MinT relies on a permissioned network which provides basic security guarantees, like the public key infrastructure (PKI) and digital signature, data integrity [2], identity authentication [6] and access control [38], etc.. The rationale behind the MinT is described as follows:

**1. Containerized PoW Miner:** The edge layer in MinT consists of various types of IoT devices, like smart cameras in a surveillance system or smart meters connected to a power grid. To follow an ideal “one cup-one vote” Nakamoto consensus protocol, devices are only allowed to launch PoW containers as miners to participate the blockchain network, and all containers are assigned the same computation resource for PoW mining algorithm. Each miner has the same probability of generating blocks and being rewarded accordingly due to the uniform computation distribution of the network. Thus, these containerized PoW miners construct a fair mining blockchain network disregarding devices’ capability.



**Figure 1.** Illustration of MinT System Architecture.

**2. Microservice-oriented Service:** MinT utilizes an intermediate fog layer to provide middle-ware services for devices at edge and cloud level. To address heterogeneity of IoT systems, a lightweight Microservice-oriented architecture (MoA) is adopted as a fundamental service infrastructure to support functionality, such as data aggregation and microservice management, and security mechanisms, like encryption/decryption, identity verification and access control, etc.. Each microservice unit exposes a set of RESTful web-service APIs for interaction. The fine-granularity and loose-coupling features of the MoA framework allows for fast development and easy deployment among heterogeneous platforms using non-standard development.

**3. DT enabled Fair Mining Intelligence:** As dishonest containerized miners could use extra computing power than they have been permitted, MinT relies on DT technology and intelligent services on a fog/cloud server to maintain a fair mining network at the edge layer. By aggregating data flows from distributed miners, mirroring miners (logic objects) that are associated with their physical counterparts are created. These miner twins monitor the usage of containerized miners running on devices. By analyzing the real-time status of miner twins and historical statistics, abusers can be detected and preventive actions can be triggered to deter identified misbehaving miners; such that the MinT ensures a fair mining blockchain network.

#### 4. Miner Twin Enabled Fair-Mining Mechanism

This section provides a comprehensive overview of MinT based fair mining mechanism such that readers can understand key components and workflow. Then, we describe miner twin process including key parameters selection. Following that, we offer details on lightweight SSA based anomaly detection and byzantine tolerant PoB consensus algorithm.

##### 4.1. MinT Workflow for Fair Mining

Figure 2 illustrates the workflow of the fair-mining mechanism in MinT system. The upstream data flow starts from the containerized miners and aggregates the fog servers installed with different modules. The fog server firstly normalize the data from all physical miners which reports to it under its jurisdiction. The fog server can either construct logical miners that mirror these new physical miners or update status of existing logical miners. The fog server further encrypts its local logical twinning miners and forward them to the cloud.

On receiving the encrypted data from multiple fog servers, the cloud server aggregates the information into a logical miners pool to represent a system level twinning PoW network. Using the live feed from the logical twin and the historical data, MinT uses an intelligent model for fair mining strategy. Given a fair mining algorithm, the upstream

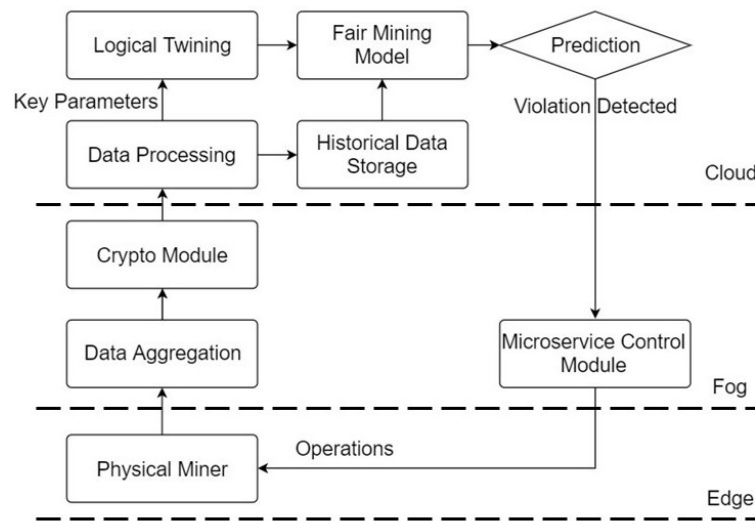


Figure 2. Miner Twins based Fair-Mining Flowchart.

data flow starts from the predication. The predicted status will be compared to the actual footage, using anomaly detection algorithm MinT identifies dishonest miners who violate the fair PoW consensus and sends orders to the Microservice Control Module on a fog layer accordingly, which will take further actions on the “outlaws”.

#### 4.2. Miner Twin Process

To mirror the physical miner, several parameters are extracted for the logical miner, including central CPU usage ( $C$ ), global GPU usage ( $G$ ), memory usage ( $M$ ), and I/O bandwidth ( $B$ ). Since PoW depends on computation intensive algorithms, the CPU usage and GPU usage are chosen as the *Key Parameters* according to the selection of calculation module, while memory, I/O bandwidth and other metrics are considered as *Contributing Parameters*. To avoid falling behind other miners, the physical miner will normally use all the allocated CPU/GPU resources.

Because the system resource allocated to each miner is restricted but identical, the data can be normalized in form of percentage, for example  $c = \frac{C}{C_{set}} \times 100\%$  where  $C_{set}$  is the preset CPU limit and  $c$  is the normalized value. Given an assumption that a containerized miner can only use its CPU to do the PoW algorithm, then for a miner  $k$ , the parameter vector of its Physical Object (physical miner) with timestamp  $i$  would be  $PO_{ki} = (c_{ki}, g_{ki}, m_{ki}, b_{ki})$ , and the Key Parameter is  $c_{ki}$ . And the vector for the Logical Object (logic miner) can be represented as  $LO_{ki} = (c_{ki}, g_{ki}, m_{ki}, b_{ki})$  and the Key Parameter is  $c_{ki}$ .

#### 4.3. Fast Anomaly Detection for Fair Mining

Fast and accurate identifying of the misbehaved miners is the essential step to ensure fair mining, where MinT adopts the Singular Spectrum Analysis (SSA) algorithm to achieve this goal. SSA is recognized as one of the quickest sequential change-point detection approaches for processing time series problems [39]. By decomposing and reconstructing the interested time series, SSA extracts certain components of the origin series like periodic pattern, noises, trends, etc. SSA is widely used in solving problems like smoothing, extraction of seasonality components, as well as study the structure in some minor time series and change-point detection [36].

Unlike traditional methods, SSA is non-parametric and does not require prior knowledge of the parametric model of the considered time series data. Although SSA uses some statistical concepts, it does not need any statistical assumptions about the target series. Moreover, SSA algorithm can be used for processing time series with relatively small size,

which make this method more suitable for edge-fog scenarios [40]. The SSA algorithm can be described as follows:

**1. Embedding:** The target of SSA is a one-dimensional time series  $\mathbb{X} = [x_1, \dots, x_N]$ , where  $N$  is the series length. By choosing proper window length  $L$ , one can transfer the times series into multi-dimensional series of vectors  $\vec{X}_i$ . Combine these vectors results in the trajectory matrix  $X = [\vec{X}_1, \vec{X}_2, \dots, \vec{X}_K]$ , where  $K = N - L + 1$ . The multi-dimensional vectors  $\vec{X}_i = (x_i, \dots, x_{L+i-1})'$ ,  $i = 1, \dots, K$ , are also called lagged vectors.

**2. Singular Value Decomposition:** After singular value decomposing the trajectory matrix  $X$ , the eigenvalues are denoted by  $\lambda_1, \dots, \lambda_L$  in decreasing order of magnitude and the corresponding eigenvectors  $U_1, \dots, U_L$  where the matrix  $U = [U_1, U_2, \dots, U_L]$ ,  $\|U_i\| = 1$  is orthogonal. Then, the eigentriples are  $(\sqrt{\lambda_i}, U_i, V_i)$ , by denoting  $V_i = X'U_i/\sqrt{\lambda_i}$ . Supposing the rank of  $X$  is  $d$ , then the trajectory matrix as  $X = X_1 + \dots + X_d$ .

**3. Grouping and Reconstructing:** The next step is to group the matrices  $X_i$  into certain groups and calculate the sum within these groups. So we denote a subset indices  $I = i_1, i_2, \dots, i_l$  where  $l < L$ . So the corresponding matrix is  $X_I = X_{i_1} + \dots + X_{i_l}$ .

**4. Diagonal Averaging:** Using diagonal averaging, we can transfer  $X_I$  into time series  $\mathbb{X}_I$ .

$$\mathbb{X}_I(i) = \begin{cases} \frac{1}{l} \sum_{j=1}^i x_{j,i-j+1} & \text{for } 1 \leq i < L \\ \frac{1}{l} \sum_{j=1}^L x_{j,i-j+1} & \text{for } L \leq i \leq K \\ \frac{1}{N-i+1} \sum_{j=i-K+1}^{N-K+1} x_{j,i-j+1} & \text{for } K \leq i \leq N \end{cases}$$

By selecting certain subset indices  $I = i_1, i_2, \dots, i_l$ , one can reconstruct the time series. By observing the distance between the  $l$ -dimensional matrix and the test time series matrix, we can detect the anomaly by identifying a significant increase of the distance. The SSA based Change-Point detection utilized in the paper can be described in following stages:

**Stage 1: Construct Base Matrix** First construct the base matrix (or target matrix) according to the four steps of the SSA algorithm. Given the target time series  $\mathbb{X} = [x_{n+1}, \dots, x_{n+N}]$ , embed it into the trajectory matrix  $X = [\vec{X}_1, \vec{X}_2, \dots, \vec{X}_K]$ , where  $K = N - L + 1$ . And the columns of the trajectory matrix are the vectors:

$$\vec{X}_i = (x_{n+i}, \dots, x_{n+L+i-1})', i = 1, \dots, K$$

Then conduct the SVD and get  $L$  eigenvectors which can be grouped into certain subset  $I = i_1, i_2, \dots, i_l, l < L$ .

**Stage 2: Construct Test Matrix** Similarly, construct the test matrix of size  $L \times Q$ :

$$X_{test} = [X_{p+1}^{\vec{}}, X_{p+2}^{\vec{}}, \dots, X_p^{\vec{}} + Q],$$

where  $q = p + Q$ . And the columns of the matrix are the vectors:

$$\vec{X}_j = (x_{n+j}, \dots, x_{n+L+j-1})', j = p + 1, \dots, p + Q$$

**Stage 3: Compute the Detection Statistics** In this stage, we first compute  $D_{n,l,p,q}$ , the sum of the squared Euclidean distances between the  $l$ -dimensional subspace from the base matrix and the vectors  $\vec{X}_j (j = p + 1, \dots, p + Q)$  from the test matrix.

$$D_{n,l,p,q} = \sum_{j=p+1}^q ((\vec{X}_j)^T \vec{X}_j - (\vec{X}_j)^T U U^T \vec{X}_j)$$

Then we give the normalized sum of squared distances  $S_n = \frac{1}{\mu_{n,l}} \tilde{D}_{n,l,p,q}$ .  $\mu_{n,l} = \tilde{D}_{m,l,0,K}$  is the estimator and we make the hypothesis that no change of time series structure occurs at the time intervals where  $m$  is the largest value of  $m \leq n$ .

We also compute the Cumulative Sum as the final score for the anomaly detection.

$$W_1 = S_1, W_{n+1} = \max\{0, W_n + S_{n+1} - S_n - \kappa/\sqrt{LQ}\}, n \geq 1$$

where  $\kappa$  is a constant and in this paper we set  $\kappa = 1/(3\sqrt{LQ})$  [41].

**Stage 4: Set threshold and make decisions** To detect the change of the time series, we could check the value of  $D_{n,l,p,q}$ ,  $S_n$  and  $W_n$ . Basically the large value of the three detection statistics indicates the change or the anomaly and the algorithm announces the structural change if we observe  $W_n > h$  for some  $n$  where  $h$  is the threshold given by

$$h = \frac{2t_\alpha}{LQ} \sqrt{\frac{1}{3}Q(3LQ - Q^2 + 1)}$$

and  $t_\alpha$  is the  $1 - \alpha$ -quantile of the standard normal distribution [42].

#### 4.4. Proof-of-Behavior Consensus Algorithm for Fair Mining Enforcement

Above mentioned SSA based detection can identify a single misbehaved miner based on its own footnote, however, it cannot handle byzantine scenarios that multiple compromised miners by an adversary collude to violate fair mining policies. By observing a miner's running operations, the calculated cumulative sum (CUSUM)-type  $W$  can indicate a miner's behavior. Inspired by deepfake detection in video surveillance systems [43,44], our MinT relies on a novel *Proof-of-Behavior* consensus algorithm that leverages CUSUM-type  $W$  calculated in SSA algorithm to detect multiple dishonest miners in distributed byzantine tolerant scenarios.

We consider a mining network  $\mathcal{N}$  including  $n_i$  miners, where  $i \in \{1, k\}$  and  $k = |\mathcal{N}|$ . All dishonest miners are denoted by  $m_i \in \mathcal{M}$  and their fraction is  $f = |\mathcal{M}|/|\mathcal{N}|$ . The CUSUM-type  $W_i$  of miner  $n_i$  denotes a behavior vector  $W_i = \{w_1, w_2, \dots, w_d\}$ , where  $d$  is the SSA detection time window. Finally, each twin can maintain a global view of collected behavior vectors, which is a matrix  $G = \{W_1, W_2, \dots, W_k\}$ . The PoB firstly generates a behavior score  $s(i)$  for each miner  $n_i$ , which is a sum of relative Euclidean distances between other miners' behavior vector. Then, a  $W_i \in G$  with minimal behavior score will be selected as a benchmark  $W^*$ .

The PoB consensus algorithm aims to chooses a vector  $W$  which is deviates least from the distribution of  $G$ . However, an adversary can compromise multiple miners that generate large vectors to force "honest" miners to choose a byzantine behavior vector as the ground truth one. Thus, our PoB algorithm adopts a *Krum* aggregation rule [45] to guarantee byzantine tolerance. We assume that honest miners within network  $\mathcal{N}$  store  $G$  including  $n \geq 2f + 3$  vectors in which at most  $f$  vectors are generated by byzantine nodes in  $\mathcal{M}$ . For  $W_j$  belongs to the  $n - f - 2$  closest vectors to  $W_i$ , where  $i \neq j$ , we denote  $i \rightarrow j$ . So we could define the consensus score:

$$s(i) = \sum_{i \rightarrow j} \|W_i - W_j\|^2$$

Then each nodes can compute behavior scores  $s(1), \dots, s(k)$  that are associated with miners  $n_1, \dots, n_k$  separately. By calculating the minimum behavior score

$$s^* = \min_{i \in \{1, \dots, k\}} (s(i)),$$

all honest miners choose a behavior vector  $W_i$  that satisfies  $s(i) = s^*$  as the ground truth  $W^*$ . Given assumption that an adversary controls no more than  $f$  miners, all honest miners can reach an agreement on the unique  $W^*$ .

## 5. Experimental Study

In this section, a proof-of-concept prototype implementation and experimental configuration are described. Following that we evaluate effectiveness of proposed MinT solution based on numerical results. Finally, we discuss performance and security properties provided by MinT.



### 5.1. Experimental Setup

A proof-of-concept test platform is created, in which 16 Raspberry Pis (RPi) are adopted as the edge devices. Each RPi is empowered with quad-core Cortex-A72 CPU @1.5GHz and an installed RAM with 4GB memory running Raspbian OS based on Debian. The single-board computer (SBC) is capable of carrying containerized PoW module to participate the blockchain network. A desktop functions as a fog server, which has Intel Core i7-7700K CPU and a RAM of 16 GB memory. All the RPis are connected to a fog server via local area network (LAN).

As the GPU is not available on the RPi, we select CPU-based PoW algorithm for container construction. For fast deployment, Docker [46] is adopted as the microservice container that is affordable to RPis and transmit the data from the physical miner to a fog server through RESTfull APIs. Each of the miner containers is configured with and restricted to one CPU core, 500MB memory and 10 percent of system I/O bandwidth. The collected data is stored in forms of vector as described in Section 4.2.

As the PoW algorithm is executed on CPU, samples of the key parameter  $C$  are collected and the historical data vector  $c_{hi}$  is used to obtain the statistic profile, where  $h = 1, \dots, 16$  and  $i = 0, 1, \dots$ . For SSA based change-point detection, we define  $N = 24$  according to the size of the data sets,  $L = 12$  to the half size of  $N$ ,  $p = 12$ , and  $q = 24$ . We deliberately set  $p \geq K$  so that the base and test matrices would not coincide. After visual inspection of the components of the decomposition of the whole time series, we choose certain  $l$  to represent ignoring the noise components.

### 5.2. Experimental Results

All 16 miners are default running at 100% of the assigned system resources under the jurisdiction of the fog server. Four different test scenarios are considered in our experimental study. To verify SSA based detection on single misbehaved miner, we firstly conduct test cases that only one dishonest miner uses double assigned computation power on mining given different parameters combination. Then we consider a more stealthy single miner violation, which incremental increases the computing power from 20% up to 50%. To validate effectiveness of PoB based detection, we simulate a byzantine network, in which two miners act as byzantine nodes while other 14 miners are honest members. Finally, we evaluate false positive rates at the network level with different threshold settings.

(a)	(b)	(c)	(d)
Av-	Si-	Si-	Si-
er-	gle	gle	gle
age,	Miner,	Miner,	Miner,
$l =$	$l =$	$l =$	$l =$
4,	4, 8,	4,	4,
$q =$	$q =$	$q =$	$q =$
$p =$	$p =$	$p =$	$p =$
12.	12 12.	24.	24.

**Figure 3.** SSA detection on single miner violation with different parameters combination.

#### 5.2.1. SSA detection on static single miner violation

In this scenario, one dishonest miner uses twice as much CPU power as the assigned amount at  $t = 200$ s. Figure 3(a) presents the network level observation at the fog server. The blue line is the average CPU usage for all 16 miners in this blockchain network, and the red line is the  $w_n$  value calculated using SSA algorithm as the score. And the green line is the threshold  $h = 0.607$  which is computed with  $t_\alpha = 1.2815$ . As shown by Fig. 3(a), the fluctuation in the average CPU utility incurs a low peak in the distance score. However, applying the SSA algorithm on each miner twin individually avoids the false negative. Figure 3(b) shows that a significant peak is observed at  $t = 200$ s.

We also studied the impacts of different selections of the SSA parameters varying  $l$  and  $q - p$  combination. Figure 3(c) shows the consequence of increasing the value of  $l$  from

4 to 8, but with the same matrix size. The larger  $l$  leads to a more noise part with the signal; therefore, it would be more difficult to find a change in the signal time series. And if the  $l$  is too small which would cause underfitting, we may miss some part of the signal. Due to the limited space, the figure is not included here.

Meanwhile, the matrix size  $q - p$  also has significant impact on the detection distance score. Figure 3(d) shows that by increasing the value of  $q - p$  to 24 while  $l = 4$ , the distance (red) line is smoother than in (b).

### 5.2.2. SSA detection on adaptive single miner violation

The second scenario considers more stealthy behavior of a violator, which increases the computing power slowly, from 20% to 50% taking multiple steps at time point  $t = 125s$ ,  $t = 175s$ ,  $t = 225s$ ,  $t = 275s$ . Figure 4(a) shows detection results that a miner increases 20% CPU usage at each time point. Figures 4(b)-(d) show similar results of cases when the CPU usage increases by 30%, 40%, and 50% respectively. Obviously, the SSA based anomaly detection is able to detect the changes in the structure of the time series data and identify the corresponding violation on mining power. However, the critical issue is how to select a threshold to ensure a high detection accuracy and minimize the false positive/negative rate.

(a) (b)(c) (d)  
Ex- Ex- Ex-  
tra tra tra  
CPU CPU CPU  
us- us- us-  
age age age  
at at at  
20%. 30%. 40%. 50%.

**Figure 4.** SSA detection on single miner violation with additive CPU usage.

### 5.2.3. PoB based Fair Mining Detection Effectiveness

We take an observation of 20 minutes on the 16 miners running at 100% of the assigned system resource. And two of the miners act as the byzantine (dishonest) workers which would gain extra 10% at the 9th and 10th minute. As shown in Figure 5(a), the  $W$  vectors from dishonest workers varies from honest ones when the byzantine workers gain more computing power. And during the two minutes where violation occurs, the resulting consensus scores associated with the byzantine nodes are much larger, as shown in Figure 5(b).

(a) (b)  
Be- Be-  
hav- hav-  
ior ior  
vec- scores  
tors dis-  
in tri-  
9th bu-  
minution.

**Figure 5.** Behavior score distribution with sequential time spots.

### 5.2.4. Fair mining violation detection performance analysis

The fourth scenario is designed that mainly tests the false positive rate from the network level observation at the fog server with different threshold settings. Figure 6 shows the false alarm rates when two of the sixteen miners gain extra system resource from 10% to 80%. The false alarm rate is calculated by comparing the averaged the  $W$  value with the threshold  $h$ . When we decrease  $h$  from 0.6 to 0.03, the false alarm rate increases rapidly

at the beginning and then slowly approaches to one. With the increasing percentage of the computing power the dishonest miner gains, the false alarm rate grows.

**Figure 6.** False alarm rate with different threshold  $h$ .

### 5.3. Discussions

The experimental results presented in this section is merely a preliminary study on top of a proof-of-concept platform. It validates the feasibility to identify dishonest miners based on digital twins integrating the SSA and novel POB consensus algorithms. The experimental system covers the edge and fog layers of the proposed MinT architecture that envisions large scale IoT networks including edge devices under various type of fog servers connected to a cloud data center. Furthermore, our PoB consensus algorithm adopts Krum rule in behavior score calculation, which only chooses  $n - f - 2$  closet behavior vectors and precludes those  $f - 1$  malicious vectors that are far away from the center of distribution. Given assumption that an adversary cannot control more than  $f$  nodes of a mining network  $\mathcal{N}$  that satisfies  $n \geq 2f + 3$ , all honest participants can still output the same benchmark behavior vector  $W^*$ .

There are a lot of open questions that have not been addressed, meanwhile, some of them are on the list of our on-going efforts.

- More comprehensive study on SSA is to be conducted to answer questions such as, how to select an optimal/sub-optimal threshold? How to minimize the detection delay as scaling up miners?
- The mechanisms that ensure security and authenticity of the data transmitted from PO to LO are among the tasks of top priority. Although PoB consensus is promising to guarantee byzantine fault tolerance in mining violation detection, the threat model based on attack scenarios in SSA detection needs more investigation, like communication security between miner and twin and container's robustness given failed or compromised conditions.
- A comprehensive understanding of computation cost and network latency in the twinning process is mandatory. We have to tackle performance and scalability issues to bring MinT into practice.

## 6. Conclusions and Future work

In this paper, we propose MinT, an edge-fog-cloud architecture to enable a fair PoW consensus mechanism leveraging miner twins. Experimentally the paper validated the feasibility of the concept of using DT to monitor the miners' behaviors and deter selfish nodes who violate the fair-mining rule. The reported preliminary results are based on quick change point detection and a PoB consensus algorithm to catch violators, however, more intelligent solutions are need to solve large scale, hierarchical IoT networks in the real world using MinT model. Our future work includes the following.

Besides a comprehensive investigation on anomaly detection using SSA method, AI/ML based algorithms will be investigated for efficient, accurate detection of dishonest miners in the blockchain network. Secondly, on top of our edge-fog-cloud based smart surveillance system [47], we will construct a large scale MinT system with a complete edge-fog-cloud architecture, implementing miner twins at both fog and cloud layers and collecting more data for ML model training and testing purposes. In addition, since the miners and their digital twins are implemented as microservices in MinT, security and privacy of the miner twins are among the top concerns; specifically data security and microservice-to-microservice authentication and authorization will be investigated as the core of a solid MinT architecture.

- 1 **Author Contributions:** Conceptualization, Q.Q., R.X. and Y.C.; methodology, Q.Q. and R.X.; software,
- 2 Q.Q. and R.X.; validation, Q.Q., R.X. and Y.C.; formal analysis, Q.Q. and R.X.; investigation, Y.C.;
- 3 resources, Y.C., E.B. and A.A.; data creation, Q.Q.; writing—original draft preparation, Q.Q. and

4 R.X.; writing—review and editing, Y.C., E.B. and A.A.; visualization, Q.Q. and R.X.; supervision, Y.C.  
5 and E.B.; project administration, Y.C. and A.A.; funding acquisition, Y.C. All authors have read and  
6 agreed to the published version of the manuscript.

7 **Funding:** This work is partially supported by the U.S. National Science Foundation (NSF) via grants  
8 CNS-2141468.

9 **Acknowledgments:** The views and conclusions contained herein are those of the authors and should  
10 not be interpreted as necessarily representing the official policies or endorsements, either expressed  
11 or implied, of the U. S. Air Force.

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

### 13 Abbreviations

14 The following abbreviations are used in this manuscript:

15	AI	Artificial Intelligence
	AR	Augmented Reality
	CUSUM	Cumulative Sum
	DDDAS	Dynamic Data Driven Applications Systems
	DLT	Distributed Ledger Technology
	DT	Digital Twins
	FMaaS	Fair Mining as a Service
	EPDS	Electric Propulsion Drive Systems
	ICT	Information and Communication Technology
	IoT	Internet of Things
	LAN	Local Area Network
	LO	Logical Object
16	MinT	Miner Twins
	ML	Machine Learning
	MoA	Microservice-oriented architecture
	P2P	Peer-to-Peer
	PBFT	Practical Byzantine Fault Tolerance
	PO	Physical Object
	PoB	Proof-of-Behavior
	PoS	Proof-of-Stake
	PoW	Proof-of-Work
	SBC	Single Board Computer
	SSA	Singular Spectrum Analysis
	VR	Virtual Reality

### References

1. Novo, O. Blockchain meets IoT: An architecture for scalable access management in IoT. *IEEE Internet of Things Journal* **2018**, *5*, 1184–1195.
2. Nikouei, S.Y.; Xu, R.; Nagothu, D.; Chen, Y.; Aved, A.; Blasch, E. Real-time index authentication for event-oriented surveillance video query using blockchain. 2018 IEEE International Smart Cities Conference (ISC2). IEEE, 2018, pp. 1–8.
3. Fitwi, A.; Chen, Y. Secure and Privacy-Preserving Stored Surveillance Video Sharing atop Permissioned Blockchain. *arXiv preprint arXiv:2104.05617* **2021**.
4. Xu, R.; Chen, Y. Fed-DDM: A Federated Ledgers based Framework for Hierarchical Decentralized Data Marketplaces. *arXiv preprint arXiv:2104.05583* **2021**.
5. Xu, R.; Zhai, Z.; Chen, Y.; Lum, J.K. BIT: A blockchain integrated time banking system for community exchange economy. 2020 IEEE International Smart Cities Conference (ISC2). IEEE, 2020, pp. 1–8.
6. Xu, R.; Chen, Y.; Blasch, E.; Chen, G. Exploration of blockchain-enabled decentralized capability-based access control strategy for space situation awareness. *Optical Engineering* **2019**, *58*, 041609.
7. Xu, R.; Chen, Y.; Blasch, E.; Chen, G. Blendcac: A blockchain-enabled decentralized capability-based access control for iots. 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). IEEE, 2018, pp. 1027–1034.
8. Xu, R.; Chen, Y.; Blasch, E.; Chen, G. Blendcac: A smart contract enabled decentralized capability-based access control mechanism for the iot. *Computers* **2018**, *7*, 39.

9. Barricelli, B.R.; Casiraghi, E.; Fogli, D. A survey on digital twin: Definitions, characteristics, applications, and design implications. *IEEE Access* **2019**, *7*, 167653–167671.
10. Blasch, E.; Ravela, S.; Aved, A. *Handbook of dynamic data driven applications systems*; Springer, 2018.
11. Yaqoob, I.; Salah, K.; Uddin, M.; Jayaraman, R.; Omar, M.; Imran, M. Blockchain for digital twins: Recent advances and future research challenges. *IEEE Network* **2020**, *34*, 290–298.
12. Xu, R.; Chen, Y.; Blasch, E. Microchain: a Light Hierarchical Consensus Protocol for IoT System. *Blockchain Applications in IoT: Principles and Practices* **2021**.
13. Castro, M.; Liskov, B.; others. Practical Byzantine fault tolerance. OSDI, 1999, Vol. 99, pp. 173–186.
14. Alsabah, H.; Capponi, A. Pitfalls of Bitcoin's Proof-of-Work: R&D arms race and mining centralization. *Available at SSRN 3273982* **2020**.
15. Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system. Technical report, Manubot, 2019.
16. Wang, W.; Hoang, D.T.; Hu, P.; Xiong, Z.; Niyato, D.; Wang, P.; Wen, Y.; Kim, D.I. A survey on consensus mechanisms and mining strategy management in blockchain networks. *IEEE Access* **2019**, *7*, 22328–22370.
17. King, S.; Nadal, S. Ppcoin: Peer-to-peer crypto-currency with proof-of-stake. *self-published paper, August* **2012**, 19.
18. Grieves, M. Digital twin: manufacturing excellence through virtual factory replication. *White paper* **2014**, *1*, 1–7.
19. Erkoyuncu, J.A.; Butala, P.; Roy, R.; others. Digital twins: Understanding the added value of integrated models for through-life engineering services. *Procedia Manufacturing* **2018**, *16*, 139–146.
20. Van Schalkwyk, P. The Ultimate Guide to Digital Twins. <https://xmpro.com/digital-twins-the-ultimate-guide/> **2019**.
21. Khan, L.U.; Saad, W.; Niyato, D.; Han, Z.; Hong, C.S. Digital-Twin-Enabled 6G: Vision, Architectural Trends, and Future Directions. *arXiv preprint arXiv:2102.12169* **2021**.
22. Blasch, E.; Lambert, D.A. *High-level information fusion management and systems design*; Artech House, 2012.
23. Bilberg, A.; Malik, A.A. Digital twin driven human-robot collaborative assembly. *CIRP Annals* **2019**, *68*, 499–502.
24. Sun, X.; Bao, J.; Li, J.; Zhang, Y.; Liu, S.; Zhou, B. A digital twin-driven approach for the assembly-commissioning of high precision products. *Robotics and Computer-Integrated Manufacturing* **2020**, *61*, 101839.
25. Karadeniz, A.M.; Arif, İ.; Kanak, A.; Ergün, S. Digital twin of egastronomic things: A case study for ice cream machines. 2019 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2019, pp. 1–4.
26. Rassölkin, A.; Vaimann, T.; Kallaste, A.; Kuts, V. Digital twin for propulsion drive of autonomous electric vehicle. 2019 IEEE 60th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON). IEEE, 2019, pp. 1–4.
27. Chen, X.; Kang, E.; Shiraishi, S.; Preciado, V.M.; Jiang, Z. Digital behavioral twins for safe connected cars. Proceedings of the 21th ACM/IEEE International Conference on Model Driven Engineering Languages and Systems, 2018, pp. 144–153.
28. Kapteyn, M.G.; Knezevic, D.J.; Willcox, K. Toward predictive digital twins via component-based reduced-order models and interpretable machine learning. AIAA Scitech 2020 Forum, 2020, p. 0418.
29. Pargmann, H.; Euhäusen, D.; Faber, R. Intelligent big data processing for wind farm monitoring and analysis based on cloud-technologies and digital twins: A quantitative approach. 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA). IEEE, 2018, pp. 233–237.
30. El Saddik, A. Digital twins: The convergence of multimedia technologies. *IEEE multimedia* **2018**, *25*, 87–92.
31. Liu, Y.; Zhang, L.; Yang, Y.; Zhou, L.; Ren, L.; Wang, F.; Liu, R.; Pang, Z.; Deen, M.J. A novel cloud-based framework for the elderly healthcare services using digital twin. *IEEE Access* **2019**, *7*, 49088–49101.
32. Laaki, H.; Miche, Y.; Tammi, K. Prototyping a digital twin for real time remote control over mobile networks: Application of remote surgery. *IEEE Access* **2019**, *7*, 20325–20336.
33. Saad, W.; Bennis, M.; Chen, M. A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. *IEEE network* **2019**, *34*, 134–142.
34. Nguyen, H.X.; Trestian, R.; To, D.; Tatipamula, M. Digital twin for 5G and beyond. *IEEE Communications Magazine* **2021**, *59*, 10–15.
35. Suhail, S.; Hussain, R.; Jurdak, R.; Oracevic, A.; Salah, K.; Hong, C.S. Blockchain-based Digital Twins: Research Trends, Issues, and Future Challenges. *arXiv preprint arXiv:2103.11585* **2021**.
36. Hassani, H. Singular spectrum analysis: methodology and comparison **2007**.
37. Dong, Q.; Yang, Z.; Chen, Y.; Li, X.; Zeng, K. Anomaly detection in cognitive radio networks exploiting singular spectrum analysis. International Conference on Mathematical Methods, Models, and Architectures for Computer Network Security. Springer, 2017, pp. 247–259.
38. Xu, R.; Chen, Y.; Blasch, E.; Chen, G. A federated capability-based access control mechanism for internet of things (IOTs). Sensors and Systems for Space Applications XI. International Society for Optics and Photonics, 2018, Vol. 10641, p. 106410U.
39. Polunchenko, A.S.; Sokolov, G.; Du, W. Quickest change-point detection: A bird's eye view. *arXiv preprint arXiv:1310.3285* **2013**.
40. Yang, Z.; Chen, N.; Chen, Y.; Zhou, N. A novel PMU fog based early anomaly detection for an efficient wide area PMU network. 2018 IEEE 2nd International Conference on Fog and Edge Computing (ICFEC). IEEE, 2018, pp. 1–10.
41. Moskvina, V.; Zhigljavsky, A. Application of the singular spectrum analysis for change-point detection in time series. PhD thesis, CARDIFF, 2001.
42. Moskvina, V.; Zhigljavsky, A. An algorithm based on singular spectrum analysis for change-point detection. *Communications in Statistics-Simulation and Computation* **2003**, *32*, 319–352.

- 
43. Nagothu, D.; Xu, R.; Chen, Y.; Blasch, E.; Aved, A. DeFake: Decentralized ENF-Consensus Based DeepFake Detection in Video Conferencing. *IEEE 23rd International Workshop on Multimedia Signal Processing*; , 2021.
  44. Xu, R.; Nagothu, D.; Chen, Y. EconLedger: A Proof-of-ENF Consensus Based Lightweight Distributed Ledger for IoVT Networks. *Future Internet* **2021**, *13*, 248.
  45. Blanchard, P.; El Mhamdi, E.M.; Guerraoui, R.; Stainer, J. Machine learning with adversaries: Byzantine tolerant gradient descent. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 118–128.
  46. Merkel, D. Docker: lightweight linux containers for consistent development and deployment. *Linux journal* **2014**, *2014*, 2.
  47. Xu, R.; Nikouei, S.Y.; Nagothu, D.; Fitwi, A.; Chen, Y. BlendSPS: A BLockchain-ENabled Decentralized Smart Public Safety System. *Smart Cities* **2020**, *3*, 928–951.