Differential Privacy-based Blockchain for Industrial Internet of Things

Keke Gai, *Member, IEEE,* Yulu Wu, Liehuang Zhu, *Member, IEEE,* Zijian Zhang, *Member, IEEE,* Meikang Qiu, *Senior Member, IEEE*

Abstract—Contemporarily, two emerging techniques, blockchain and edge computing, are driving a dramatical rapid growth in the field of *Internet-of-Things* (IoT). Benefits of applying edge computing is an adoptable complementarity for cloud computing; blockchain is an alternative for constructing transparent secure environment for data storage/governance. Instead of using these two techniques independently, in this paper, we propose a novel approach that integrates IoT with edge computing and blockchain, which is called *Blockchain-based Internet-of-Edge* (BIoE) model. The proposed model, designed for a scalable and controllable IoT system, sufficiently exploits advantages of edge computing and blockchain to establish a privacy-preserving mechanism while considering other constraints, such as energy cost. We implement experiment evaluations running on Ethereum. According to our data collections, the proposed model improves privacy protections without lowering down the performance in an energy-efficient manner.

Index Terms—Internet-of-Things, blockchain, privacy-preserving, differential privacy, edge computing, task allocation

1 Introduction

The extensiveness of mobile devices in contemporary society has been remarkably driving an integrated network environment, such as *Internet-of-Things* (IoT) [1]–[5]. Introducing cloud computing into IoT is widely accepted as a centralized computing solution, by which heavy computation tasks or multi-functions can be delivered. The cloud deployment aids those devices with limited computing capability to expand the scope of service offerings. A great impact has been made in various domains, e.g. tele-health [6], mobile banking, and mobile rentals. It is observable that a cloud solution is a presentation of centralized computing, which generally simplifies configuration of the hardware/software at the user end [7].

Despite numerous merits of implementing cloud computing, a pure centralized service deployment is not the only option any more, with the enhancement of the computation capability occurring in mobility. For instance, energy wastes taken place during the communication between end users and cloud servers may be not worthy, as some mobile devices or local nearby equipment can offer similar services with less energy costs. The prevalence of the local IoT deployment further strengthens the on-premise computing [8]. Thus, as an emerging technique, edge computing is becoming an alternative for supplementing cloud systems. In a connected environment, edge devices can be involved in an IoT system, so that it is adoptable to migrate tasks to idle edge machines [9], [10].

Moreover, another technology buzzword currently influencing the networking deployment is blockchain. The gene of blockchain is protecting privacy, as using aliases can mask true identities for all involvers in the blockchain system. In addition, blockchain is a type of decentralized ledger storage system, which determines the information stored on a chain is not changeable. Other characteristics of implementing blockchain mostly derive from its decentralization setting, such as fault tolerance, attack resistance, and avoiding third-party risks. Considering the environment of IoT, blockchain can efficiently play a role of storing information by its decentralization features.

Even though widespread utilizations of emerging technologies above, building up a blueprint of integrating IoT with edge computing and blockchain still is ambiguous in presentation. The interconnection between these mechanisms is not perceivable yet. This paper focuses on developing a privacy-preserving scalable task allocation strategy as well as considers other constraints in practice. The proposed approach is called *Blockchain-based Internet-of-Edge* (BIoE) model. Fig. 1 presents the architecture of our model that exhibits a high level structure of BIoE. Blockchain, as shown in the figure, is taking a critical role who takes responsibility for information saving and participating in the task allocation. A few parameters can be applied during the decision making, such as edge nodes,

- *K. Gai (first author), Y. Wu, L. Zhu and Z. Zhang are with the School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China, 100081, {gaikeke@bit.edu.cn, 2120171080@bit.edu.cn, liehuangz@bit.edu.cn, zhangzijian@bit.edu.cn}.*
- *M. Qiu is with the Department of Electrical Engineering, Columbia University, New York, NY, USA, 10027, qiumeikang@yahoo.com.*
- *L. Zhu is the corresponding author (liehuangz@bit.edu.cn).*
- *This work is supported by National Natural Science Foundation of China (Grant No. 61972034), and partially supported by Beijing Institute of Technology Research Fund Program for Young Scholars (Dr. Keke Gai).*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2948094, IEEE Transactions on Industrial Informatics

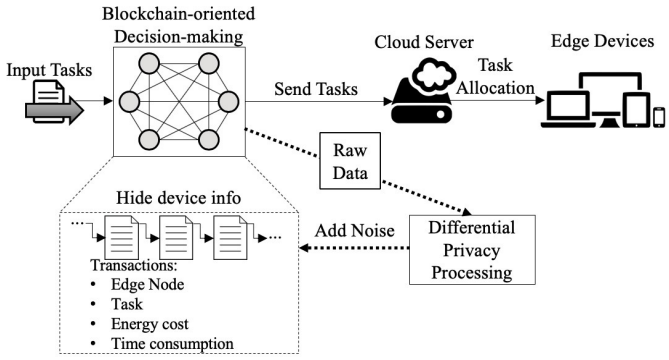

Fig. 1: The architecture of the proposed model.

tasks, energy cost or time consumption.

The primary intention of this work is to utilize blockchain to construct a privacy-preserving environment for task allocations. We bring in blockchain to the field to achieve a few goals, including enhancing edge nodes' trustworthiness, securing communications, and achieving optimal task allocations. As widely accepted, blockchain can contribute to decentralization applications as well as trustworthy services [11]–[13]. However, One of the drawbacks of directly using blockchain in the energy-related information storage is leaking edge nodes' identities. A differential privacy approach is proposed to protect identities in this work, by which noises are added to those data stored in blocks.

From the perspective of the task allocation, the objective of allocations considers both edge devices and cloud servers. Blockchain system provides a decision-making platform, in which all involvers have a fair opportunity to obtain the allocation, depending on the priority setting. We consider all equipment that takes part in task deliveries within the network as the involver. The key problem solved by our model is that it naturally hides the identity of the edge device during the task allocation in a cloud/edge scenario. In this work, we consider one-type constraint situation when configuring strategy priority, which combines both energy and time costs. Details of the proposed model are presented in succeeding sections.

Main contributions of this work are summarized as follows:

1) This work proposes a privacy-preserving approach that combines blockchain, edge computing, and IoT. We use the alias feature in blockchain so that all edge devices in the connected environment are changeable and scalable due to the decentralization setting. Our approach is a traceable mechanism for solving the task allocation problem in edge computing that utilizes blockchain's characteristics. Each block stores allocation information for the purpose of improving/adjusting tolerance errors.
2) This work has explored the implementation of differential privacy technique in a blockchain system, in order to prevent information on blocks from data mining-based attacks.

The rest of this paper is organized by the following order: Section 2 provides a brief summary about recent relevant studies. Next, details of the proposed model are given in Section 3. Moreover, we present core algorithms in our model in Section 4 and partial experiment results gathered from our evaluations are displayed and analyzed in Section 5. Finally, Section 6 concludes this work.

2 Related Work

2.1 Privacy Problem in Task Allocation

Many prior studies have evaluated that *Reinforcement Learning* (RL) techniques have been adopted in task allocations [14]–[16]. Among these attempts, a representative method is to learn feedbacks/ estimations from edge nodes to construct parameters and produce reward functions. A continuous learning generally is needed to achieve the optimal reward function. Even though many previous studies have explored to combine RL with task allocations, our research stands on a viewpoint that considers optimal task allocations as well as privacy protections.

Task allocation problem, known as task scheduling in some scenarios, is a classic research topic that recently has attracted more attentions due to the emergence of various service-oriented deployments, such as cloud computing and IoT [17], [18]. Most explorations on this field concentrates on optimization objectives [19], such as saving energy cost or increasing efficiency. Most optimization researches rarely considered privacy protection issues. In this section, we summarized previous representative privacy-driven studies in the problem of the task allocation.

We observed that many prior studies emphasized the balance between the security/privacy and the performance. A general research direction was to guarantee one facet's performance (e.g. privacy protection), while increasing that of the other facet, such as work efficiency. Thus, this sort of research mainly abstracts the optimization problem considering two or a few constraints. For instance, Gong et al. [20] presented an approach that supported mobile devices to share resources with the ad hoc mobile clouds while considering geographic information protections. This approach was based on differential privacy and balanced a few variables, such as privacy, utility, and system overhead.

In the embedded system environment, the work [21] accomplished an attempt on securing communications of the controller area network for automotive electronic systems. This work mainly solved a mapping problem that matched both security and safety requirements. In addition, the work [22] formulated the multiprocessor system-on-chip task scheduling into a multi-dimensional optimization problem, which considered a security constraint. The proposed heuristic algorithm could be also applied in a cloud implementation. Other work addressed a similar research topic included [23], [24].

Distinguishing from prior work, our work introduces the implementation of blockchain to naturally protect privacy of

the allocation infrastructure. One of the cores of the privacy protection in task allocations is hiding the identity of the device. There is no alternative mechanism in a blockchain system. System optimization can be well embedded in the operation of constructing blocks in our approach. The next section provides a brief review on blockchain techniques in contemporary cloud/edge systems.

2.2 Blockchain in Cloud/Edge Computing

Blockchain is a group of relevant techniques using blocks for transparent, traceable, and anti-tamper operations.Many recent studies have emphasized features of the blockchain, as mentioned above, to explore its potential implementations. We observed that a variety of domains were trying to utilize this technology to establish a safe, fair, and auditable environment.

One popular domain was using blockchain in cloud-based medical cares [25], [26]. The privacy leakage threat in this domain derives from the medical information sharing in a connected environment. Unknown physicians or unexpected service providers might had access to the data, albeit patients, known as data owners, had limited knowledge about the information usage behind the service interface. This type of the research had been explored by many recent studies [27].

For example, Xia et al. [28] developed a blockchain-based medical system that monitored the usage of medical data. The monitoring operations depended on recording all activities of the data access, which were implemented by constructing a block for each activity. This method utilized the traceability feature of blockchains, so that all data access-relevant actions were recorded in a tamper-proof manner. Another work [29] improved the security of medical data by using blockchain from a different perspective. This work used blockchain from a perspective of attribute-based signature; however, the operating principle was similar to Xie et al.'s approach [28], which also utilized blockchain's traceable characteristics. In each block, access actions were stored by recording data users' attributes and the results of matching multi-authority.

Using blockchain in medical data sharing was only a sample of research direction. As a matter of fact, there were many domains for the implementation of blockchain, such as data storage, auditing, financial service, etc [30]–[32]. Combining cloud/edge computing with blockchain was incondite, as two systems generally were running separately for distinct function units. Cloud/edge computing was mostly considered the medium of the service model [33]; blockchain was playing a role of constructing decentralizations for achieving certain functions [34]. Nevertheless, a seamless integration of blockchain and cloud/edge computing was rarely explored.

Besides, many fields have used blockchain technology to guarantee privacy-preserving. For example, Ke et al. [35] constructed a privacy-preserving blockchain-based mobile crowdsensing system by using both blockchain and edge computing. This system proposed a reputation management scheme in order to to protect user privacy and defend against malicious users. Axin et al. [36] proposed an efficient and privacy-preserving encryption scheme to trace users' attributes in the field of data sharing. This scheme used characteristics of blockchain to ensure integrity and non-repudiation of data to guarantee the security and privacy of data.

In summary, our work is an attempt on consolidating edge computing and blockchain into a new amalgamation form. Distinguishing from prior work, our attempt is trying to merge blockchain with edge computing, which enables blockchain to be a technical component of task allocations. We will present detailed description about our model in the next section.

3 Proposed Model

3.1 Threat Environment

Since all data stored in blocks are available to all blockchain users, the adversary of our model has a full access to the data. Data mining-based attack methods are deemed main malicious activities. Main threats defined in our model are twofold.

The first threat is the malicious user who launches malevolent activities for mapping infrastructure of edge computing. We assume that the adversary intends to have a stealthy look at the structure of the edge system. Even though data in the blockchain, which are edge nodes, tasks, energy cost and time consumption, are only limited to the record requirement, e.g. allocation plan, these accurate information can be used as a reference for launch a linkage attack. Considering an insider attack, all adversaries have legal roles to possess authorizations for requesting services and a supportive database is accessible. The infrastructure of the edge system can be mapped through numerous attempts launched by single/multiple malicious users.

In addition, we consider the other threat of the malicious user, but adversarial actions focus on stealing identity information, such as geographical data and energy cost information which will leak the identity information of edge nodes and let adversary know the certain tasks edge nodes execute. We assume that all adversaries have legal communication protocols with edge nodes, so that the communication time is measurable. It results in a chance of leaking geographic information as attackers can estimate the physical distance from measuring/-comparing latency.

3.2 Model Design

As described in prior sections, our BIoE model implements a differential privacy technique to screen raw allocation data and utilizes blockchain to support task allocations in edge computing. There are a number of crucial components in this model, which include *Smart Contract* (SC), *Edge Nodes* (ENs), *Optimization Server* (OpS) and blocks (for recording actions, energy cost and time consumption of edge nodes performing tasks). We provide a short summary for each role as follows.

The role of SC in our model is transferring tasks to OpS to process them between ENs and record corresponding actions, energy cost and time consumption. Besides, ENs need to perform tasks and compute time consumption and energy consuming which are provided to achieve instant reward. In addition,

TABLE 1: Main notations used in this paper and the definitions.

Notations	Definitions
SC	Smart Contract, referring to a kind of computer trading protocol executing contract terms and plays a role of transferring tasks and recording data
ENs	Edge Nodes, referring to edge devices which compose industrial internet of things
OpS	Optimization Server which is used to Q-learning technique to achieve task allocation for ENs
W_j	Tasks which need to be distributed to ENs to execute, where $j \in m$, m is the amount of tasks
N_i	Notation for edge nodes, where $i \in k$, k is the amount of ENs
E_{t+1}^{i/W_j}	Energy cost for EN N_i executing task W_j at time t+1
T_{t+1}^{i/W_j}	Time consumption for EN N_i executing task W_j at time t+1
r_{t+1}^{i}	Referring to the reward function at time t+1 when EN N_i executes task W_j
α	Balance factor between energy cost and time consumption for EN executing tasks which affects the weight of reward function
ϵ	The parameter to decide greedy algorithm
$n_{s,max}$	The amount of state set for Q-learning algorithm
$n_{s,t}$	The amount of states currently explored
ϵ_{min}	The minimum of exploration percentage of ϵ-greedy algorithm
ϵ_{max}	The maximum of exploration percentage of ϵ-greedy algorithm
$Q(s_t, a_t)$	Referring to the Q-value function for state s_t and action a_t at time t
γ	Referring to a discount factor to instant reward
$Q^*(s, a)$	The optimal Q-value function
$\pi^*(s)$	The optimal strategy
$DP(E_i^{W_j})$	The energy cost for EN N_i executing task W_j which is processed by differential privacy algorithm

OpS is used to execute corresponding Q-learning algorithm to achieve optimal Q-value function which is future reward and optimal strategy for allocating tasks to ENs and communicate with smart contract on blockchain. Finally, blockchain system consists of a chain of blocks for recording corresponding data. Our model designs the functionality of blocks for the purpose of the traceability and adjustment.

In our model, SC sends these tasks needed to be processed by ENs to OpS. OpS uses ϵ-greedy algorithm and strategy to choose an EN to complete a task and achieves energy cost and time consumption of EN completing corresponding task. OpS calculates the energy cost and time consumption to get instant reward and Q-value function. In addition, OpS achieves final reward and final optimal Q-value function and the optimal strategy determining certain ENs to complete certain tasks. Finally, the OpS will distribute corresponding tasks to ENs according to the optimal strategy. When ENs complete tasks, they will send actions, energy cost and time consumption of ENs performing tasks to OpS, where energy cost data is processed by differential privacy technology. Then, SC stores actions, energy cost and time consumption of ENs performing tasks on the blockchain.

We set some entities, tasks $W = \{W_1, W_2, \ldots, W_m\}$, ENs $N = \{N_1, N_2, \ldots, N_k\}$, SC on blockchain and OpS, where m is the amount of tasks and k is the amount of ENs. In our model, OpS is the agent. The state space S of our model is that uncertain EN executes uncertain task. And action space A is OpS chooses one EN to execute one task. We set reward function at time $T + 1$ by Eq.(1).

$$r_{t+1}^{i} = \alpha E_{t+1}^{i/W_j} + (1 - \alpha) T_{t+1}^{i/W_j}. \quad (1)$$

In Eq.(1), α is the balance factor between energy cost and time consumption which affects the weight of reward function. E_{t+1}^{i/W_j} and T_{t+1}^{i/W_j} is energy cost and time consumption by EN N_i executing task W_j. The larger α is, the greater the effect of energy cost on reward function is. The smaller α is, the greater the effect of time consumption on reward function is.

For task allocation, OpS chooses ϵ-greedy algorithm to choose ENs to process tasks. The ϵ-greedy algorithm can be represented by Eq.(2).

$$\epsilon = min(\epsilon_{max}, \epsilon_{min} + \mu(n_{s,max} - n_{s,t})/(n_{s,max}). \quad (2)$$

In Eq.(2), $\mu \in [0, 1]$, $n_{s,max}$ is the amount of state set, $n_{s,t}$ is the amount of states currently explored. ϵ_{min} and ϵ_{max} is the minimum and maximum of exploration percentage of ϵ-greedy algorithm respectively.

Our Q-value function is defined by Eq.(3).

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha(r + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)). \quad (3)$$

In Eq.(3), $\gamma \in [0, 1]$ is a discount factor which is important to instant reward. The larger γ is, more important Q-value function is. The smaller γ is, more important reward function is.

OpS will distribute tasks by the optimal strategy by Eq.(4).

$$\pi^*(s) = argmax Q^*(s, a). \quad (4)$$

In Eq.(4), $Q^*(s, a)$ means the optimal Q-value function.

After OpS achieves the optimal strategy $\pi^*(s)$, OpS will distribute tasks to corresponding ENs. We sets an four-tuple $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ to represent that an EN N_i executes an task W_j and produces energy cost $DP(E_i^{W_j})$ and time consumption $T_i^{W_j}$, where $DP(E_i^{W_j})$ means $E_i^{W_j}$ is processed by differential privacy algorithm. SC needs to store the four-tuple $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ on the blockchain for the purpose of traceability and adjustments to task allocation.

A differential privacy algorithm is used to protect data's privacy by adding noises into queried results. Currently differential privacy algorithm utilizes two main methods which are the Laplace's mechanism and the exponential mechanism. In our model, we uses the Laplace's mechanism [37] to add noises into energy cost set. In [37], their differential privacy protection method is represented by Eq.(5).

$$Pr[M(Q,D)=S] \propto exp(\frac{\epsilon}{\triangle Q}|S-Q(D)|_1). \quad (5)$$

In Eq.(5), it includes some parameters as follows. $M(\cdot)$ is the differential privacy protection algorithm based on the Laplace mechanism. The function Q is the mapping function of data set D. The set S is the noise set produced by $M(\cdot)$ which subjects to the Laplace's distribution and the mean. The function sensitivity $\triangle Q$ and the privacy protection level ϵ which is in direct proportion to the privacy protection degree and inversely proportional to data availability. Besides, their method uses the probability density function represented by Eq.(6) to make added noises subjecting to the Laplace's distribution.

$$Pr(x,\lambda) = \frac{1}{2\lambda} e^{\frac{-|x|}{\lambda}}. \quad (6)$$

In Eq.(6), $\lambda = \frac{\triangle Q}{\epsilon}$ and x is data. They use the differential privacy protection algorithm $M(\cdot)$ to add noises to queried results which is represented by Eq.(7).

$$\begin{aligned} M(Q,D) = Q(D) + (Lap_1(\frac{\triangle Q}{\epsilon}) \\ (Lap_2(\frac{\triangle Q}{\epsilon}), \ldots, (Lap_m(\frac{\triangle Q}{\epsilon})). \end{aligned} \quad (7)$$

In Eq.(7), $Lap_j(\frac{\triangle Q}{\epsilon})$, $i \in k$, is queried data's noise subjecting to the Laplace's distribution. In addition, $\triangle Q$ is the maximum of two queried results for two adjacent data set D and D' which can be represented by Eq.(8).

$$\triangle Q = max\{|Q(D) - Q(D')|_1\}. \quad (8)$$

In our model, we set the privacy protection level ϵ and achieve the data set E (energy cost set). Then, we generate the noise $Lap_j(\frac{\triangle Q}{\epsilon})$ to satisfy the probability density function $Pr(E_i^{W_j}, \lambda)$ and

$$Pr(E_i^{W_j}, \lambda) = \frac{1}{2\lambda} e \frac{-E_i^{W_j}}{\lambda}. \quad (9)$$

For Eq.(9), we set $\lambda = \frac{\triangle Q}{\epsilon}$ and $\triangle Q = E_i^{W_j}$. For protecting the privacy of ENs, we add noises $Lap_j(\frac{\triangle Q}{\epsilon})$ represented by Eq.(10).

$$DP(E_i^{W_j}) = E_i^{W_j} + Lap_j(\frac{\triangle Q}{\epsilon}). \quad (10)$$

We present main algorithms in the next section.

Algorithm 4.1 Allocation Task Algorithm

Require: Input tasks (W[m]), ENs (N(k)), discount factor γ, balance factor α

Ensure: The optimal strategy $\pi^*(s)$

1: OpS initial $Q(s,a)=0$, $\pi(s)$, $s=s_0$
2: **for** t=1,2,... **do**
3: OpS ensures its state s_t
4: OpS chooses an EN N_i to process a task W_j according to $\pi(s)$ and ϵ-greedy algorithm
5: N_i produces energy cost E_t^{i/W_j} and time consumption T_t^{i/W_j}
6: OpS computes reward function $r_t^i = \alpha E_t^{i/W_j} + (1-\alpha)T_t^{i/W_j}$
7: OpS computes $Q(s_t,a_t) = Q(s_t,a_t) + \alpha(r + \gamma Q(s_{t+1},a_{t+1}) - Q(s_t,a_t))$
8: OpS computes strategy $\pi(s) = argmaxQ(s,a_{t+1})$ and sets $s=s_t$
9: **end for**
10: $\pi^*(s) = \pi(s)$
11: **return** $\pi^*(s)$

4 ALGORITHMS

4.1 Allocation Tasks Algorithm

Allocation Tasks algorithm is designed for allocating tasks between edge nodes. We use a RL technique to construct an approximate optimal solution from iterations. The reason for implementing the RL technique is to increase the adoptability in a continuously changing application scenarios. A few recent studies [38], [39] have proved the performance of RL in task allocations. Algorithm 4.1 presents pseudo codes of Allocation Tasks algorithm. Main inputs of this algorithm include an input task, denoted by M[m]. An input ENs, denoted by N(k), discount factor γ and balance factor α, as shown in pseudo codes. The output of Allocation Tasks algorithm is the optimal strategy $\pi^*(s)$.

We present main phases of Allocation Tasks algorithm in the followings:

1) First, OpS needs to initial $Q(s,a)=0$ and set the strategy $\pi(s_0)$. Then, OpS needs to operate t times loops to achieve the optimal strategy $\pi^*(s)$.
2) In this loop, OpS first ensures its state s_t and choose an EN N_i to process a task W_j according to $\pi^*(s_t)$ and ϵ-greedy algorithm. N_i executes W_j and produces corresponding energy cost E_t^{i/W_j} and time consumption T_t^{i/W_j}. Then OpS uses E_t^{i/W_j} and T_t^{i/W_j} to compute reward function $r_t^i = \alpha E_t^{i/W_j} + (1-\alpha)T_t^{i/W_j}$. At last, OpS computes strategy $\pi(s_t) = argmaxQ(s,a_{t+1})$.
3) OpS needs to achieve the optimal strategy $\pi^*(s)$ to allocate tasks between ENs. So, $\pi^*(s)$ can be achieved when the t times loop is over and $\pi^*(s) = \pi(s)$.

In summary, this algorithm enables OpS to allocate tasks between ENs by a RL technique, Q-learning.

Algorithm 4.2 Allocation Block Creation Algorithm

Require: Input tasks (W[m]), ENs (N(k)) and the privacy protection level ϵ

Ensure: The four-tuple $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ set

1: SC sends tasks W[m] to OpS
2: OpS call Algorithm 4.1 to achieve the optimal strategy $\pi^*(s)$ to distributes W[m] to N[k]
3: OpS achieves $\langle N_i, W_j, E_i^{W_j}, T_i^{W_j} \rangle$ set
4: **for** j=0; j<m; j++ **do**
5: OpS generates the noise $Lap_j(\frac{\triangle Q}{\epsilon})$
6: OpS adds noise $Lap_j(\frac{\triangle Q}{\epsilon})$ to queried result $E_i^{W_j}$ to achieve $DP(E_i^{W_j})$ where $i \in k$
7: **end for**
8: OpS sends $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ set to SC
9: SC stores $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ set on the blockchain
10: **return** $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ set

4.2 Allocation Block Creation Algorithm

Allocation Block Creation Tasks algorithm is designed for that SC stores corresponding actions, energy cost and time consumption of an EN executing a task. Main inputs of the algorithm are tasks W[m], ENs N(k) and the privacy protection level ϵ. The output of the algorithm is the four-tuple $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ set.

We present main phases of Allocation Block Creation Tasks algorithm in the followings:

1) First, SC sends tasks W[m] to OpS. Then, OpS call Algorithm 4.1 to use Q-learning technique to achieve the optimal strategy $\pi^*(s)$ to distributes tasks W[m] to ENs N[k]. OpS chooses certain ENs to execute certain tasks to achieve energy cost and time consumption. OpS achieves $\langle N_i, W_j, E_i^{W_j}, T_i^{W_j} \rangle$ set.
2) To ensure the system's traceability, OpS needs to send SC $\langle N_i, W_j, E_i^{W_j}, T_i^{W_j} \rangle$ set. However, SC cannot achieve $E_i^{W_j}$ due to the privacy of ENs. Therefore we choose differential privacy protection method the Laplace mechanism to add noises to $E_i^{W_j}$ to guarantee the privacy of energy cost set. We generate the noise $Lap_j(\frac{\triangle Q}{\epsilon})$ which satisfies the probability density function $Pr(E_i^{W_j}, \lambda) = \frac{1}{2\lambda} e \frac{-E_i^{W_j}}{\lambda}$, where $\lambda = \frac{\triangle Q}{\epsilon}$ and $\triangle Q = E_i^{W_j}$. Then, OpS computes $DP(E_i^{W_j}) = E_i^{W_j} + Lap_j(\frac{\triangle Q}{\epsilon})$.
3) When OpS have computed $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ set, OpS sends $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ set to SC. SC will store $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ set on the blockchain.

In summary, this algorithm enables SC to stores the four-tuple $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ set on the blockchain to ensuring the traceability and adjustments of task allocation.

5 Evaluations and Findings

We implemented an experiment evaluation to assess the performance of our BIoE model. Energy-saving efficiency was not an emphasis in our evaluation, since our approach inherently had optimal solutions under the given criterion (energy cost counted in a unit time). In our evaluation, we emphasized the feasibility from the perspective of privacy protection so that energy cost, time consumption and energy cost processed by differential privacy protection are measured. We presented our experiment configuration and partial results in Section 5.1 and 5.2, respectively. Meanwhile, a security analysis was shown in Section 5.3 and a discussion about limitations and future work was given in Section 5.4.

5.1 Experiment Configuration

The principle of configuring our experiments was simulating an edge-based IoT application. Applying edge computing in IoT system generally consisted of a limited number of edge devices (nodes), which primarily was deployed for a specific purpose/system. Hence, our proposed model had a smaller scope of the edge node, which was distinct from a public-oriented blockchain system.

The software configuration included an Ethereum client Geth (1.8.3-stable) running on a computer (MacBook Pro 2017 version) as well as an Ethereum Wallet 0.10.0. The hardware configuration included an macOS 10.13.4 operating system, a CPU with 2.3GHz, an i5 version Intel Core, and a memory with 8 GB of 2133 MHz LPDDR3. The deterministic programming for task allocations and differential privacy protection algorithm were written by python computing language that was running on an Pycharm CE version 2017.3.1.

In the experiment evaluation, main measurement objectives included data packing-up time, gas cost for packing up data, energy cost of an EN executing a task, time consumption and energy cost processed by differential privacy protection. We presented a few evaluation results in the next section.

5.2 Experiment Results

Due to page length limit, we demonstrated partial evaluation results gathered from our experiments in this section. We selected a number of representative cases that had a variety of amount of tasks to depict the variation trend of the examined variable, which were 20, 50, 100 and 200.

Fig. 2 displayed results of the time length for packing up paired parameters to the block, which were ENs N_i set, tasks W_j set, energy cost processed by differential privacy protection $DP(E_i^{W_j})$ set and time consumption $T_i^{W_j}$ set. According to our data collections, the time scope of the block creation for packing up these data was within a period of 1s-15s. The trend of data packing up to blockchain is a linear growth, which depicted that the time costs had a positive relationship with the amount of tasks.

Besides, Fig. 3 showed the results of gas cost for packing up above data to the block. According to Fig. 3, the gas scope

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2948094, IEEE Transactions on Industrial Informatics

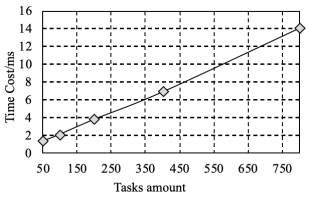

Fig. 2: Time cost for packing up the four-tuple set to blockchain.

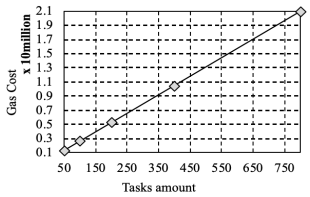

Fig. 3: Gas cost for packing up the four-tuple set to blockchain.

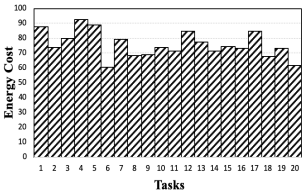

Fig. 4: Energy cost for an EN processing a task when the amount of tasks is 20.

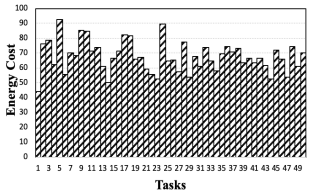

Fig. 5: Energy cost for an EN processing a task when the amount of tasks is 50.

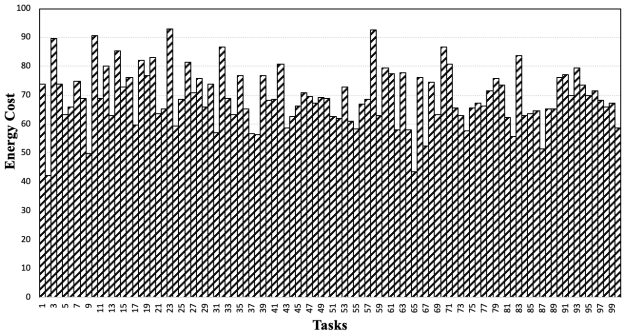

Fig. 6: Energy cost for an EN processing a task when the amount of tasks is 100.

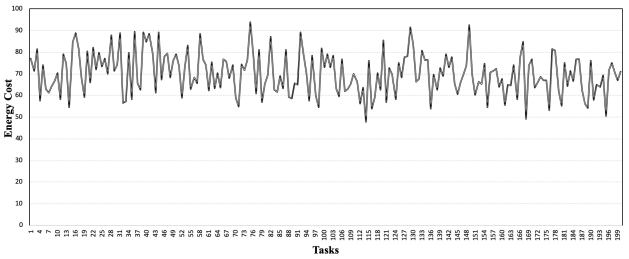

Fig. 7: Energy cost for an EN processing a task when the amount of tasks is 200.

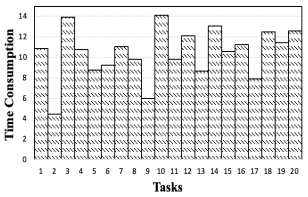

Fig. 8: Time consumption for an EN processing a task when the amount of tasks is 20.

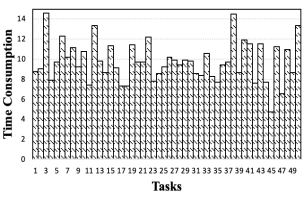

Fig. 9: Time consumption for an EN processing a task when the amount of tasks is 50.

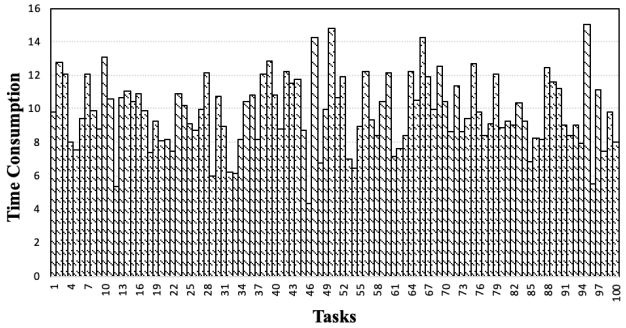

Fig. 10: Time consumption for an EN processing a task when the amount of tasks is 100.

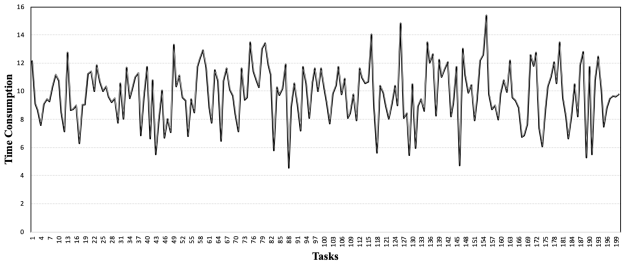

Fig. 11: Time consumption for an EN processing a task when the amount of tasks is 200.

of packing up data to blockchain was within a period of $1 \times 10^5 gas - 21 \times 10^5$gas. Also, gas cost of packing up data to blockchain is linear with the amount of tasks.

Fig.s 4 -7 exhibited energy cost of an EN executing a certain task which obeys the optimal strategy $\pi^*(s)$. The difference is the variety of the amount of tasks: 20, 50, 100 and 200. The amount of energy cost is related to the performance of ENs and the difficulty degree of completing tasks. Similarly, from Fig. 8 to Fig. 11, they displayed the time consumption of an EN executing a certain task at the amount of tasks: 20, 50, 100 and 200 . Also, the amount of time consumption is related to the performance of ENs and the difficulty degree of completing tasks.

Fig. 12, Fig. 13, Fig. 14 and Fig. 15 indicated $DP(E_i^{W_j})$ set which is energy cost set protected by differential privacy using the Laplace mechanism related to Fig. 4, Fig. 5 , Fig. 6 and Fig. 7. Depicted from these Figures, $DP(E_i^{W_j})$ was not much different from $E_i^{W_j}$ for the purpose of availability of energy cost set and $DP(E_i^{W_j})$ protected the privacy of ENs.

Fig.s 16 - 19 showed comparisons between the original

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2948094, IEEE Transactions on Industrial Informatics

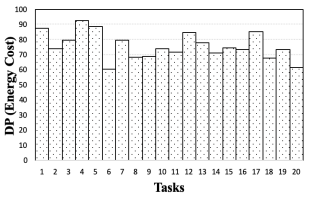

Fig. 12: Energy cost protected by differential privacy for an EN processing a task when the amount of tasks is 20.

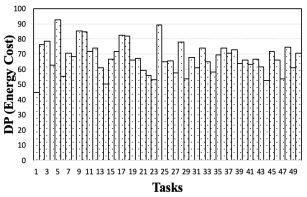

Fig. 13: Energy cost protected by differential privacy for an EN processing a task when the amount of tasks is 50.

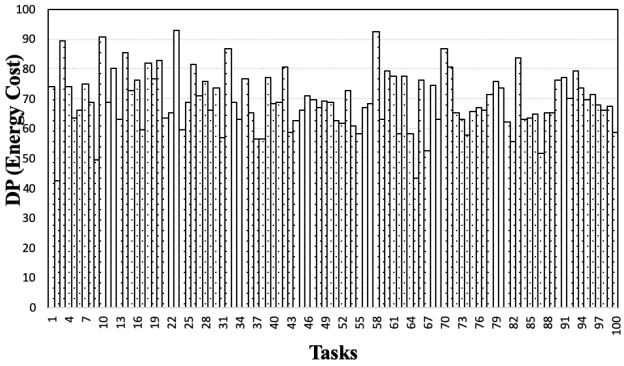

Fig. 14: Energy cost protected by differential privacy for an EN processing a task when the amount of tasks is 100.

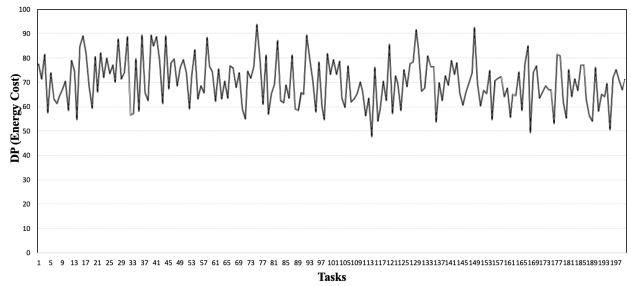

Fig. 15: Energy cost protected by differential privacy for an EN processing a task when the amount of tasks is 200.

energy cost data and DP-ed (energy cost) data, which are processed by differential privacy technology. In these figures, we could learn that difference value between energy cost and DP(energy cost) is among $-6 * 10^{-9} \sim 6 * 10^{-9}$. These difference values are not significant, but they could efficiently prevent malicious users from obtaining the original data and analyzing the location of edge nodes.

In summary, main findings of our evaluations included: (1) the amount of tasks in the system had a positive relationship with the block creation time lengths and gas cost; (2) energy cost and time consumption were correlate with performance of ENs and difficulty degree of completing tasks(3) differential privacy protection method had no effect on data availability.

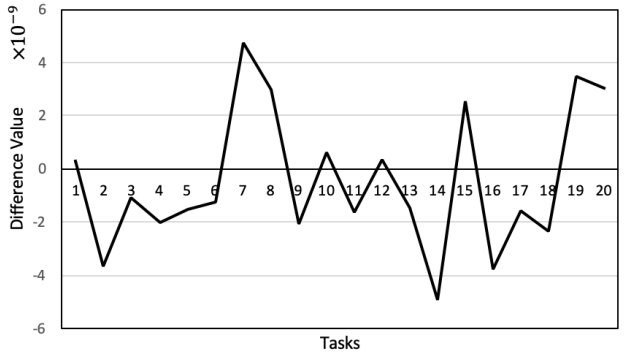

Fig. 16: Comparison between energy cost and DP (energy cost) when the amount of tasks is 20).

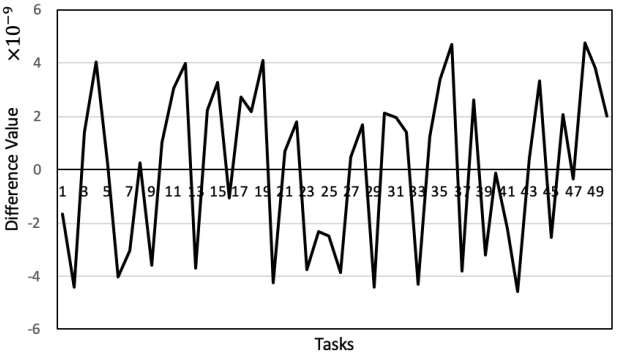

Fig. 17: Comparison between energy cost and DP (energy cost) when the amount of tasks is 50).

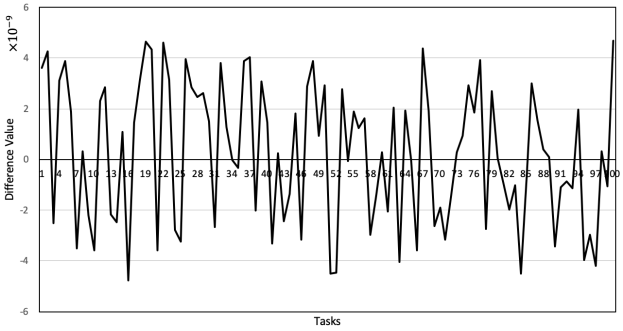

Fig. 18: Comparison between energy cost and DP (energy cost) when the amount of tasks is 100).

5.3 Security Analysis

We present a security analysis on the proposed differential privacy-based blockchain system, which are associated with the pre-defined threat environment given in Section 3.1. Based on the threat assumption, adversaries have a full access to all data stored in blocks. Two types of threats, malicious users launching malevolent activities for mapping infrastructure of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2948094, IEEE Transactions on Industrial Informatics

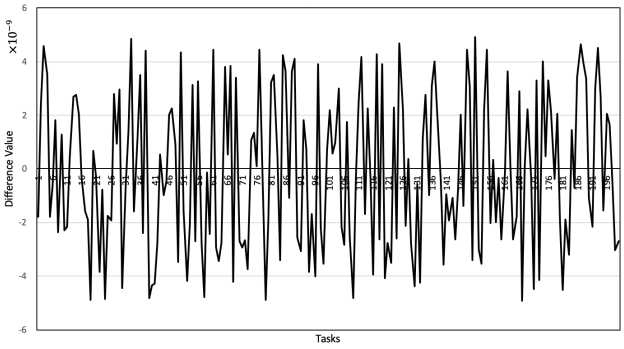

Fig. 19: Comparison between energy cost and DP (energy cost) when the amount of tasks is 200).

edge computing and malicious users stealing identity information of edge nodes, our model needs to face and solve them.

In our BIoE model, we use an OpS to distribute tasks to ENs and SC to store the four-tuple $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ on the blockchain. To the first type of threats, without adding nosies, adversaries can easily obtain real identities and behavior privacy from mining allocation information or launching a linkage attack to map infrastructure of edge computing by accessing supportive database from open source. Fortunately, our BIoE model uses a differential privacy protection method (Laplace distribution mechanism) to add noises into the energy cost set, such that a distortion is made on the protect target set. Our observation is that utilizing the Laplace mechanism can successfully screen the real allocation information of ENs while considering the requirement of iterations. Concerning data mining-based attacks, noises can increase the complexity of the feature extraction and information retrievals, from the perspective of the adversary. Moreover, added noises also are adoptable to protect users' identities to avoid the second type of threats, as data matching operations are hardly done between data stored on blockchain and other supportive database for processed data. Therefore, implementing our BIoE model can efficiently enhance the privacy-preserving capability.

5.4 Discussions

The novelty of our work was observable. We creatively fused a few emerging techniques, including edge computing, IoT, and blockchain. The proposed model could fully employ the advantage of each technique, so that it could guarantee task allocation manipulations while ensuring privacy protections and intelligent controls. Despite the novelty of our model, there were a few limitations/challenges that would be addressed in our future work. In this section, we highlighted a number of researchable problems that were found during our study.

We noticed that SC had limited relationship with the task allocation as task allocations were mostly completed by OpS, . The function of SC was storing the four-tuple $\langle N_i, W_j, DP(E_i^{W_j}), T_i^{W_j} \rangle$ on the blockchain. Because of above problems, we assumed that OpS is secure. The potential risk was tasks could be leaked to attackers when OpS was not secure.

The other problem we found was that noises added to energy cost set had not much effect on energy cost set's data availability. Our model aimed to guarantee the availability of data, noises we set could made a small impact on energy cost set. Because of noises we set, the privacy of ENs could be influenced by the degree of distortion to energy cost set not large. Addressing the problems above, our research would continuously explore our model from these perspectives in our future work.

6 Conclusions

This paper had a focus on designing a privacy-preserving scheme for implementing edge computing in IoT. The proposed approach was called BIoE model that utilized blockchain techniques in task allocations. Three designed objectives were achieved, which enabled an edge-based IoT system to be task allocation functional, privacy-preserving, and tamper-resistant. Our evaluation had depicted that our model could reach the design intention.

Acknowledgement

This work is supported by National Natural Science Foundation of China (Grant No. 61972034), and partially supported by Beijing Institute of Technology Research Fund Program for Young Scholars (Dr. Keke Gai).

References

[1] Q. Yang, B. Zhu, and S. Wu. An architecture of cloud-assisted information dissemination in vehicular networks. *IEEE Access*, 4:2764–2770, 2016.

[2] Z. Wang, H. Song, D. Watkins, K. Ong, P. Xue, Q. Yang, and X. Shi. Cyber-physical systems for water sustainability: challenges and opportunities. *IEEE Communications Magazine*, 53(5):216–222, 2015.

[3] H. Mahdikhani and R. Lu. Achieving privacy-preserving multi dot-product query in fog computing-enhanced IoT. In *2017 IEEE Global Communications Conference*, pages 1–6, Marina Bay Sands, Singapore, 2017. IEEE.

[4] T. Wang, J. Y. Zhou, A. F. Liu, M. Z. A. Bhuiyan, G. J. Wang, and W. J. Jia. Fog-based computing and storage offloading for data synchronization in iot. *IEEE Internet of Things Journal*, 6(3):4272–4282, 2019.

[5] T. Wang, L. Qiu, G. Xu, A. K. Sangaiah, and A. Liu. Energy-efficient and trustworthy data collection protocol based on mobile fog computing in internet of things. *IEEE Transactions on Industrial Informatics*, pages 1–1, 2019.

[6] C. Zhang, L. Zhu, C. Xu, and R. Lu. PPDP: An efficient and privacy-preserving disease prediction scheme in cloud-based e-healthcare system. *Future Generation Computer Systems*, 79:16–25, 2018.

[7] M. Díaz, C. Martín, and B. Rubio. State-of-the-art, challenges, and open issues in the integration of internet of things and cloud computing. *Journal of Network and Computer Applications*, 67:99–117, 2016.

[8] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao. A survey on internet of things: architecture, enabling technologies, security and privacy, and applications. *IEEE Internet of Things Journal*, 4(5):1125 – 1142, 2017.

[9] L. Lei, H. Xu, X. Xiong, K. Zheng, and W. Xiang. Joint computation offloading and multiuser scheduling using approximate dynamic programming in nb-iot edge computing system. *IEEE Internet of Things Journal*, 6(3):5345–5362, 2019.

[10] T. Wang, H. Luo, J. X. Zheng, and M. Xie. Crowdsourcing mechanism for trust evaluation in cpcs based on intelligent mobile edge computing. *ACM Transactions on Intelligent Systems and Technology*, pages 1–1, 2019.

[11] J. A. Jaroodi and N. Mohamed. Blockchain in industries: A survey. *IEEE Access*, 7:36500–36515, 2019.

[12] S. Biswas, K. Sharif, F. Li, B. Nour, and Y. Wang. A scalable blockchain framework for secure transactions in iot. *IEEE Internet of Things Journal*, 6(3):4650–4659, 2019.

[13] M. Dabbagh, M. Sookhak, and N. S. Safa. The evolution of blockchain: A bibliometric study. *IEEE Access*, 7:19212–19221, 2019.

[14] S. Lu, R. Tessier, and W. P. Burleson. Reinforcement learning for thermal-aware many-core task allocation. In *The 25th edition on Great Lakes Symposium on VLSI*, pages 379–384, 2015.

[15] D. B. Noureddine, A. Gharbi, and S. B. Ahmed. Multi-agent deep reinforcement learning for task allocation in dynamic environment. In *The 12th International Conference on Software Technologies*, pages 17–26, 2017.

[16] J. Yao and N. Ansari. Energy-aware task allocation for mobile iot by online reinforcement learning. In *2019 IEEE International Conference on Communications*, pages 1–6, 2019.

[17] C. Giovanelli, O. Kilkki, S. Sierla, I. Seilonen, and V. Vyatkin. Task allocation algorithm for energy resources providing frequency containment reserves. *IEEE Transactions on Industrial Informatics*, PP(99):1, 2018.

[18] J. Wang, Y. Wang, D. Zhang, F. Wang, H. Xiong, C. Chen, Q. Lv, and Z. Qiu. Multi-task allocation in mobile crowd sensing with individual task quality assurance. *IEEE Transactions on Mobile Computing*, PP(99):1, 2018.

[19] J. M. Luna, C. T. Abdallah, and G. L. Heileman. Probabilistic optimization of resource distribution and encryption for data storage in the cloud. *IEEE Transactions on Cloud Computing*, 6(2):428–439, 2018.

[20] Y. Gong, C. Zhang, Y. Fang, and J. Sun. Protecting location privacy for task allocation in ad hoc mobile cloud computing. *IEEE Transactions on Emerging Topics in Computing*, 6(1):110–121, 2018.

[21] C. Lin, Q. Zhu, C. Phung, and A. Sangiovanni-Vincentelli. Security-aware mapping for CAN-based real-time distributed automotive systems. In *Proceedings of the International Conference on Computer-Aided Design*, pages 115–121. IEEE Press, 2013.

[22] C. Liu, J. Rajendran, C. Yang, and R. Karri. Shielding heterogeneous MPSoCs from untrustworthy 3PIPs through security-driven task scheduling. *IEEE Transactions on Emerging Topics in Computing*, 2(4):461–472, 2014.

[23] S. Basu, M. Karuppiah, K. Selvakumar, K. C. Li, S. H. Islam, M. M. Hassan, and M. Z. A. Bhuiyan. An intelligent/cognitive model of task scheduling for iot applications in cloud computing environment. *Future Generation Computer System*, 88:254–261, 2018.

[24] L. Zeng, B. Veeravalli, and X. Li. SABA: A security-aware and budget-aware workflow scheduling strategy in clouds. *Journal of parallel and Distributed computing*, 75:141–151, 2015.

[25] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen. A survey on the security of blockchain systems. *Future Generation Computer Systems*, PP:1–1, 2017.

[26] H. Tian, J. He, and Y. Ding. Medical data management on blockchain with privacy. *J. Medical Systems*, 43(2):26:1–26:6, 2019.

[27] C. Esposito, A. De Santis, G. Tortora, H. Chang, and K.K.R. Choo. Blockchain: A panacea for healthcare cloud-based data security and privacy? *IEEE Cloud Computing*, 5(1):31–37, 2018.

[28] Q. Xia, E. Sifah, K. Asamoah, J. Gao, X. Du, and M. Guizani. MeDShare: Trust-less medical data sharing among cloud service providers via blockchain. *IEEE Access*, 5:14757–14767, 2017.

[29] R. Guo, H. Shi, Q. Zhao, and D. Zheng. Secure attribute-based signature scheme with multiple authorities for blockchain in electronic health records systems. *IEEE Access*, 776(99):1–12, 2018.

[30] J. Kang, R. Yu, X. Huang, S. Maharjan, Y. Zhang, and E. Hossain. Enabling localized peer-to-peer electricity trading among plug-in hybrid electric vehicles using consortium blockchains. *IEEE Transactions on Industrial Informatics*, 13(6):3154–3164, 2017.

[31] A. Lei, H. Cruickshank, Y. Cao, P. Asuquo, C. Ogah, and Z. Sun. Blockchain-based dynamic key management for heterogeneous intelligent transportation systems. *IEEE Internet of Things Journal*, 4(6):1832–1843, 2017.

[32] K. Christidis and M. Devetsikiotis. Blockchains and smart contracts for the Internet of Things. *IEEE Access*, 4:2292–2303, 2016.

[33] X. Chen, L. Jiao, W. Li, and X. Fu. Efficient multi-user computation offloading for mobile-edge cloud computing. *IEEE/ACM Transactions on Networking*, 24(5):2795–2808, 2016.

[34] S. Underwood. Blockchain beyond bitcoin. *Communications of the ACM*, 59(11):15–17, 2016.

[35] K. Zhao, S. Tang, B. Zhao, and Y. Wu. Dynamic and privacy-preserving reputation management for blockchain-based mobile crowdsensing. *IEEE Access*, 7:74694–74710, 2019.

[36] A. Wu, Y. Zhang, X. Zheng, R. Guo, Q. Zhao, and D. Zheng. Efficient and privacy-preserving traceable attribute-based encryption in blockchain. *Annals of Telecommunications*, 74(7):401–411, Aug 2019.

[37] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography*, pages 265–284, 2006.

[38] J. Yang, X. You, G. Wu, M. M. Hassan, A. Almogren, and J. Guna. Application of reinforcement learning in UAV cluster task scheduling. *Future Generation Computer System*, 95:140–148, 2019.

[39] S. Mostafavi, F. Ahmadi, and M. A. Sarram. Reinforcement-learning-based foresighted task scheduling in cloud computing. *CoRR*, abs/1810.04718, 2018.

Keke Gai [M 17'-S 13'] received the B.Eng. degree majored in automation, from Nanjing University of Science and Technology, Nanjing, China, in 2004, the M.E.T. (Master's of Educational Technology) degree in educational technology from the University of British Columbia, Vancouver, BC, Canada, in 2010, the MBA degree in business Administration, in 2009, M.S. degree in information technology, in 2014, from the Lawrence Technological University, Southfield, MI, USA, and the Ph.D. degree in computer science from Pace University, New York, NY, USA.

He is currently an Associate Professor at the School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China. He has published 3 books and more than 100 peer-reviewed journal/conference papers, including 6 ESI Highly Cited Papers. He has been granted 5 IEEE Best Paper awards (e.g. TrustCom 18' and HPCC 18') and 2 IEEE Best Student Paper awards in recent 5 years. His research interests include cyber security, blockchain, edge computing, cloud computing, and reinforcement learning.

Yulu Wu is currently a Master student majored in Computer Science at the School of Computer Science & Technology, Beijing Institute of Technology. Her research interests include cybersecurity, blockchain, and cloud computing.

Liehuang Zhu received his Ph.D. degree in computer science from Beijing Institute of Technology, Beijing, China, in 2004, the M.E. (Master of Engineering) degree and B.E. (Bachelor of Engineering) degree from Wuhan University, Wuhan, China, in 2001 and 1998, respectively.

He is currently a professor at School of Computer Science & Technology, Beijing Institute of Technology, Beijing, China. He has published more than 100 peer-reviewed journal or conference papers, including 10+ IEEE/ACM Transactions papers (IEEE TIFS, IEEE TII, IEEE TVT, IEEE TSG, Information Sciences, IEEE Network, Computer & Security, etc.). He has been granted a number of IEEE Best Paper Awards, including IWQoS 17', TrustCom 18'. His research interests include security protocol analysis and design, wireless sensor networks, and cloud computing.

Zijian Zhang is an Associate Professor in the School of Computer Science and Technology at Beijing Institute of Technology now. He was a visiting scholar in the Ubiquitous Security and Privacy Research Laboratory, State University of New York at Buffalo. He has published more than 40 papers in prestigious journals and magazines, e.g., TPSC, TCC, INS. His paper has been granted as an IEEE Best Paper Award by 2017 IEEE/ACM International Symposium on Quality of Service. His research interests include authentication and key agreement protocol, entity identification and behavior recognition.

Meikang Qiu [SM 07'] received the B.E. and M.E. degrees in engineering, from Shanghai Jiao Tong University, Shanghai, China, in 1992 and 1998, respectively, M.S. degree in computer science in 2003 and the Ph.D. degrees in computer science from the University of Texas at Dallas, Richardson, TX, USA, in 2007.

He is currently an Adjunct Professor at Columbia University, New York, NY, USA, and a Distinguished Professor in Computer Science at Shenzhen University, Shenzhen, China. He has published 15 books, 400 peer-reviewed journal/conference papers, and three registered patents. His research interests include cybersecurity, machine learning, big data, cloud computing, heterogeneous systems, and embedded systems.