01 Jan 2023

Distributed Detection Over Blockchain-Aided Internet Of Things In The Presence Of Attacks

Yiming Jiang

Jiangfan Zhang
Missouri University of Science and Technology, zhangjiangf@mst.edu

Recommended Citation

Y. Jiang and J. Zhang, "Distributed Detection Over Blockchain-Aided Internet Of Things In The Presence Of Attacks," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 3445 - 3460, Institute of Electrical and Electronics Engineers, Jan 2023.
The definitive version is available at https://doi.org/10.1109/TIFS.2023.3279984

Distributed Detection Over Blockchain-Aided Internet of Things in the Presence of Attacks

Yiming Jiang and Jiangfan Zhang, *Member, IEEE*

***Abstract*— Distributed detection over a blockchain-aided Internet of Things (BIoT) network in the presence of attacks is considered, where the integrated blockchain is employed to secure data exchanges over the BIoT as well as data storage at the agents of the BIoT. We consider a general adversary model where attackers jointly exploit the vulnerability of IoT devices and that of the blockchain employed in the BIoT. The optimal attacking strategy which minimizes the Kullback-Leibler divergence is pursued. It can be shown that this optimization problem is nonconvex, and hence it is generally intractable to find the globally optimal solution to such a problem. To overcome this issue, we first propose a relaxation method that can convert the original nonconvex optimization problem into a convex optimization problem, and then the analytic expression for the optimal solution to the relaxed convex optimization problem is derived. The optimal value of the relaxed convex optimization problem provides a detection performance guarantee for the BIoT in the presence of attacks. In addition, we develop a coordinate descent algorithm which is based on a capped water-filling method to solve the relaxed convex optimization problem, and moreover, we show that the convergence of the proposed coordinate descent algorithm can be guaranteed.**

***Index Terms*— Blockchain, double-spending attack, Internet of Things, distributed detection, Kullback-Leibler divergence, capped water-filling.**

I. Introduction

With the rapid development of smart devices and high-speed networks, the Internet of Things (IoT) has recently brought about an unprecedented increase in sensor resources and the deployment of sensor-like objects in safety-critical applications of vital societal interest, such as smart grids, healthcare informatics, manufacturing, and smart city [1], [2].

Typically, an IoT network consists of spatially distributed mutually-distrusting devices which sequentially generate and process exclusive data of a physical phenomenon of interest, and share their processed data with other devices over the network. In a conventional IoT (CIoT) network which is equipped with a cloud (or a fusion center), IoT devices transfer their data to a cloud where the IoT devices' data are stored and processed. Thereby the CIoT is vulnerable to a single point of failure since if the cloud does not function normally, the CIoT is paralyzed. The data stored in the cloud is also at risk of being modified or deleted by malicious attackers aiming to hack into the cloud. Moreover, the CIoT is vulnerable to some other security threats as well, including attacks on data exchanges between IoT devices and the cloud and impersonation of IoT devices. This has recently led to great interest in studying the vulnerability of the CIoT in various applications, see [3], [4], [5], [6], [7], [8], [9], [10] and the references therein.

Blockchain technology, an emerging secure distributed database technology that revolutionizes the way information is secured, distributed, and shared, has attracted enormous attention in recent years due to the following vital components [11], [12]: (i) a chronologically ordered sequence of blocks that are cryptographically linked to each other and are shared, stored and synchronized over a network; (ii) strong cryptography enabling secure data storage and secure data exchanges, and (iii) a mutual consensus protocol that enables verification and validation of the authenticity and the integrity of stored and exchanged data, and thus enables mutual trust over a network instead of relying on a central authority.

By taking advantage of the security-by-design and distributed nature without needing any central authority, blockchain lately has been integrated into IoT networks, and this kind of newly emerging blockchain-aided IoT (BIoT) network has been applied to a great deal of security-related applications, such as smart grids [13], [14], vehicular networks [15], [16], and smart city [17], [18]. Blockchain provides feasible solutions to address many common security threats to IoT networks. For example, in a blockchain network, by virtue of a consensus protocol, each node maintains a local copy of a blockchain which can be guaranteed to be identical to other nodes' copies. If one node's copy of the blockchain is maliciously corrupted, it can be retrieved from other nodes in the blockchain network, and hence a single point of failure can therefore be prevented. Moreover, the communications over a blockchain network are secured by the cryptographic algorithms of the blockchain network which can prevent attacks that manipulate either transmitted messages or the identities of senders, such as the man-in-the-middle attack and the Internet protocol address spoofing attack [7], [8], [19], [20].

However, blockchain technologies cannot eradicate all security threats to the BIoT, and the vulnerability of the BIoT is determined by the vulnerability of IoT devices and that of the blockchain employed in the BIoT. This is because on one

Manuscript received 3 February 2023; revised 20 April 2023; accepted 10 May 2023. Date of publication 25 May 2023; date of current version 12 June 2023. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Dusit Niyato. *(Corresponding author: Jiangfan Zhang.)*

The authors are with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO 65409 USA (e-mail: yjk7z@mst.edu; jiangfanzhang@mst.edu).

Digital Object Identifier 10.1109/TIFS.2023.3279984

hand, IoT devices in a BIoT are typically resource-limited and low-cost which makes them vulnerable to hacking. If hacked IoT devices in a BIoT deliberately transmit misleading data, then any task performed by the BIoT which unwittingly utilizes the data from the hacked IoT devices can be significantly compromised. On the other hand, the data stored in a blockchain is not perfectly secure. As mentioned in [21], [22], and [23], blockchains are prone to the double-spending attack (DSA), which is considered one of the most devastating attacks against blockchains. The DSA aims at falsifying the data which have already been stored in existing blocks of a blockchain. DSAs can be extremely injurious to blockchain networks. For example, one of the Bitcoin forks, Bitcoin Gold, suffered double-spending attacks in 2018, and again in 2020, with more than 17 million dollars lost in total. For a BIoT, a successful DSA can falsify the data stored in its blockchain without being perceived, and hence seriously compromise the data security of the BIoT.

Detection problems which aim to distinguish between different hypotheses are prevalent among BIoT applications, such as intrusion detection [15], [24], anomaly detection [9], [25], and object detection [17], [26]. In a BIoT, the blockchain which stores the data from IoT devices are shared, distributed, and synchronized across different agents. Hence, the decision between different hypotheses can be distributedly made at different agents which maintain local copies of the blockchain. Moreover, the decisions made by different agents can be guaranteed to reach a consensus since different agents' local copies of the blockchain are the same owing to the consensus protocol of the blockchain. In this paper, we are interested in distributed detection over a BIoT in the presence of attacks which jointly exploit the vulnerability of IoT devices and that of the blockchain employed in the BIoT.

A. Summary of Results and Main Contributions

We consider a detection task over a BIoT which aims to make a decision between two hypotheses based on IoT devices' data stored in its blockchain. A general BIoT model is considered which is generalized from most existing works on BIoT applications where a blockchain is integrated into an IoT network to secure data storage and exchanges [16], [18], [27], [28], [29], [30], [31], [32], [33], [34]. It is worth mentioning that if the task of the BIoT application which makes use of data stored in the employed blockchain has to be accomplished by a time instant, such as the detection task considered in this paper, then from the perspective of this task, the blockchain employed in the BIoT should be considered finitely long. For example, for a practical detection task, a decision should be made within a limited period of time by using a limited amount of data stored in a finite number of blocks of the blockchain. Once the decision is made, the growth of the blockchain can be regarded as being terminated from the perspective of the detection task since the data stored in the blocks that are generated after the decision is made do not affect this detection task.

Based on the general BIoT model, we adopt the Kullback-Leibler divergence (KLD) as the performance metric, which is one of the most popular performance metrics for detection problems [35], and develop the detection performance guarantee for the BIoT in the presence of attacks which jointly exploit the vulnerability of IoT devices and that of the blockchain employed in the BIoT. To be specific, considering the attacks which jointly attack the IoT devices and the blockchain of a BIoT, we pursue the detection performance guarantee by minimizing the KLD between two hypotheses over all possible malicious-data distributions. However, it can be shown that this minimization problem is nonconvex, and hence it is generally intractable to find the globally optimal solution for such a problem. To overcome this issue, we first propose a relaxation method to convert the nonconvex optimization problem into a convex optimization problem, and then the analytic expression for the optimal solution to the relaxed convex optimization problem is derived. In addition, we develop a coordinate descent algorithm which is based on a capped water-filling method to solve the relaxed convex optimization problem, and moreover, we show that the convergence of the proposed coordinate descent algorithm can be guaranteed. The optimal value of the relaxed convex optimization problem which is obtained from the proposed algorithm provides a performance guarantee for the BIoT in the presence of attacks.

B. Related Work

Recently, there has been great interest in integrating blockchain into diverse IoT applications to enhance data security, see [13], [16], [18], [27], [28], [29], [30], [31], [32], [33], [34], [36] and the references therein. In [33], the authors propose a blockchain-based vehicular announcement network where a blockchain is employed to protect data against tampering and make them widely available and accessible over the network. In [16], a peer-to-peer vehicle network is proposed where traffic information is stored in a blockchain which secures information sharing. The blockchains in these works are considered as highly secure, immutable, and transparent distributed ledgers so that the data stored in the blockchains are assumed to be perfectly secure in these works. However, the data stored in a blockchain cannot be perfectly secured. For example, blockchains are vulnerable to the DSA which, if successful, can falsify the data stored in the blockchain without being perceived. In contrast, in this paper, we take into account the vulnerability of the blockchain employed in a BIoT when studying the performance of BIoT applications.

The performance guarantee for distributed detection in the presence of attacks has been studied in various applications, see [5], [7], [37], [38] for instance. In [5], the KLD is adopted as the performance metric, and the authors develop the minimum KLD for distributed detection over a CIoT. The system and adversary models considered in [5] are different from those considered in this paper, and it is shown that the KLD minimization problem in [5] is a convex optimization problem. In contrast, we consider a distributed detection problem over a BIoT in the presence of attacks, and the minimization of the KLD for this problem can be shown to be nonconvex. In [37], the performance limit of collaborative spectrum sensing over a CIoT in the presence of attacks is analyzed where KLD is also adopted as the performance metric. A zero-sum game is employed in [37] to model the interaction between a fusion

center and an attacker. Since decision makers are assumed to be unable to interact with attackers in the model considered in this paper, the game theory approach developed in [37] cannot be applied to the problem considered in this paper.

The paper is organized as follows. In section II, the BIoT model and the adversary model are described. The optimal attacking strategy and the detection performance guarantee for the BIoT are investigated in Section III. Numerical simulations are presented in Section IV, and Section V provides our conclusions.

II. Blockchain-Aided IoT Detection Network And Adversary Models

A. Blockchain-Aided IoT Detection Network Model

We consider a general BIoT model which is generalized from most existing works on BIoT applications [13], [16], [18], [27], [28], [29], [30], [31], [32], [33], [34].

Unlike the CIoT where IoT devices send their data to a cloud for further processing, the BIoT employs a blockchain which eliminates the need of a central authority to verify and store IoT devices' data. According to functions, a BIoT which performs a detection task is usually composed of three types of agents, that is, IoT devices (or "things"), miners, and decision makers.

The first type of agent, the IoT device, forms the IoT layer of the BIoT. IoT devices are usually embedded with sensors, communication modules, software, and other technologies for the purpose of producing data and exchanging data with other agents over the BIoT. The second type of agent, the miner, forms the blockchain layer of the BIoT. The miners are assumed to have massive memory and computational power, and their duties are generating blocks which store the information produced by IoT devices. The third type of agent, the decision maker, aims to make an accurate and consistent decision based on the information stored in its local copy of the blockchain. It is worth mentioning that in a BIoT, the agents owning local copies of the blockchain can play the role of a decision maker and make their own decisions. Moreover, their decisions can be ensured to reach an agreement due to the blockchain consensus protocol. Hence, a single agent may play different roles simultaneously in a BIoT. For example, a miner can also play the role of a decision maker. In view of the popularity of the Proof-of-Work (PoW) consensus protocol, we assume that the BIoT considered in this paper employs the PoW consensus protocol. The BIoT employing other consensus protocols will be considered in future work. The working mechanism of the BIoT is similar to that of a PoW blockchain [21], [39], which are briefly summarized as follows.

1) Data Model: Let N denote the number of IoT devices in the IoT layer, and each IoT device makes measurements under the same unknown binary hypothesis (i.e., $\mathcal{H}_0$ or $\mathcal{H}_1$) over time. Let $\mathbf{x}_{j,l}$ denote the vector measurement made at the j-th IoT device in the l-th measurement sampling interval, $\forall l = 1, 2, \ldots$ and $\forall j = 1, 2, \ldots, N$. The j-th IoT device processes its raw measurement $\mathbf{x}_{j,l}$ to produce a discrete data $u_{j,l}$ by employing a function $\mathcal{Q}_j(\cdot)$, that is, $u_{j,l} = \mathcal{Q}_j(\mathbf{x}_{j,l}) \in \mathcal{K} \triangleq \{0, \ldots, |\mathcal{K}| - 1\}$ for each j and l, where $\mathcal{K}$ is the alphabet set of $u_{j,l}$ with cardinality equal to $|\mathcal{K}|$, and $u_{j,l}$ denotes the j-th IoT device's data produced in the l-th measurement sampling interval, $\forall j = 1, 2, \ldots, N$. It is worth mentioning that the process $\mathcal{Q}_j(\cdot)$ implemented at each IoT device may be necessary in practice. For example, $\mathcal{Q}_j(\cdot)$ can be an analog-to-digital converter employed at the j-th IoT device which is generally required for digital processing and digital communications. We assume that $\{u_{j,l}\}_{j,l}$ are statistically independent and identically distributed, and the alphabet sets of $u_{j,l}$ under $\mathcal{H}_1$ and $\mathcal{H}_0$ are the same. The probability mass functions (PMFs) of $u_{j,l}$ are denoted by

$$\mathbf{p}^H = [p_0^H, \ldots, p_{|\mathcal{K}|-1}^H]^T \text{ and } \mathbf{q}^H = [q_0^H, \ldots, q_{|\mathcal{K}|-1}^H]^T, \quad (1)$$

under hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively. As such, $\forall k = 0, 1, \ldots, |\mathcal{K}| - 1$,

$$\Pr\{u_{j,l} = k|\mathcal{H}_1\} = p_k^H > 0 \text{ and } \Pr\{u_{j,l} = k|\mathcal{H}_0\} = q_k^H > 0. \quad (2)$$

2) Data Exchange: In each measurement sampling interval, every IoT device produces a data $u_{j,l}$, and the data $u_{j,l}$ will be sent along with its index l to every miner. The communications over the BIoT are secured by asymmetric encryption which is similar to that of the PoW blockchain [21], [39]. To be specific, each agent of a BIoT owns a public-private key pair that forms the digital identity of the agent. The public key is created from the private key and is available at the other agents, while the private key is only available at its owner. A secure hash algorithm (SHA), e.g., SHA-256 and SHA-512 [40], is used in the data encryption process of every data exchange between two agents of the BIoT. Consider the data exchanges between the j-th IoT device and a miner as an example. In the l-th measurement sampling interval, the j-th IoT device firstly processes its message that contains data $u_{j,l}$ and the data index l by employing an SHA, and obtains a message digest. It then encrypts the message digest via its private key by using a digital signature algorithm, e.g., Elliptic Curve Digital Signature Algorithm [12], and produces a digital signature. Finally, the j-th IoT device sends the data package consisting of the message and the corresponding digital signature to the miners of the BIoT.

Once a miner receives the data package from an IoT device, it first decrypts the received digital signature via the public key of the IoT device, and obtains a message digest. Then, the miner processes the received message via the SHA to obtain another message digest. Only if these two message digests exactly match with each other, the authenticity of the received message from the IoT device is verified and the received message will be used in future processes. Otherwise, the received data package will be discarded and retransmission can take place. Note that the received digital signature at the miner can only be decrypted via the public key of the sender. It is worth mentioning that it is computationally intractable for an attacker to either find a different message which yields the same message digest or generate a valid digital signature for a fake message digest without the private key of the sender [12], [40]. Thus, the authenticity of the data package received at the miner and the identity of the sender can be validated and secured, which can prevent impersonation of the sender.

3) Block Mining and Consensus Protocol: Each miner maintains a local copy of the blockchain which is a chronologically ordered sequence of cryptographically linked blocks. For each l, after collecting and verifying the authenticity of the messages $\{u_{j,l}\}_{\forall j,l}$ from all the IoT devices, each miner constructs a local block which is a candidate for the l-th block of the blockchain. To be specific, each miner puts $\{u_{j,l}\}_{\forall j,l}$, the data index l, and the digital signatures associated with the IoT device messages into a block with a header. The header of a block consists of a discrete timestamp, a difficulty value, the hash value of the last block (parent block) of the longest branch in the miner's local copy of the blockchain, a Merkle Root which is the root of the Merkle tree constructed by recursively hashing pairs of data packages until there is only one hash [21], [39], and a number called nonce which is the solution to a puzzle problem. The hash value of the parent block in the header cryptographically links the block to its parent block.

Next, the miners compete with each other in solving a difficult PoW puzzle for their local blocks, which is called mining. The PoW puzzle is to find a nonce for a block such that the hash value of the block is smaller than a given value, i.e., the Difficulty in the header [21], [39]. Each miner searches for such a nonce for its local block via brute-force search. Once a miner solves its PoW puzzle, i.e., find a valid nonce value for its local block which meets the hash value requirement, it has successfully mined its local block, and it broadcasts its local block to all the other miners and the decision makers. After the other miners and the decision makers complete the verification of the authenticity of the data in the received block and the verification that the hash value of the received block is indeed smaller than the difficulty value and the data index stored in the received block is just one greater than that in its parent block, they add this block after its parent block in their local copies of the blockchain, and switch to work on mining the next block. The block mining process described above can be considered as a hashing competition among the miners, where the probability that a miner solves its PoW puzzle first among the miners is proportional to its hash rate which is defined as the ratio of its computational power to the total computational power in the network [21].

We assume that every decision maker has enough memory resources to store a full copy of the blockchain, and each decision maker only employs the IoT devices' data stored in the longest branch of its local copy of the blockchain to make its decision between two hypotheses. Note that for detection problems, it is generally true that the more the data, the smaller the decision error [35]. For a detection task, each decision maker only makes its decision when the number of blocks in the longest branch of its local copy of the blockchain grows to L so that the decision error can be guaranteed to meet some prescribed requirement. As such, from the perspective of this detection task, the growth of the blockchain can be considered being terminated once the longest branch of the blockchain employed in the BIoT grows to L blocks.

B. Adversary Model

The vulnerability of a BIoT comes from both the vulnerability of its IoT devices and that of the blockchain employed in the BIoT, which provides adversaries an opportunity to undermine the detection performance of the BIoT. In particular, with the goal of misleading the decision makers of the BIoT into making erroneous decisions, adversaries can jointly exploit both the vulnerability of IoT devices and that of the blockchain employed in the BIoT to falsify the data utilized by the decision makers.

1) Attacks Against IoT Devices: In order to falsify IoT devices' data in the blockchain without being perceived in the verification and validation processes described in Section II-A, an attacker first has to hack into IoT devices and obtain their private keys to generate valid falsified data packages[1]. If an IoT device has been hacked and controlled by an attacker, the attacker can use its private key to generate valid falsified data packages and send them to miners. Once the blocks which contain valid falsified data packages are successfully mined, the valid falsified data packages will be stored in the decision makers' local copies of the blockchain without being perceived.

Generally, cybersecurity measures are taken at IoT devices to keep attackers from hacking into the IoT devices and stealing their private keys. To this end, an attacker may not be able to hack into every IoT device within a limited time period. But the longer the period that the attacker spends on hacking into an IoT device, the higher the probability that the attacker successfully hacks into the target IoT device. For example, as a basic preventative measure, IoT devices can be equipped with password protection to prevent hacking. In consequence, an attacker has to employ a brute force attack which works through all possible keystrokes hoping to guess the password correctly. The longer the hacking period, the higher the probability that the attacker guesses the password correctly [41].

2) Attacks Against Blockchain: As pointed out in [21], [22], and [39], if an attacker aims to falsify the data which have already been stored in the blockchain of a BIoT, it has to launch a successful double-spending attack[2] on the blockchain. To be specific, the attacker has to generate valid counterfeit blocks[3] containing valid falsified data packages (i.e., solve PoW puzzles for the valid counterfeit blocks) to form a counterfeit branch in the blockchain which must surpass the authentic branch. A DSA is deemed successful if its counterfeit branch grows to L blocks before the authentic branch does, and hence its counterfeit branch is the longest branch in the blockchain when decision makers make their decisions. Under a successful DSA, the decisions will be made based on

[1] Valid falsified data packages mean those which can pass the cryptographic algorithms based authenticity verification process employed by the BIoT.

[2] Double-spending attacks were firstly introduced in financial blockchain applications [21], [39], [42]. That we use the same terminology is because the attacking mechanism here is the same as that of double-spending attacks against blockchain financial applications.

[3] If a block contains any falsified data package, this block is referred to as a counterfeit block. Otherwise, it is called an authentic block.

the falsified data stored in a counterfeit branch, and therefore, can be seriously misled.

In this paper, we consider an adversary model which jointly exploits the vulnerability of IoT devices and that of the blockchain employed in the BIoT. The IoT devices usually stay in sleeping mode during long intervals of inactivity to reduce energy consumption and reduce the risk of being targeted by an attacker [43], [44], [45]. Once the BIoT starts to work, its IoT devices wake up to make measurements of an unknown hypothesis and connect to the Internet to transfer their data to the miners in the network, which provides attackers an opportunity to hack into these IoT devices. We assume that there is a portion of miners, called malicious miners, which are under the command of an attacker. The attacker first commands the malicious miners to attempt to hack into every IoT device via the Internet, while the honest miners, which are not controlled by the attacker, follow the standard blockchain protocol to mine blocks for storing IoT devices' data, which form an authentic branch in the blockchain. At a time instant t_0, we assume that the probability that an IoT device has been hacked and controlled by the attacker is α ($\alpha \geq 0$), and the honest miners have built L_0 ($L_0 \geq 0$) blocks in the authentic branch of the blockchain. Thus, the expected percentage of IoT devices which are hacked and controlled by the attacker is α at t_0. If an IoT device has been hacked by the attacker, then the IoT device is called a malicious IoT device. Otherwise, it is called an honest IoT device. Starting from the time instant t_0, the malicious IoT devices' data can be deliberately falsified by the attacker since the attacker has already controlled these IoT devices and obtained their private keys to generate valid falsified data packages. As a result, the blocks of the authentic branch of the blockchain whose indices are greater than L_0 are counterfeit blocks since in these blocks, the data from the malicious IoT devices are falsified by the attacker without the need to launch a DSA. Next, in order to falsify the malicious IoT devices' data which have already been stored in the authentic branch's blocks whose indices are smaller than or equal to L_0, we assume that the attacker devotes all the computational power of the malicious miners to launching a DSA on the blockchain. Let $\mathcal{A}$ denote the set of malicious IoT devices, and we assume that the attacker aims at falsifying the malicious IoT devices' data $\{u_{j,l}\}_{j\in\mathcal{A}, L_A\leq l\leq L_0}$, which have been stored in the authentic branch's blocks with indices greater than or equal to L_A, to $\{\tilde{u}_{j,l}\}_{j\in\mathcal{A}, L_A\leq l\leq L_0}$ where[4] $0 \leq L_A \leq L_0$. For the cases where $L_0 > 0$, the counterfeit branch built by the malicious miners diverges from the authentic branch at the $(L_A - 1)$-th block of the authentic branch. It is worth mentioning that the counterfeit branch includes the authentic blocks with indices smaller than L_A and the counterfeit blocks in the counterfeit branch whose indices are greater than or equal to L_A. The authentic branch consists of the authentic blocks with indices from 1 to L_0 and the counterfeit blocks in the authentic branch whose indices are greater than or equal to $(L_0 + 1)$. Once the attacker extends the counterfeit branch to become longer than the authentic branch, all the honest miners switch from extending the authentic branch to working on extending the counterfeit branch according to the PoW consensus protocol [21], [22], [39], [42]. Therefore, the authentic branch stops growing, and the counterfeit branch will remain the longest branch in the blockchain as time goes by which implies that a successful DSA has been launched.

It is worth mentioning that the honest IoT devices' data in counterfeit blocks are authentic and cannot be falsified by the attacker since the attacker does not have their private keys to generate valid falsified data packages for the honest IoT devices. In addition, the attacker has to command the malicious miners to find a valid nonce value for each counterfeit block in the counterfeit branch so that the counterfeit block can be accepted by the decision makers and the other miners in their local copies of the blockchain due to the PoW consensus protocol. If the counterfeit branch built by the attacker is the longest branch in the blockchain when the decision makers make their decisions, the counterfeit branch is considered as the only valid branch by the decision makers according to the PoW consensus protocol. In consequence, the attacker may be able to fool the decision makers into reaching erroneous decisions since every decision maker employs the falsified IoT devices' data stored in the counterfeit branch to make its decision.

We assume that if the j-th IoT device is controlled by the attacker, the attacker falsifies the j-th IoT device's data $\{u_{j,l}\}_{l=L_A}^{L}$ to $\{\tilde{u}_{j,l}\}_{l=L_A}^{L}$ which is an independent and identically distributed sequence and follows the malicious-data PMFs

$$\mathbf{p}^B = [p_0^B, \ldots, p_{|\mathcal{K}|-1}^B]^T \text{ and } \mathbf{q}^B = [q_0^B, \ldots, q_{|\mathcal{K}|-1}^B]^T, \quad (3)$$

under hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively. As such, if the j-th IoT device is malicious, then $\forall l = L_A, L_A + 1, \ldots, L$ and $\forall k = 0, 1, \ldots, |\mathcal{K}| - 1$,

$$\Pr\{\tilde{u}_{j,l} = k|\mathcal{H}_1\} = p_k^B \quad \text{and} \quad \Pr\{\tilde{u}_{j,l} = k|\mathcal{H}_0\} = q_k^B. \quad (4)$$

It is worth mentioning that if $L_0 = L$, then when the attacker starts to launch a DSA, the authentic branch has already grown to L blocks, and each decision maker has already made its decision. Therefore, it is pointless for the attacker to launch a DSA to impair the detection performance of the BIoT, and hence the case that $L_0 = L$ is a trivial case. The case where $\alpha = 0$ is also a trivial case since if $\alpha = 0$, the attacker cannot falsify any data stored in the blockchain since the attacker does not obtain any IoT device's private key to generate valid falsified data. Before proceeding, to avoid these trivial cases, we make the following assumption throughout this paper.

Assumption 1: We assume that $0 < \alpha \leq 1$ *and* $L_A \leq L_0 < L$.

III. Optimal Attacking Strategy and Guaranteed Detection Performance

In this section, we investigate the detection performance of a BIoT in the presence of attacks, and pursue a guaranteed detection performance under any attacks. It can be shown that

[4] If $L_0 = 0$, then the attacker can falsify all the malicious IoT devices' data stored in the blockchain without the need to launch a DSA. Therefore, when $L_0 = 0$, we define $L_A = 0$ which signifies that the attacker does not launch a DSA. For the cases where $L_0 > 0$, L_A is greater than 0.

the probability that an attacker successfully launches a DSA is a constant depending on L_0, L_A, L, and the malicious miners' hash rate [46]. We use P_s to denote the probability that the attacker successfully launches a DSA on the BIoT.

When the longest branch of the blockchain grows to L blocks, let $\hat{\mathbf{u}}_j \triangleq [\hat{u}_{j,1}, \hat{u}_{j,2}, \ldots, \hat{u}_{j,L}]^T$ and $\mathbf{U} \triangleq [\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \ldots, \hat{\mathbf{u}}_N]$ denote the data of the j-th IoT device and all IoT devices' data stored in the longest branch of the blockchain, respectively. Note that $\hat{u}_{j,l} = u_{j,l}$, $\forall l \in \{1, 2, \ldots, L_A - 1\}$, while[5] $\forall l \in \{L_A^+, L_A^+ + 1, \ldots, L\}$, $\hat{u}_{j,l}$ can be either $u_{j,l}$ or $\tilde{u}_{j,l}$ depending on whether the j-th IoT device is malicious and whether the DSA launched by the attacker is successful. Define $\mathbf{R}_i \triangleq [\mathbf{r}_i^{(1)}, \mathbf{r}_i^{(2)}, \ldots, \mathbf{r}_i^{(N)}] \in \mathcal{R}$ where $\mathbf{r}_i^{(j)} \triangleq [r_{i,1}^{(j)}, r_{i,2}^{(j)}, \ldots, r_{i,L}^{(j)}]^T \in \mathcal{O}$, $\forall j \in \{1, 2, \ldots, N\}$, and $r_{i,l}^{(j)} \in \mathcal{K}$ for any i and l. $\mathcal{R}$ denotes the set of all possible $\mathbf{R}_i$ with cardinality $|\mathcal{R}| = |\mathcal{O}|^N$, and $\mathcal{O}$ denotes the set of all possible $\mathbf{r}_i^{(j)}$ with cardinality $|\mathcal{O}| = |\mathcal{K}|^L$. Under hypothesis $\mathcal{H}_1$, for any $\mathbf{R}_i \in \mathcal{R}$,

$$\begin{aligned}
&\Pr\{\mathbf{U} = \mathbf{R}_i|\mathcal{H}_1\} \\
&= \Pr\{\mathbf{U} = \mathbf{R}_i|\mathcal{E}, \mathcal{H}_1\} \times \Pr\{\mathcal{E}|\mathcal{H}_1\} \\
&\quad + \Pr\left\{\mathbf{U} = \mathbf{R}_i \middle| \mathcal{E}^{\mathrm{C}}, \mathcal{H}_1\right\} \times \Pr\left\{\mathcal{E}^{\mathrm{C}} \middle| \mathcal{H}_1\right\} \\
&= P_s \left(\prod_{j=1}^{N} \Pr\left\{\hat{\mathbf{u}}_j = \mathbf{r}_i^{(j)} \middle| \mathcal{E}, \mathcal{H}_1\right\}\right) \\
&\quad + (1 - P_s) \left(\prod_{j=1}^{N} \Pr\left\{\hat{\mathbf{u}}_j = \mathbf{r}_i^{(j)} \middle| \mathcal{E}^{\mathrm{C}}, \mathcal{H}_1\right\}\right), \qquad (5)
\end{aligned}$$

due to the fact that $\{\hat{\mathbf{u}}_j\}_{j=1}^N$ are statistically independent. The notation $\mathcal{E}$ in (5) stands for the event that the DSA is successful, and $\mathcal{E}^{\mathrm{C}}$ is the event that the DSA is not successful.

If the j-th IoT device is honest, then $\hat{u}_{j,l} = u_{j,l}, \forall l \in \{1, 2, \ldots, L\}$. On the other hand, if the j-th IoT device is malicious, then $\hat{u}_{j,l} = u_{j,l}, \forall l \in \{1, 2, \ldots, L_A - 1\}$ and $\hat{u}_{j,l} = \tilde{u}_{j,l}, \forall l \in \{L_0, L_0+1, \ldots, L\}$. Moreover, if the DSA launched by the attacker is successful, i.e., $\mathcal{E}$ happens, then $\hat{u}_{j,l} = \tilde{u}_{j,l}$, $\forall l \in \{L_A^+, L_A^+ + 1, \ldots, L_0 - 1\}$. Otherwise, $\hat{u}_{j,l} = u_{j,l}$, $\forall l \in \{L_A^+, L_A^+ + 1, \ldots, L_0 - 1\}$. As such, we can obtain[6]

$$\begin{aligned}
&\Pr\left\{\hat{\mathbf{u}}_j = \mathbf{r}_i^{(j)} \middle| \mathcal{E}, \mathcal{H}_1\right\} \\
&= \Pr\left\{\hat{\mathbf{u}}_j = \mathbf{r}_i^{(j)} \middle| j \notin \mathcal{A}, \mathcal{E}, \mathcal{H}_1\right\} \\
&\quad \times \Pr\{j \notin \mathcal{A}| \mathcal{E}, \mathcal{H}_1\} \\
&\quad + \Pr\left\{\hat{\mathbf{u}}_j = \mathbf{r}_i^{(j)} \middle| j \in \mathcal{A}, \mathcal{E}, \mathcal{H}_1\right\} \\
&\quad \times \Pr\{j \in \mathcal{A}| \mathcal{E}, \mathcal{H}_1\} \\
&= (1 - \alpha) \prod_{l=1}^{L} p_{r_{i,l}^{(j)}}^H + \alpha \prod_{l=1}^{L_A - 1} p_{r_{i,l}^{(j)}}^H \prod_{l=L_A^+}^{L} p_{r_{i,l}^{(j)}}^B, \qquad (6)
\end{aligned}$$

[5]We define $L_A^+ \triangleq \max\{L_A, 1\}$ to subsume the case where $L_A = 0$ and $L_0 = 0$.

[6]We define $\prod_{l_1}^{l_2}(\cdot) = 1$ if $l_2 < l_1$.

due to the fact that $\{\hat{u}_{j,l}\}_l$ are statistically independent, and similar to (6), we can obtain

$$\begin{aligned}
&\Pr\left\{\hat{\mathbf{u}}_j = \mathbf{r}_i^{(j)} \middle| \mathcal{E}^{\mathrm{C}}, \mathcal{H}_1\right\} \\
&\qquad = (1 - \alpha) \prod_{l=1}^{L} p_{r_{i,l}^{(j)}}^H + \alpha \prod_{l=1}^{L_0} p_{r_{i,l}^{(j)}}^H \prod_{l=L_0+1}^{L} p_{r_{i,l}^{(j)}}^B. \qquad (7)
\end{aligned}$$

From (5), (6), and (7), we can obtain that under $\mathcal{H}_1$, for any $\mathbf{R}_i \in \mathcal{R}$,

$$\begin{aligned}
&\Pr\{\mathbf{U} = \mathbf{R}_i|\mathcal{H}_1\} \\
&= P_s \left\{\prod_{j=1}^{N} \left[(1 - \alpha) \prod_{l=1}^{L} p_{r_{i,l}^{(j)}}^H + \alpha \prod_{l=1}^{L_A - 1} p_{r_{i,l}^{(j)}}^H \prod_{l=L_A^+}^{L} p_{r_{i,l}^{(j)}}^B\right]\right\} \\
&\quad + (1 - P_s) \left\{\prod_{j=1}^{N} \left[(1 - \alpha) \prod_{l=1}^{L} p_{r_{i,l}^{(j)}}^H + \alpha \prod_{l=1}^{L_0} p_{r_{i,l}^{(j)}}^H \right.\right. \\
&\quad \left.\left. \times \prod_{l=L_0+1}^{L} p_{r_{i,l}^{(j)}}^B\right]\right\} \qquad (8)
\end{aligned}$$

Similar to (8), we also can obtain that under hypothesis $\mathcal{H}_0$, for any $\mathbf{R}_i \in \mathcal{R}$,

$$\begin{aligned}
&\Pr\{\mathbf{U} = \mathbf{R}_i|\mathcal{H}_0\} \\
&= P_s \left\{\prod_{j=1}^{N} \left[(1 - \alpha) \prod_{l=1}^{L} q_{r_{i,l}^{(j)}}^H + \alpha \prod_{l=1}^{L_A - 1} q_{r_{i,l}^{(j)}}^H \prod_{l=L_A^+}^{L} q_{r_{i,l}^{(j)}}^B\right]\right\} \\
&\quad + (1 - P_s) \left\{\prod_{j=1}^{N} \left[(1 - \alpha) \prod_{l=1}^{L} q_{r_{i,l}^{(j)}}^H + \alpha \prod_{l=1}^{L_0} q_{r_{i,l}^{(j)}}^H \right.\right. \\
&\quad \left.\left. \times \prod_{l=L_0+1}^{L} q_{r_{i,l}^{(j)}}^B\right]\right\}. \qquad (9)
\end{aligned}$$

By evaluating (8) and (9) for all possible $\mathbf{R}_i$, we can obtain the probability mass functions $\mathbf{p} \triangleq [p_1, p_2, \ldots, p_{|\mathcal{R}|}]^T$ and $\mathbf{q} \triangleq [q_1, q_2, \ldots, q_{|\mathcal{R}|}]^T$ of $\mathbf{U}$ under hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively, where $p_i \triangleq \Pr\{\mathbf{U} = \mathbf{R}_i|\mathcal{H}_1\}$ and $q_i \triangleq \Pr\{\mathbf{U} = \mathbf{R}_i|\mathcal{H}_0\}$ are the probabilities of the event $\{\mathbf{U} = \mathbf{R}_i\}$ under hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively.

According to Stein's lemma, the KLD indicates the best error exponent of the miss probability under the Neyman-Pearson setup [47]. To this end, we choose the KLD as the performance metric for the BIoT, and pursue a guaranteed detection performance of the BIoT for all possible malicious-data distributions $\mathbf{p}^B$ and $\mathbf{q}^B$.

It is seen from (8) and (9) that the probability mass functions $\mathbf{p}$ and $\mathbf{q}$ are functions of $\mathbf{p}^B$ and $\mathbf{q}^B$ in (3), and hence, the KLD $D(\mathbf{q}||\mathbf{p})$ between $\mathbf{q}$ and $\mathbf{p}$ is also a function of $\mathbf{p}^B$ and $\mathbf{q}^B$,

which can be represented by $D_0(\mathbf{p}^B, \mathbf{q}^B)$, that is,

$$D_0\left(\mathbf{p}^B, \mathbf{q}^B\right) \triangleq D(\mathbf{q}||\mathbf{p}) = \sum_{i=1}^{|\mathcal{R}|} \Pr\{\mathbf{U}=\mathbf{R}_i|\mathcal{H}_0\} \log \frac{\Pr\{\mathbf{U}=\mathbf{R}_i|\mathcal{H}_0\}}{\Pr\{\mathbf{U}=\mathbf{R}_i|\mathcal{H}_1\}}. \tag{10}$$

A guaranteed KLD for all possible malicious-data distributions can be obtained by solving the following optimization problem

$$\min_{\mathbf{p}^B, \mathbf{q}^B} \; D_0\left(\mathbf{p}^B, \mathbf{q}^B\right) \tag{11a}$$

$$\text{s.t. } 0 \le p_k^B \le 1, 0 \le q_k^B \le 1, \quad \forall k \in \mathcal{K}, \tag{11b}$$

$$\sum_{k\in\mathcal{K}} p_k^B = 1, \quad \sum_{k\in\mathcal{K}} q_k^B = 1. \tag{11c}$$

In general, the objective function $D_0(\mathbf{p}^B, \mathbf{q}^B)$ in (11a) is a nonconvex function of $\mathbf{p}^B$ and $\mathbf{q}^B$. Thus, it is generally intractable to obtain the optimal solution to (11). To address this issue, we propose a relaxation method which converts the nonconvex optimization problem in (11) into a convex optimization problem, which is elaborated below.

Note that for each $\mathbf{R}_i \in \mathcal{R}$, $\Pr\{\mathbf{U}=\mathbf{R}_i|\mathcal{H}_1\}$ in (8) can be rewritten as

$$\Pr\{\mathbf{U}=\mathbf{R}_i|\mathcal{H}_1\} = \underbrace{(1-\alpha)^N \prod_{l_1=1}^{L} p^H_{r^{(1)}_{i,l_1}} \prod_{l_2=1}^{L} p^H_{r^{(2)}_{i,l_2}} \cdots \prod_{l_N=1}^{L} p^H_{r^{(N)}_{i,l_N}}}_{\triangleq \Psi_i} + \sum_{m\in\mathcal{M}} \Gamma_{i,m} P^{(1)}_{i,m} + \sum_{m\in\mathcal{M}} \Lambda_{i,m} P^{(2)}_{i,m}, \tag{12}$$

where Ψ_i is the probability that $\mathbf{U}=\mathbf{R}_i$ and all the N IoT devices are honest under $\mathcal{H}_1$. $\mathcal{M} \triangleq \{1, 2, \ldots, 2^N-1\}$ whose elements are used to specify all possible states of the N IoT devices except for the state that all the N IoT devices are honest. The state of the j-th IoT device is indicated by $S_m^N(j)$ which is the j-th digit of the N-bit-long binary sequence $\mathbf{S}_m^N$ converted from a decimal number m. We define that $S_m^N(j) = 1$ indicates that the j-th IoT device is malicious, and $S_m^N(j) = 0$ indicates that the j-th IoT device is honest. In (12), $\Gamma_{i,m}$, $P^{(1)}_{i,m}$, $\Lambda_{i,m}$, and $P^{(2)}_{i,m}$ are defined as[7]

$$\Gamma_{i,m} \triangleq P_s(1-\alpha)^{N-n_m}\alpha^{n_m} \prod_{j=1}^{N} \gamma^{(j)}_{i,m}, \tag{13}$$

$$P^{(1)}_{i,m} \triangleq \prod_{j=1}^{N} p^{(1)}_{i,j,m}, \tag{14}$$

$$\Lambda_{i,m} \triangleq (1-P_s)(1-\alpha)^{N-n_m}\alpha^{n_m} \prod_{j=1}^{N} \lambda^{(j)}_{i,m}, \tag{15}$$

$$P^{(2)}_{i,m} \triangleq \prod_{j=1}^{N} p^{(2)}_{i,j,m}, \tag{16}$$

[7]For simplicity of notation, we define $0^0 = 1$, and hence when $\alpha = 1$ and $n_m = N$ (i.e., $m = 2^N - 1$), we have $(1-\alpha)^{N-n_m} = 1$.

$$\gamma^{(j)}_{i,m} \triangleq \begin{cases} \prod_{l=1}^{L_A-1} p^H_{r^{(j)}_{i,l}}, & \text{if } S_m^N(j) = 1, \\ \prod_{l=1}^{L} p^H_{r^{(j)}_{i,l}}, & \text{if } S_m^N(j) = 0, \end{cases} \tag{17}$$

$$p^{(1)}_{i,j,m} \triangleq \begin{cases} \prod_{l=L_A^+}^{L} p^B_{r^{(j)}_{i,l}}, & \text{if } S_m^N(j) = 1, \\ 1, & \text{if } S_m^N(j) = 0, \end{cases} \tag{18}$$

$$\lambda^{(j)}_{i,m} \triangleq \begin{cases} \prod_{l=1}^{L_0} p^H_{r^{(j)}_{i,l}}, & \text{if } S_m^N(j) = 1, \\ \prod_{l=1}^{L} p^H_{r^{(j)}_{i,l}}, & \text{if } S_m^N(j) = 0, \end{cases} \tag{19}$$

$$p^{(2)}_{i,j,m} \triangleq \begin{cases} \prod_{l=L_0+1}^{L} p^B_{r^{(j)}_{i,l}}, & \text{if } S_m^N(j) = 1, \\ 1, & \text{if } S_m^N(j) = 0, \end{cases} \tag{20}$$

and n_m is the number of malicious IoT devices when the states of the N IoT devices are specified by $\mathbf{S}_m^N$. In other words,

$$n_m = \sum_{j=1}^{N} \mathbb{1}_{\{1\}}\left[S_m^N(j)\right], \tag{21}$$

where $\mathbb{1}_{\mathcal{X}}[x] = 1$, if $x \in \mathcal{X}$, and $\mathbb{1}_{\mathcal{X}}[x] = 0$, if $x \notin \mathcal{X}$. For example, if $N = 3$ and $m = 1$, then $\mathbf{S}_1^3 = 001$, $S_1^3(1) = 0$, $S_1^3(3) = 1$, and $n_1 = 1$. As such, $\Gamma_{i,m}P^{(1)}_{i,m}$ is the probability that $\mathbf{U}=\mathbf{R}_i$, the DSA is successful, and the states of the N IoT devices are specified by $\mathbf{S}_m^N$ under $\mathcal{H}_1$. $\Lambda_{i,m}P^{(2)}_{i,m}$ is the probability that $\mathbf{U}=\mathbf{R}_i$, the DSA is not successful, and the states of the N IoT devices are specified by $\mathbf{S}_m^N$ under $\mathcal{H}_1$.

Similar to (12), $\Pr\{\mathbf{U}=\mathbf{R}_i|\mathcal{H}_0\}$ in (9) can be rewritten as

$$\Pr\{\mathbf{U}=\mathbf{R}_i|\mathcal{H}_0\} = \underbrace{(1-\alpha)^N \prod_{l_1=1}^{L} q^H_{r^{(1)}_{i,l_1}} \prod_{l_2=1}^{L} q^H_{r^{(2)}_{i,l_2}} \cdots \prod_{l_N=1}^{L} q^H_{r^{(N)}_{i,l_N}}}_{\triangleq \Theta_i} + \sum_{m\in\mathcal{M}} \Delta_{i,m} Q^{(1)}_{i,m} + \sum_{m\in\mathcal{M}} \Phi_{i,m} Q^{(2)}_{i,m}, \tag{22}$$

where

$$\Delta_{i,m} \triangleq P_s(1-\alpha)^{N-n_m}\alpha^{n_m} \prod_{j=1}^{N} \delta^{(j)}_{i,m}, \tag{23}$$

$$Q^{(1)}_{i,m} \triangleq \prod_{j=1}^{N} q^{(1)}_{i,j,m}, \tag{24}$$

$$\Phi_{i,m} \triangleq (1-P_s)(1-\alpha)^{N-n_m}\alpha^{n_m} \prod_{j=1}^{N} \phi^{(j)}_{i,m}, \tag{25}$$

$$Q^{(2)}_{i,m} \triangleq \prod_{j=1}^{N} q^{(2)}_{i,j,m}, \tag{26}$$

$$\delta^{(j)}_{i,m} \triangleq \begin{cases} \prod_{l=1}^{L_A-1} q^H_{r^{(j)}_{i,l}}, & \text{if } S_m^N(j) = 1, \\ \prod_{l=1}^{L} q^H_{r^{(j)}_{i,l}}, & \text{if } S_m^N(j) = 0, \end{cases} \tag{27}$$

$$q^{(1)}_{i,j,m} \triangleq \begin{cases} \prod_{l=L_A^+}^{L} q^B_{r^{(j)}_{i,l}}, & \text{if } S_m^N(j) = 1, \\ 1, & \text{if } S_m^N(j) = 0, \end{cases} \tag{28}$$

$$\phi_{i,m}^{(j)} \triangleq \begin{cases} \prod_{l=1}^{L_0} q_{r_{i,l}^{(j)}}^{H}, & \text{if } S_m^N(j)=1, \\ \prod_{l=1}^{L} q_{r_{i,l}^{(j)}}^{H}, & \text{if } S_m^N(j)=0, \end{cases} \tag{29}$$

$$q_{i,j,m}^{(2)} \triangleq \begin{cases} \prod_{l=L_0+1}^{L} q_{r_{i,l}^{(j)}}^{B}, & \text{if } S_m^N(j)=1, \\ 1, & \text{if } S_m^N(j)=0. \end{cases} \tag{30}$$

In fact, $\Delta_{i,m}Q_{i,m}^{(1)}$ is the probability that $\mathbf{U}=\mathbf{R}_i$, the DSA is successful, and the states of the N IoT devices are indicated by $\mathbf{S}_m^N$ under $\mathcal{H}_0$. $\Phi_{i,m}Q_{i,m}^{(2)}$ is the probability that $\mathbf{U}=\mathbf{R}_i$, the DSA is not successful, and the states of the N IoT devices are specified by $\mathbf{S}_m^N$ under $\mathcal{H}_0$. Note that from their definitions, $P_{i,m}^{(1)}$, $P_{i,m}^{(2)}$, $Q_{i,m}^{(1)}$, and $Q_{i,m}^{(2)}$ are products of quantities between zero and one. Hence we have $P_{i,m}^{(1)}, P_{i,m}^{(2)}, Q_{i,m}^{(1)}, Q_{i,m}^{(2)} \in [0,1]$ for any $i=1,2,\ldots,|\mathcal{R}|$ and $m \in \mathcal{M}$.

For simplicity of notation, we define $\mathcal{I} \triangleq \{1,2,\ldots,|\mathcal{R}|\}$. Note from (2), (13), (15), (17), (19), (23), (25), (27), and (29) that for different $i \in \mathcal{I}$, the signs of $\Gamma_{i,m}$, $\Lambda_{i,m}$, $\Delta_{i,m}$, and $\Phi_{i,m}$ remain unchanged, respectively. Hence, we can define $\mathcal{M}_\Gamma \triangleq \{m \in \mathcal{M} | \Gamma_{i,m} \neq 0\}$, $\mathcal{M}_\Lambda \triangleq \{m \in \mathcal{M} | \Lambda_{i,m} \neq 0\}$, $\mathcal{M}_\Delta \triangleq \{m \in \mathcal{M} | \Delta_{i,m} \neq 0\}$, and $\mathcal{M}_\Phi \triangleq \{m \in \mathcal{M} | \Phi_{i,m} \neq 0\}$ for any $i \in \mathcal{I}$. We also define $\mathbf{P}_m^{(1)}$, $\forall m \in \mathcal{M}_\Gamma$, $\mathbf{P}_m^{(2)}$, $\forall m \in \mathcal{M}_\Lambda$, $\mathbf{Q}_m^{(1)}$, $\forall m \in \mathcal{M}_\Delta$, and $\mathbf{Q}_m^{(2)}$, $\forall m \in \mathcal{M}_\Phi$ as vectors stacking $\{P_{i,m}^{(1)}\}_{i\in\mathcal{I}}$, $\{P_{i,m}^{(2)}\}_{i\in\mathcal{I}}$, $\{Q_{i,m}^{(1)}\}_{i\in\mathcal{I}}$, and $\{Q_{i,m}^{(2)}\}_{i\in\mathcal{I}}$, respectively. Note from (12) that if $\Gamma_{i,m}=0$, then $P_{i,m}^{(1)}$ does not affect $\Pr\{\mathbf{U}=\mathbf{R}_i|\mathcal{H}_1\}$, and thus does not affect $D_0(\mathbf{p}^B,\mathbf{q}^B)$. Similarly, if either $\Lambda_{i,m}=0$, $\Delta_{i,m}=0$, or $\Phi_{i,m}=0$, then the corresponding $P_{i,m}^{(2)}$, $Q_{i,m}^{(1)}$, or $Q_{i,m}^{(2)}$ does not affect $D_0(\mathbf{p}^B,\mathbf{q}^B)$. In light of this, we define $\boldsymbol{\theta}$ as a one by $[\sum_{i\in\mathcal{I}}(|\mathcal{M}_\Gamma|+|\mathcal{M}_\Lambda|+|\mathcal{M}_\Delta|+|\mathcal{M}_\Phi|)]$ vector stacking $\{P_{i,m}^{(1)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Gamma}$, $\{P_{i,m}^{(2)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Lambda}$, $\{Q_{i,m}^{(1)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Delta}$, and $\{Q_{i,m}^{(2)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Phi}$ which are the parameters of $D_0(\mathbf{p}^B,\mathbf{q}^B)$. It is seen from (8), (9), (10), (12), and (22) that $D_0(\mathbf{p}^B,\mathbf{q}^B)$ can be rewritten as

$$\begin{aligned} D_1(\boldsymbol{\theta}) &\triangleq D_0\left(\mathbf{p}^B,\mathbf{q}^B\right) \\ &= \sum_{i\in\mathcal{I}}\left(\Theta_i + \sum_{m\in\mathcal{M}}\Delta_{i,m}Q_{i,m}^{(1)} + \sum_{m\in\mathcal{M}}\Phi_{i,m}Q_{i,m}^{(2)}\right) \\ &\quad \times \ln\frac{\Theta_i + \sum_{m\in\mathcal{M}}\Delta_{i,m}Q_{i,m}^{(1)} + \sum_{m\in\mathcal{M}}\Phi_{i,m}Q_{i,m}^{(2)}}{\Psi_i + \sum_{m\in\mathcal{M}}\Gamma_{i,m}P_{i,m}^{(1)} + \sum_{m\in\mathcal{M}}\Lambda_{i,m}P_{i,m}^{(2)}} \\ &= \sum_{i\in\mathcal{I}}\left(\Theta_i + \sum_{m\in\mathcal{M}_\Delta}\Delta_{i,m}Q_{i,m}^{(1)} + \sum_{m\in\mathcal{M}_\Phi}\Phi_{i,m}Q_{i,m}^{(2)}\right) \\ &\quad \times \ln\frac{\Theta_i + \sum_{m\in\mathcal{M}_\Delta}\Delta_{i,m}Q_{i,m}^{(1)} + \sum_{m\in\mathcal{M}_\Phi}\Phi_{i,m}Q_{i,m}^{(2)}}{\Psi_i + \sum_{m\in\mathcal{M}_\Gamma}\Gamma_{i,m}P_{i,m}^{(1)} + \sum_{m\in\mathcal{M}_\Lambda}\Lambda_{i,m}P_{i,m}^{(2)}}. \end{aligned} \tag{31}$$

Note that if $S_m^N(j)=1$ which indicates that the j-th IoT device is malicious, then $\gamma_{i,m}^{(j)}p_{i,j,m}^{(1)} = \prod_{l=1}^{L_A-1}p_{r_{i,l}^{(j)}}^{H}\prod_{l=L_A^+}^{L}p_{r_{i,l}^{(j)}}^{B} = \Pr\{\hat{\mathbf{u}}_j=\mathbf{r}_i^{(j)}|j\in\mathcal{A},\mathcal{E},\mathcal{H}_1\}$. If $S_m^N(j)=0$ which indicates that the j-th IoT device is honest, then $\gamma_{i,m}^{(j)}p_{i,j,m}^{(1)} = \prod_{l=1}^{L}p_{r_{i,l}^{(j)}}^{H} = \Pr\{\hat{\mathbf{u}}_j=\mathbf{r}_i^{(j)}|j\notin\mathcal{A},\mathcal{E},\mathcal{H}_1\}$. Hence, we can obtain that

$$\begin{aligned} \sum_{i\in\mathcal{I}}\prod_{j=1}^{N}\gamma_{i,m}^{(j)}p_{i,j,m}^{(1)} &= \sum_{i\in\mathcal{I}}\prod_{j=1}^{N}\Pr\{\hat{\mathbf{u}}_j=\mathbf{r}_i^{(j)}|\mathcal{Y}_j,\mathcal{H}_1\} \\ &= \sum_{\mathbf{r}_i^{(1)}\in\mathcal{O}}\sum_{\mathbf{r}_i^{(2)}\in\mathcal{O}}\cdots\sum_{\mathbf{r}_i^{(N)}\in\mathcal{O}}\prod_{j=1}^{N}\Pr\{\hat{\mathbf{u}}_j=\mathbf{r}_i^{(j)}|\mathcal{Y}_j,\mathcal{H}_1\} \\ &= 1, \end{aligned} \tag{32}$$

where $\mathcal{Y}_j \in \{\{j\in\mathcal{A},\mathcal{E}\},\{j\notin\mathcal{A},\mathcal{E}\}\}$. By following a similar argument, we can obtain

$$\sum_{i\in\mathcal{I}}\prod_{j=1}^{N}\lambda_{i,m}^{(j)}p_{i,j,m}^{(2)} = 1. \tag{33}$$

By employing (11c), (32), (33) and $\sum_{k\in\mathcal{K}}p_k^H=1$, we can obtain two constraints on $P_{i,m}^{(1)}$ and $P_{i,m}^{(2)}$ that

$$\begin{aligned} \sum_{i\in\mathcal{I}}\Gamma_{i,m}P_{i,m}^{(1)} &= P_s(1-\alpha)^{N-n_m}\alpha^{n_m}\sum_{i\in\mathcal{I}}\prod_{j=1}^{N}\gamma_{i,m}^{(j)}p_{i,j,m}^{(1)} \\ &= P_s(1-\alpha)^{N-n_m}\alpha^{n_m}, \quad \forall m\in\mathcal{M}_\Gamma, \end{aligned} \tag{34}$$

$$\begin{aligned} \sum_{i\in\mathcal{I}}\Lambda_{i,m}P_{i,m}^{(2)} &= (1-P_s)(1-\alpha)^{N-n_m}\alpha^{n_m}\sum_{i\in\mathcal{I}}\prod_{j=1}^{N}\lambda_{i,m}^{(j)}p_{i,j,m}^{(2)} \\ &= (1-P_s)(1-\alpha)^{N-n_m}\alpha^{n_m}, \forall m\in\mathcal{M}_\Lambda. \end{aligned} \tag{35}$$

Similar to (34) and (35), by employing (11c) and $\sum_{k\in\mathcal{O}}q_k^H=1$, we also can obtain two constraints on $Q_{i,m}^{(1)}$ and $Q_{i,m}^{(2)}$ that

$$\sum_{i\in\mathcal{I}}\Delta_{i,m}Q_{i,m}^{(1)} = P_s(1-\alpha)^{N-n_m}\alpha^{n_m}, \quad \forall m\in\mathcal{M}_\Delta, \tag{36}$$

$$\sum_{i\in\mathcal{I}}\Phi_{i,m}Q_{i,m}^{(2)} = (1-P_s)(1-\alpha)^{N-n_m}\alpha^{n_m}, \quad \forall m\in\mathcal{M}_\Phi. \tag{37}$$

Hence, the optimization problem in (11) can be relaxed to the following optimization problem

$$\min_{\boldsymbol{\theta}} \; D_1(\boldsymbol{\theta}) \tag{38a}$$

$$\text{s.t. } P_{i,m}^{(1)} \in [0,1], \quad \forall i\in\mathcal{I},\ \forall m\in\mathcal{M}_\Gamma, \tag{38b}$$

$$P_{i,m}^{(2)} \in [0,1], \quad \forall i\in\mathcal{I},\ \forall m\in\mathcal{M}_\Lambda, \tag{38c}$$

$$Q_{i,m}^{(1)} \in [0,1], \quad \forall i\in\mathcal{I},\ \forall m\in\mathcal{M}_\Delta, \tag{38d}$$

$$Q_{i,m}^{(2)} \in [0,1], \quad \forall i\in\mathcal{I},\ \forall m\in\mathcal{M}_\Phi, \tag{38e}$$

$$\text{constraints (34)–(37).}$$

Moreover, the optimal value of (38) provides a lower bound on that of (11a), and hence indicates a guaranteed performance for the distributed detection over the BIoT in the presence of attacks.

Lemma 1: Unlike the original optimization problem in (11) which is nonconvex, the optimization problem in (38) is a convex optimization problem.

Proof: Refer to Appendix A. ∎

Note that the feasible set of the problem in (38) is the intersection of the closed sets specified by (38b)–(38e) and the

affine hyperplanes specified by (34)–(37). Hence, the feasible set of the problem in (38) is closed. In addition, since the feasible set of the problem in (38) is bounded and the objective function in (38) is continuous, we know from Weirstrass' Theorem that an optimal solution to the optimization problem in (38) exists [48]. In the following theorem, we provide the analytic form of the optimal solution to the optimization problem in (38).

Theorem 1: The optimal solution $\{P_{i,m}^{*(1)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Gamma}$, $\{P_{i,m}^{*(2)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Lambda}$, $\{Q_{i,m}^{*(1)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Delta}$, *and* $\{Q_{i,m}^{*(2)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Phi}$ *to the optimization problem in (38) has the following analytic expressions that* $\forall i \in \mathcal{I}$,

$$P_{i,m}^{*(1)} = \min\left\{\left[\frac{\left(\begin{array}{c}\zeta_{\boldsymbol{p},m}^{*(1)} A_i^* - \Psi_i - \sum_{m'\in\mathcal{M}_\Lambda}\Lambda_{i,m'}P_{i,m'}^{*(2)} \\ -\sum_{m'\in\mathcal{M}_\Gamma\setminus\{m\}}\Gamma_{i,m'}P_{i,m'}^{*(1)}\end{array}\right)}{\Gamma_{i,m}}\right]^+, 1\right\}, \quad \forall m\in\mathcal{M}_\Gamma, \tag{39}$$

$$P_{i,m}^{*(2)} = \min\left\{\left[\frac{\left(\begin{array}{c}\zeta_{\boldsymbol{p},m}^{*(2)} A_i^* - \Psi_i - \sum_{m'\in\mathcal{M}_\Gamma}\Gamma_{i,m'}P_{i,m'}^{*(1)} \\ -\sum_{m'\in\mathcal{M}_\Lambda\setminus\{m\}}\Lambda_{i,m'}P_{i,m'}^{*(2)}\end{array}\right)}{\Lambda_{i,m}}\right]^+, 1\right\}, \quad \forall m\in\mathcal{M}_\Lambda, \tag{40}$$

$$Q_{i,m}^{*(1)} = \min\left\{\left[\frac{\left(\begin{array}{c}\zeta_{\boldsymbol{q},m}^{*(1)} B_i^* - \Theta_i - \sum_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)} \\ -\sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)}\end{array}\right)}{\Delta_{i,m}}\right]^+, 1\right\}, \quad \forall m\in\mathcal{M}_\Delta, \tag{41}$$

$$Q_{i,m}^{*(2)} = \min\left\{\left[\frac{\left(\begin{array}{c}\zeta_{\boldsymbol{q},m}^{*(2)} B_i^* - \Theta_i - \sum_{m'\in\mathcal{M}_\Delta}\Delta_{i,m'}Q_{i,m'}^{*(1)} \\ -\sum_{m'\in\mathcal{M}_\Phi\setminus\{m\}}\Phi_{i,m'}Q_{i,m'}^{*(2)}\end{array}\right)}{\Phi_{i,m}}\right]^+, 1\right\}, \quad \forall m\in\mathcal{M}_\Phi, \tag{42}$$

where $[x]^+ \triangleq \max\{0,x\}$ *for any* x, $A_i^* \triangleq \Theta_i + \sum_{m\in\mathcal{M}_\Delta}\Delta_{i,m}Q_{i,m}^{*(1)} + \sum_{m\in\mathcal{M}_\Phi}\Phi_{i,m}Q_{i,m}^{*(2)}$, *and* $B_i^* \triangleq \Psi_i + \sum_{m\in\mathcal{M}_\Gamma}\Gamma_{i,m}P_{i,m}^{*(1)} + \sum_{m\in\mathcal{M}_\Lambda}\Lambda_{i,m}P_{i,m}^{*(2)}$. $\zeta_{\boldsymbol{p},m}^{*(1)}$, $\zeta_{\boldsymbol{p},m}^{*(2)}$, $\zeta_{\boldsymbol{q},m}^{*(1)}$, *and* $\zeta_{\boldsymbol{q},m}^{*(2)}$ *are positive constants which ensure* $\sum_{i\in\mathcal{I}}\Gamma_{i,m}P_{i,m}^{*(1)} = P_s(1-\alpha)^{N-n_m}\alpha^{n_m}$, $\sum_{i\in\mathcal{I}}\Lambda_{i,m}P_{i,m}^{*(2)} = (1-P_s)(1-\alpha)^{N-n_m}\alpha^{n_m}$, $\sum_{i\in\mathcal{I}}\Delta_{i,m}Q_{i,m}^{*(1)} = P_s(1-\alpha)^{N-n_m}\alpha^{n_m}$, *and* $\sum_{i\in\mathcal{I}}\Phi_{i,m}Q_{i,m}^{*(2)} = (1-P_s)(1-\alpha)^{N-n_m}\alpha^{n_m}$, *respectively.*

Proof: Refer to Appendix B. ∎

Theorem 1 provides the analytic form of the optimal solution to the problem in (38). However, as shown in (39), (40), (41), and (42), the analytic expressions of $P_{i,m}^{*(1)}$, $P_{i,m}^{*(2)}$, $Q_{i,m}^{*(1)}$, and $Q_{i,m}^{*(2)}$ are coupled with each other. In light of this, we propose a coordinate descent algorithm to obtain the optimal solution to the problem in (38) which is summarized in Algorithm 1. In order to describe the Algorithm 1 concisely, we first define $\boldsymbol{\theta}^{(t)}$ as a vector stacking $\{P_{i,m,(t)}^{(1)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Gamma}$, $\{P_{i,m,(t)}^{(2)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Lambda}$, $\{Q_{i,m,(t)}^{(1)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Delta}$,

Algorithm 1

Initialization: Arbitrarily initialize the PMFs $\mathbf{p}^B$ and $\mathbf{q}^B$ in (3) and then initialize $\mathbf{P}_m^{(1)}$, $\forall m\in\mathcal{M}_\Gamma$, $\mathbf{P}_m^{(2)}$, $\forall m\in\mathcal{M}_\Lambda$, $\mathbf{Q}_m^{(1)}$, $\forall m\in\mathcal{M}_\Delta$, and $\mathbf{Q}_m^{(2)}$, $\forall m\in\mathcal{M}_\Phi$ in $\boldsymbol{\theta}$ by using (14), (16), (24), and (26), respectively. Update $\boldsymbol{\theta}^{(0)}$ with $\boldsymbol{\theta}$. Set $t=0$, and choose a small constant $\epsilon>0$.
Calculate $D_{(0)} \triangleq D_1(\boldsymbol{\theta}^{(0)})$ by using (31).
Repeat
for $m = 1,\ldots,2^N-1$ **do**
  Calculate n_m by using (21).
  for $i = 1,\ldots,|\mathcal{R}|$ **do** ▷ minimize over $\mathbf{P}_m^{(1)}$
    if $\Gamma_{i,m}\neq 0$ **then**
      Calculate A_i by using (43) with $\boldsymbol{\theta}\leftarrow\boldsymbol{\theta}^{(t)}$.
      Update $P_{i,m,(t+1)}^{(1)}$ by using right-hand side of (39) with $A_i^*\leftarrow A_i$, $P_{i,m'}^{*(2)}\leftarrow P_{i,m'}^{(2)}$, $m'\in\mathcal{M}_\Lambda$, and $P_{i,m'}^{*(1)}\leftarrow P_{i,m'}^{(1)}$, $m'\in\mathcal{M}_\Gamma\setminus\{m\}$.
    end if
  end for
  $\mathbf{P}_m^{(1)}\leftarrow\mathbf{P}_{m,(t+1)}^{(1)}$ and update $\boldsymbol{\theta}^{(t)}$ by replacing $\mathbf{P}_{m,(t)}^{(1)}$ with $\mathbf{P}_{m,(t+1)}^{(1)}$.
  for $i = 1,\ldots,|\mathcal{R}|$ **do** ▷ minimize over $\mathbf{P}_m^{(2)}$
    if $\Lambda_{i,m}\neq 0$ **then**
      Calculate A_i by using (43) with $\boldsymbol{\theta}\leftarrow\boldsymbol{\theta}^{(t)}$.
      Update $P_{i,m,(t+1)}^{(2)}$ by using right-hand side of (40) with $A_i^*\leftarrow A_i$, $P_{i,m'}^{*(1)}\leftarrow P_{i,m'}^{(1)}$, $m'\in\mathcal{M}_\Gamma$, and $P_{i,m'}^{*(2)}\leftarrow P_{i,m'}^{(2)}$, $m'\in\mathcal{M}_\Lambda\setminus\{m\}$.
    end if
  end for
  $\mathbf{P}_m^{(2)}\leftarrow\mathbf{P}_{m,(t+1)}^{(2)}$ and update $\boldsymbol{\theta}^{(t)}$ by replacing $\mathbf{P}_{m,(t)}^{(2)}$ with $\mathbf{P}_{m,(t+1)}^{(2)}$.
  for $i = 1,\ldots,|\mathcal{R}|$ **do** ▷ minimize over $\mathbf{Q}_m^{(1)}$
    if $\Delta_{i,m}\neq 0$ **then**
      Calculate B_i by using (44) with $\boldsymbol{\theta}\leftarrow\boldsymbol{\theta}^{(t)}$.
      Update $Q_{i,m,(t+1)}^{(1)}$ by using right-hand side of (41) with $B_i^*\leftarrow B_i$, $Q_{i,m'}^{*(2)}\leftarrow Q_{i,m'}^{(2)}$, $m'\in\mathcal{M}_\Phi$, and $Q_{i,m'}^{*(1)}\leftarrow Q_{i,m'}^{(1)}$, $m'\in\mathcal{M}_\Delta\setminus\{m\}$.
    end if
  end for
  $\mathbf{Q}_m^{(1)}\leftarrow\mathbf{Q}_{m,(t+1)}^{(1)}$ and update $\boldsymbol{\theta}^{(t)}$ by replacing $\mathbf{Q}_{m,(t)}^{(1)}$ with $\mathbf{Q}_{m,(t+1)}^{(1)}$.
  for $i = 1,\ldots,|\mathcal{R}|$ **do** ▷ minimize over $\mathbf{Q}_m^{(2)}$
    if $\Phi_{i,m}\neq 0$ **then**
      Calculate B_i by using (44) with $\boldsymbol{\theta}=\boldsymbol{\theta}^{(t)}$.
      Update $Q_{i,m,(t+1)}^{(2)}$ by using right-hand side of (42) with $B_i^*\leftarrow B_i$, $Q_{i,m'}^{*(1)}\leftarrow Q_{i,m'}^{(1)}$, $m'\in\mathcal{M}_\Delta$, and $Q_{i,m'}^{*(2)}\leftarrow Q_{i,m'}^{(2)}$, $m'\in\mathcal{M}_\Phi\setminus\{m\}$.
    end if
  end for
  $\mathbf{Q}_m^{(2)}\leftarrow\mathbf{Q}_{m,(t+1)}^{(2)}$ and update $\boldsymbol{\theta}^{(t)}$ by replacing $\mathbf{Q}_{m,(t)}^{(2)}$ with $\mathbf{Q}_{m,(t+1)}^{(2)}$.
end for
$\boldsymbol{\theta}^{(t+1)}\leftarrow\boldsymbol{\theta}^{(t)}$.
$t\leftarrow t+1$. Calculate $D_{(t)}\triangleq D_1(\boldsymbol{\theta}^{(t)})$ by using (31).
Until $|D_{(t)}-D_{(t-1)}|<\epsilon$.
Output: $D_1(\boldsymbol{\theta}^{(t)})$, $\boldsymbol{\theta}^{(t)}$.

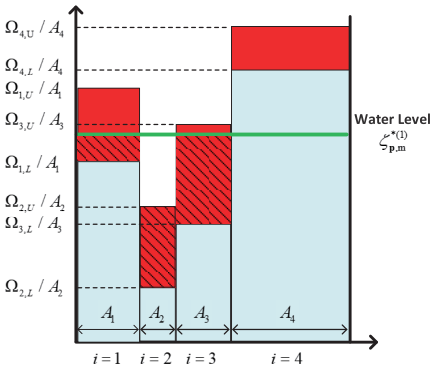

Fig. 1. The diagram of the capped water-filling method.

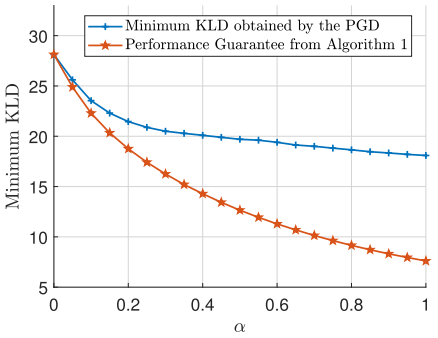

Fig. 2. Comparison between the minimum KLD and its guarantee.

and $\{Q^{(2)}_{i,m,(t)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Phi}$, and define $\mathbf{P}^{(1)}_{m,(t)}$, $\forall m\in\mathcal{M}_\Gamma$, $\mathbf{P}^{(2)}_{m,(t)}$, $\forall m\in\mathcal{M}_\Lambda$, $\mathbf{Q}^{(1)}_{m,(t)}$, $\forall m\in\mathcal{M}_\Delta$, and $\mathbf{Q}^{(2)}_{m,(t)}$, $\forall m\in\mathcal{M}_\Phi$ as vectors stacking $\{P^{(1)}_{i,m,(t)}\}_{i\in\mathcal{I}}$, $\{P^{(2)}_{i,m,(t)}\}_{i\in\mathcal{I}}$, $\{Q^{(1)}_{i,m,(t)}\}_{i\in\mathcal{I}}$, and $\{Q^{(2)}_{i,m,(t)}\}_{i\in\mathcal{I}}$, respectively. We also define

$$A_i \triangleq \Theta_i + \sum_{m\in\mathcal{M}_\Delta} \Delta_{i,m} Q^{(1)}_{i,m} + \sum_{m\in\mathcal{M}_\Phi} \Phi_{i,m} Q^{(2)}_{i,m}, \tag{43}$$

$$B_i \triangleq \Psi_i + \sum_{m\in\mathcal{M}_\Gamma} \Gamma_{i,m} P^{(1)}_{i,m} + \sum_{m\in\mathcal{M}_\Lambda} \Lambda_{i,m} P^{(2)}_{i,m}. \tag{44}$$

Theorem 2: Algorithm 1 can converge to the globally optimal solution to the problem in (38).

Proof: Since the problem in (38) is a convex optimization problem and the objective function (38a) is twice continuously differentiable, the coordinate descent method can converge to its optimal solution [49]. ∎

The constants $\zeta^{*(1)}_{\mathbf{p},m}$, $\zeta^{*(2)}_{\mathbf{p},m}$, $\zeta^{*(1)}_{\mathbf{q},m}$, and $\zeta^{*(2)}_{\mathbf{q},m}$ in (39)–(42) can be obtained by employing a variant of the water-filling procedure which is referred to as a capped water-filling method. We take Step 9 of Algorithm 1 as an example to describe this procedure, which is illustrated in Fig. 1. Note that in Step 9 of Algorithm 1, $\{\mathbf{P}^{(1)}_{m'}\}_{m'\in\mathcal{M}_\Gamma\setminus\{m\}}$, $\{\mathbf{P}^{(2)}_{m'}\}_{m'\in\mathcal{M}_\Lambda}$, $\{\mathbf{Q}^{(1)}_{m'}\}_{m'\in\mathcal{M}_\Delta}$, $\{\mathbf{Q}^{(2)}_{m'}\}_{m'\in\mathcal{M}_\Phi}$ are considered as known constants. For each $i\in\mathcal{I}$, we first calculate A_i by using (43). Then we calculate the minimum B_i by using (44) with $P^{(1)}_{i,m}=0$ and the maximum B_i by using (44) with $P^{(1)}_{i,m}=1$, which are denoted as $\Omega_{i,L}$ and $\Omega_{i,U}$, respectively. In Fig. 1 where $|\mathcal{I}|=4$, for each $i\in\mathcal{I}$, we first draw a rectangle with the area, the base, and the height equal to $\Omega_{i,U}$, A_i, and $\Omega_{i,U}/A_i$, respectively.

Then we start to gradually increase the level $\zeta^{*(1)}_{\mathbf{p},m}$ of "water" from its initial value zero. Once $\zeta^{*(1)}_{\mathbf{p},m}$ reaches $\Omega_{i,L}/A_i$, the water begins to fill the portion of the i-th rectangle that is above $\Omega_{i,L}/A_i$ (see the red parts of the rectangles in Fig. 1). As $\zeta^{*(1)}_{\mathbf{p},m}$ increases, if $\zeta^{*(1)}_{\mathbf{p},m}$ passes $\Omega_{i,U}/A_i$, then the portion of the i-th rectangle that is above $\Omega_{i,L}/A_i$ has been filled up, and hence the water can no longer fill the i-th rectangle. The water in the rectangles is indicated by the rectangles filled with inclined lines in Fig. 1. The process of increasing $\zeta^{*(1)}_{\mathbf{p},m}$ is stopped when the total area of the rectangles filled with inclined lines is equal to $P_s(1-\alpha)^{N-n_m}\alpha^{n_m}$, and the corresponding value of $\zeta^{*(1)}_{\mathbf{p},m}$ is the desired value of $\zeta^{*(1)}_{\mathbf{p},m}$ in (39) which is depicted by the green line in Fig. 1. The procedures to solve (40), (41), and (42) are similar to this capped water-filling procedure.

IV. Numerical Results

The optimal objective of the problem in (38) yielded by Algorithm 1 provides a guarantee on the detection performance of the BIoT in the presence of attacks. To corroborate this performance guarantee, we numerically investigate the optimal solution to the problem in (11) by using the projected gradient descent (PGD) method with multiple initial points [50]. In all the following simulation results, we consider the scenario where $\mathcal{K}=\{0,1\}$, $N=4$, $\mathbf{p}^H=[0.1,0.9]$, $\mathbf{q}^H=[0.9,0.1]$, and $L=4$.

As the probability α that an IoT device has been hacked into varies from 0 to 1, Fig. 2 depicts the minimum KLD obtained by using the PGD with multiple initial points and the performance guarantee yielded by Algorithm 1 for the case where $L_0=3$, $L_A=3$, and $P_s=0.0118$, which are marked with '+'s and pentagrams, respectively. It is seen from Fig. 2 that the performance guarantee yielded by Algorithm 1 provides a lower bound on the minimum KLD obtained by the PGD, which provides valuable insights into the worst detection performance of the BIoT in adversarial environments.

Next, we investigate how the value of P_s impacts the minimum KLD and the performance guarantee obtained from Algorithm 1. As α varies from 0 to 1, Fig. 3 depicts the minimum KLD obtained by using the PGD with multiple initial points and the performance guarantee yielded by Algorithm 1 with different P_s for the case where $L_0=2$ and $L_A=2$. The blue curve marked with 'x's and the orange curve marked with triangles illustrate the minimum KLD obtained by using the PGD and the performance guarantee obtained from Algorithm 1 when $P_s=0.0027$. The yellow curve marked with squares and the purple curve marked with circles illustrate the minimum KLD obtained by using the PGD and the performance guarantee obtained from Algorithm 1 when $P_s=0.0118$. The green curve marked with '+'s and the cyan curve marked with pentagrams illustrate the minimum KLD obtained by using the PGD and the performance guarantee obtained from Algorithm 1 when $P_s=0.1278$. It is seen from Fig. 3 that the performance guarantees yielded by Algorithm 1 still provide lower bounds on the minimum KLDs obtained by

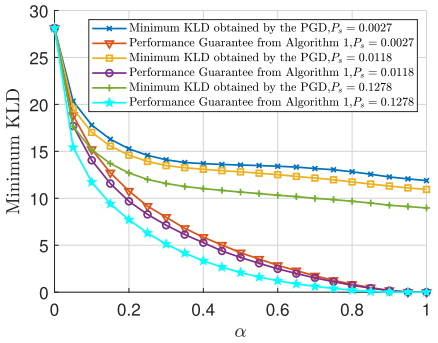

Fig. 3. Comparison between the minimum KLD and its guarantee for different P_s.

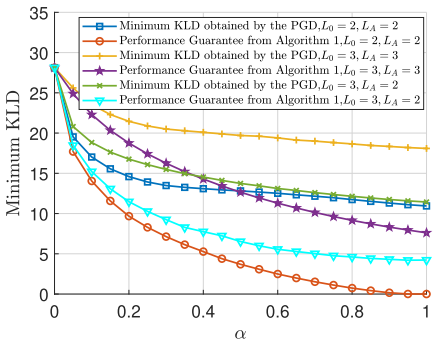

Fig. 4. Comparison between the minimum KLD and its guarantee for different L_0 and L_A.

the PGD for different P_s. In addition, Fig. 3 shows that as P_s increases, the minimum KLD and the performance guarantee both decrease. This is because as P_s increases, the probability of successfully launching a DSA becomes greater, and hence, the detection performance of the BIoT is degraded to a larger extent.

Fig. 4 depicts the minimum KLD obtained by using the PGD with multiple initial points and the performance guarantee yielded by Algorithm 1 with different L_0 and L_A for the case where $P_s = 0.0118$. The blue curve marked with squares and the orange curve marked with circles illustrate the minimum KLD obtained by using the PGD and the performance guarantee obtained from Algorithm 1 when $L_0 = 2$ and $L_A = 2$. The yellow curve marked with '+'s and the purple curve marked with pentagrams illustrate the minimum KLD obtained by using the PGD and the performance guarantee obtained from Algorithm 1 when $L_0 = 3$ and $L_A = 3$. The green curve marked with 'x's and the cyan curve marked with triangles illustrate the minimum KLD obtained by using the PGD and the performance guarantee obtained from Algorithm 1 when $L_0 = 3$ and $L_A = 2$. Fig. 4 also shows that the performance guarantee yielded by Algorithm 1 provides valid lower bounds on the minimum KLD obtained by the PGD. In addition, Fig. 4 shows that as L_0 and L_A increase, the minimum KLD and the performance guarantee both increase. This is because as L_0 and L_A increase, the numbers of counterfeit blocks in the counterfeit branch and the authentic branch both decrease. Therefore, the detection performance of the BIoT becomes better.

V. Conclusion

In this paper, we considered distributed detection over a BIoT in the presence of attacks which jointly exploit the vulnerability of IoT devices and that of the blockchain employed in the BIoT. The pursuit of a detection performance guarantee for the BIoT has been cast as a relaxed convex optimization problem, and the analytic expression for the solution to the relaxed convex optimization problem has been derived. Moreover, based on a capped water-filling method, we have developed a coordinate descent algorithm with guaranteed convergence to solve the relaxed convex optimization problem.

Appendix A
Proof of Lemma 1

Let $\tilde{\boldsymbol{\theta}}$ denote a vector stacking $\{\tilde{P}^{(1)}_{i,m}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Gamma}$, $\{\tilde{P}^{(2)}_{i,m}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Lambda}$, $\{\tilde{Q}^{(1)}_{i,m}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Delta}$, and $\{\tilde{Q}^{(2)}_{i,m}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Phi}$. Similarly, we define $\check{\boldsymbol{\theta}}$ as a vector stacking $\{\check{P}^{(1)}_{i,m}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Gamma}$, $\{\check{P}^{(2)}_{i,m}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Lambda}$, $\{\check{Q}^{(1)}_{i,m}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Delta}$, and $\{\check{Q}^{(2)}_{i,m}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Phi}$, and $\boldsymbol{\theta} \triangleq \mu\tilde{\boldsymbol{\theta}} + (1-\mu)\check{\boldsymbol{\theta}}$ for some $\mu \in [0,1]$. From (31), we can obtain

$$
\begin{aligned}
&D_1(\boldsymbol{\theta}) \\
&= \sum_{i\in\mathcal{I}} \Bigg\{ \Theta_i + \sum_{m\in\mathcal{M}_\Delta} \Delta_{i,m}\left[\mu\tilde{Q}^{(1)}_{i,m} + (1-\mu)\check{Q}^{(1)}_{i,m}\right] \\
&\quad + \sum_{m\in\mathcal{M}_\Phi} \Phi_{i,m}\left[\mu\tilde{Q}^{(2)}_{i,m} + (1-\mu)\check{Q}^{(2)}_{i,m}\right] \Bigg\} \\
&\quad \times \ln \frac{\left\{ \begin{array}{l} \Theta_i + \sum\limits_{m\in\mathcal{M}_\Delta} \Delta_{i,m}\left[\mu\tilde{Q}^{(1)}_{i,m} + (1-\mu)\check{Q}^{(1)}_{i,m}\right] \\ + \sum\limits_{m\in\mathcal{M}_\Phi} \Phi_{i,m}\left[\mu\tilde{Q}^{(2)}_{i,m} + (1-\mu)\check{Q}^{(2)}_{i,m}\right] \end{array} \right\}}{\left\{ \begin{array}{l} \Psi_i + \sum\limits_{m\in\mathcal{M}_\Gamma} \Gamma_{i,m}\left[\mu\tilde{P}^{(1)}_{i,m} + (1-\mu)\check{P}^{(1)}_{i,m}\right] \\ + \sum\limits_{m\in\mathcal{M}_\Lambda} \Lambda_{i,m}\left[\mu\tilde{P}^{(2)}_{i,m} + (1-\mu)\check{P}^{(2)}_{i,m}\right] \end{array} \right\}} \\
&\le \sum_{i\in\mathcal{I}} \mu\left(\Theta_i + \sum_{m\in\mathcal{M}_\Delta} \Delta_{i,m}\tilde{Q}^{(1)}_{i,m} + \sum_{m\in\mathcal{M}_\Phi} \Phi_{i,m}\tilde{Q}^{(2)}_{i,m}\right) \\
&\quad \times \ln \frac{\mu\left(\Theta_i + \sum\limits_{m\in\mathcal{M}_\Delta} \Delta_{i,m}\tilde{Q}^{(1)}_{i,m} + \sum\limits_{m\in\mathcal{M}_\Phi} \Phi_{i,m}\tilde{Q}^{(2)}_{i,m}\right)}{\mu\left(\Psi_i + \sum\limits_{m\in\mathcal{M}_\Gamma} \Gamma_{i,m}\tilde{P}^{(1)}_{i,m} + \sum\limits_{m\in\mathcal{M}_\Lambda} \Lambda_{i,m}\tilde{P}^{(2)}_{i,m}\right)} \\
&\quad + \sum_{i\in\mathcal{I}} (1-\mu)\left(\Theta_i + \sum_{m\in\mathcal{M}_\Delta} \Delta_{i,m}\check{Q}^{(1)}_{i,m} + \sum_{m\in\mathcal{M}_\Phi} \Phi_{i,m}\check{Q}^{(2)}_{i,m}\right) \\
&\quad \times \ln \frac{(1-\mu)\left(\Theta_i + \sum\limits_{m\in\mathcal{M}_\Delta} \Delta_{i,m}\check{Q}^{(1)}_{i,m} + \sum\limits_{m\in\mathcal{M}_\Phi} \Phi_{i,m}\check{Q}^{(2)}_{i,m}\right)}{(1-\mu)\left(\Psi_i + \sum\limits_{m\in\mathcal{M}_\Gamma} \Gamma_{i,m}\check{P}^{(1)}_{i,m} + \sum\limits_{m\in\mathcal{M}_\Lambda} \Lambda_{i,m}\check{P}^{(2)}_{i,m}\right)} \qquad (45) \\
&= \mu D_1\left(\tilde{\boldsymbol{\theta}}\right) + (1-\mu) D_1\left(\check{\boldsymbol{\theta}}\right), \qquad (46)
\end{aligned}
$$

where the inequality in (45) comes from the fact that for any nonnegative $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$, we have $\sum_{i=1}^n a_i \log \frac{a_i}{b_i} \ge (\sum_{i=1}^n a_i) \log \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}$. Therefore, $D_1(\boldsymbol{\theta})$ is a

convex function with respect to $\boldsymbol{\theta}$. Moreover, since the inequality constraint functions in (38b)–(38e) are convex and the equality constraint functions in (34)–(37) are affine functions of $\boldsymbol{\theta}$, the optimization problem in (38) is a convex optimization problem.

APPENDIX B
PROOF OF THEOREM 1

Define $\mathbf{P}_m^{*(1)}$, $\forall m \in \mathcal{M}_\Gamma$, $\mathbf{P}_m^{*(2)}$, $\forall m \in \mathcal{M}_\Lambda$, $\mathbf{Q}_m^{*(1)}$, $\forall m \in \mathcal{M}_\Delta$, and $\mathbf{Q}_m^{*(2)}$, $\forall m \in \mathcal{M}_\Phi$ as vectors stacking $\{P_{i,m}^{*(1)}\}_{i\in\mathcal{I}}$, $\{P_{i,m}^{*(2)}\}_{i\in\mathcal{I}}$, $\{Q_{i,m}^{*(1)}\}_{i\in\mathcal{I}}$, and $\{Q_{i,m}^{*(2)}\}_{i\in\mathcal{I}}$, respectively. Let $D_1^*(\mathbf{Q}_m^{(1)}) \triangleq D_1(\boldsymbol{\theta})|\{\mathbf{Q}_{m'}^{(1)} = \mathbf{Q}_{m'}^{*(1)}\}_{m'\in\mathcal{M}_\Delta\setminus\{m\}}, \{\mathbf{P}_{m'}^{(1)} = \mathbf{P}_{m'}^{*(1)}\}_{m'\in\mathcal{M}_\Gamma}, \{\mathbf{P}_{m'}^{(2)} = \mathbf{P}_{m'}^{*(2)}\}_{m'\in\mathcal{M}_\Lambda}, \{\mathbf{Q}_{m'}^{(2)} = \mathbf{Q}_{m'}^{*(2)}\}_{m'\in\mathcal{M}_\Phi}$ which indicates that we only consider $\mathbf{Q}_m^{(1)}$ in $\boldsymbol{\theta}$ as the variable of $D_1^*(\mathbf{Q}_m^{(1)})$, and the other parameters in $\boldsymbol{\theta}$ are set to be the globally optimal solutions to the optimization problem in (38). Similarly, we define $D_2^*(\mathbf{Q}_m^{(2)}) \triangleq D_1(\boldsymbol{\theta})|\{\mathbf{Q}_{m'}^{(2)} = \mathbf{Q}_{m'}^{*(2)}\}_{m'\in\mathcal{M}_\Phi\setminus\{m\}}, \{\mathbf{P}_{m'}^{(1)} = \mathbf{P}_{m'}^{*(1)}\}_{m'\in\mathcal{M}_\Gamma}, \{\mathbf{P}_{m'}^{(2)} = \mathbf{P}_{m'}^{*(2)}\}_{m'\in\mathcal{M}_\Lambda}, \{\mathbf{Q}_{m'}^{(1)} = \mathbf{Q}_{m'}^{*(1)}\}_{m'\in\mathcal{M}_\Delta}$, $D_3^*(\mathbf{P}_m^{(1)}) \triangleq D_1(\boldsymbol{\theta})|\{\mathbf{P}_{m'}^{(1)} = \mathbf{P}_{m'}^{*(1)}\}_{m'\in\mathcal{M}_\Gamma\setminus\{m\}}, \{\mathbf{P}_{m'}^{(2)} = \mathbf{P}_{m'}^{*(2)}\}_{m'\in\mathcal{M}_\Lambda}, \{\mathbf{Q}_{m'}^{(1)} = \mathbf{Q}_{m'}^{*(1)}\}_{m'\in\mathcal{M}_\Delta}, \{\mathbf{Q}_{m'}^{(2)} = \mathbf{Q}_{m'}^{*(2)}\}_{m'\in\mathcal{M}_\Phi}$, and $D_4^*(\mathbf{P}_m^{(2)}) \triangleq D_1(\boldsymbol{\theta})|\{\mathbf{P}_{m'}^{(2)} = \mathbf{P}_{m'}^{*(2)}\}_{m'\in\mathcal{M}_\Lambda\setminus\{m\}}, \{\mathbf{P}_{m'}^{(1)} = \mathbf{P}_{m'}^{*(1)}\}_{m'\in\mathcal{M}_\Gamma}, \{\mathbf{Q}_{m'}^{(1)} = \mathbf{Q}_{m'}^{*(1)}\}_{m'\in\mathcal{M}_\Delta}, \{\mathbf{Q}_{m'}^{(2)} = \mathbf{Q}_{m'}^{*(2)}\}_{m'\in\mathcal{M}_\Phi}$.

Since $\{P_{i,m}^{*(1)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Gamma}$, $\{P_{i,m}^{*(2)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Lambda}$, $\{Q_{i,m}^{*(1)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Delta}$, and $\{Q_{i,m}^{*(2)}\}_{i\in\mathcal{I},m\in\mathcal{M}_\Phi}$ are the globally optimal solutions to the optimization problem in (38), we know that

$$\mathbf{Q}_m^{*(1)} = \arg\min_{\mathbf{Q}_m^{(1)}} D_1^*\left(\mathbf{Q}_m^{(1)}\right), \forall m \in \mathcal{M}_\Delta \quad (47)$$
$$\text{s.t. constraints (36), (38d).}$$

$$\mathbf{Q}_m^{*(2)} = \arg\min_{\mathbf{Q}_m^{(2)}} D_2^*\left(\mathbf{Q}_m^{(2)}\right), \forall m \in \mathcal{M}_\Phi \quad (48)$$
$$\text{s.t. constraints (37), (38e).}$$

$$\mathbf{P}_m^{*(1)} = \arg\min_{\mathbf{P}_m^{(1)}} D_3^*\left(\mathbf{P}_m^{(1)}\right), \forall m \in \mathcal{M}_\Gamma \quad (49)$$
$$\text{s.t. constraints (34), (38b).}$$

$$\mathbf{P}_m^{*(2)} = \arg\min_{\mathbf{P}_m^{(2)}} D_4^*\left(\mathbf{P}_m^{(2)}\right), \forall m \in \mathcal{M}_\Lambda \quad (50)$$
$$\text{s.t. constraints (35), (38c).}$$

From (31), we can obtain the following property of the optimal solution.

Claim 1: The globally optimal solution to (38) cannot give rise to that $A_i^* > 0$ and $B_i^* = 0$ for any $i \in \mathcal{I}$.

Proof of Claim 1: We prove this claim by contradiction. Suppose that the globally optimal solution to (38), denoted by $\boldsymbol{\theta}^*$, gives rise to that $A_{i'}^* > 0$ and $B_{i'}^* = 0$ for some $i' \in \mathcal{I}$. We know from (31) and the definitions of A_i^* and B_i^* in Theorem 1 that $D_1(\boldsymbol{\theta}^*) = \infty$. From (12), (13), (15), (17), (19), and the definition of B_i^*, we know that $B_{i'}^* = 0$ only if $\alpha = 1$ and $\Gamma_{i',2^N-1}P_{i',2^N-1}^{*(1)} = \Lambda_{i',2^N-1}P_{i',2^N-1}^{*(2)} = 0$. Note that if $\alpha = 1$ and $P_s > 0$, then $\Gamma_{i',2^N-1} > 0$, and hence $P_{i',2^N-1}^{*(1)} = 0$. If $\alpha = 1$ and $P_s = 0$, then $\Lambda_{i',2^N-1} > 0$, and hence $P_{i',2^N-1}^{*(2)} = 0$. We first arbitrarily pick an $i^\dagger \in \{i \in \mathcal{I}|P_{i,2^N-1}^{*(1)} > 0\}$ if $P_s > 0$ (or $i^\dagger \in \{i \in \mathcal{I}|P_{i,2^N-1}^{*(2)} > 0\}$ if $P_s = 0$). It's worth mentioning that when $\alpha = 1$ we can always find such an $i^\dagger$ due to the constraint in (34) that $\sum_{i\in\mathcal{I}}\Gamma_{i,2^N-1}P_{i,2^N-1}^{*(1)} = P_s > 0$ if $P_s > 0$ (or (35) that $\sum_{i\in\mathcal{I}}\Lambda_{i,2^N-1}P_{i,2^N-1}^{*(2)} = 1 - P_s > 0$ if $P_s = 0$). We increase the value of $P_{i',2^N-1}^{*(1)}$ by some constant ϵ where $0 < \epsilon < \min\{1, (\Gamma_{i^\dagger,2^N-1}P_{i^\dagger,2^N-1}^{*(1)})/\Gamma_{i',2^N-1}\}$ if $P_s > 0$ (or increase the value of $P_{i',2^N-1}^{*(2)}$ by some constant ϵ where $0 < \epsilon < \min\{1, (\Lambda_{i^\dagger,2^N-1}P_{i^\dagger,2^N-1}^{*(2)})/\Lambda_{i',2^N-1}\}$ if $P_s = 0$), and decrease the value of $P_{i^\dagger,2^N-1}^{*(1)}$ by the constant $\epsilon\Gamma_{i',2^N-1}/\Gamma_{i^\dagger,2^N-1}$ if $P_s > 0$ (or decrease the value of $P_{i^\dagger,2^N-1}^{*(2)}$ by the constant $\epsilon\Lambda_{i',2^N-1}/\Lambda_{i^\dagger,2^N-1}$ if $P_s = 0$) so that the value of $\sum_{i\in\mathcal{I}}\Gamma_{i,2^N-1}P_{i,2^N-1}^{*(1)}$ (or $\sum_{i\in\mathcal{I}}\Lambda_{i,2^N-1}P_{i,2^N-1}^{*(2)}$) doesn't change and hence the constraint in (34) (or (35)) still holds. By doing so, the adjusted B_i^*, denoted by $\hat{B}_{i'}^*$, can be written as $\hat{B}_{i'}^* = \epsilon\Gamma_{i',2^N-1} > 0$ if $P_s > 0$ (or $\hat{B}_{i'}^* = \epsilon\Lambda_{i',2^N-1} > 0$ if $P_s = 0$). Moreover, since $P_{i^\dagger,2^N-1}^{*(1)} - (\epsilon\Gamma_{i',2^N-1}/\Gamma_{i^\dagger,2^N-1}) > 0$ (or $P_{i^\dagger,2^N-1}^{*(2)} - (\epsilon\Lambda_{i',2^N-1}/\Lambda_{i^\dagger,2^N-1}) > 0$), the adjusted $B_{i^\dagger}^*$, i.e., $\hat{B}_{i^\dagger}^*$, satisfies $\hat{B}_{i^\dagger}^* > 0$.

By following the same procedure for all $i' \in \mathcal{I}$ such that $A_{i'}^* > 0$ and $B_{i'}^* = 0$, we obtain a new solution based on the globally optimal solution, denoted by $\hat{\boldsymbol{\theta}}^*$, which gives rise to $\hat{A}_i^* > 0$ and $\hat{B}_i^* > 0$, $\forall i \in \mathcal{I}$. We know from (31) that this new solution $\hat{\boldsymbol{\theta}}^*$ leads to a finite objective value $D_1(\hat{\boldsymbol{\theta}}^*)$, and hence $D_1(\hat{\boldsymbol{\theta}}^*) < D_1(\boldsymbol{\theta}^*)$. This contradicts the optimality of the globally optimal solution $\boldsymbol{\theta}^*$ to (38). Therefore, the globally optimal solution to (38) cannot give rise to that $A_i^* > 0$ and $B_i^* = 0$ for any $i \in \mathcal{I}$. ■

We can know from Claim 1 that for the globally optimal solution to (38), if $B_i^* = 0$, then $A_i^* = 0$, and if $A_i^* > 0$, then $B_i^* > 0$. Next, we will solve (47)–(50). First, we consider the optimization problem in (47). Since the problem in (38) is a convex optimization problem, the problem in (47) is also a convex optimization problem. Noting that the inequality constraints in (38d) of the problem in (47) are affine, and the feasible set of (47) is not empty, the weaker Slater's condition for (47) holds, and hence, the Karush-Kuhn-Tucker (KKT) conditions are the necessary and sufficient conditions for the optimality of the solution to the problem in (47) [51]. Hence, we can pursue the optimal solution to the optimization problem in (47) by solving the KKT conditions for the problem in (47), which can be written as $\forall m \in \mathcal{M}_\Delta$,

$$\nabla_{\mathbf{Q}_m^{(1)}} F_1(\mathbf{Q}_m^{(1)}, \boldsymbol{\mu}_{1,m}, \boldsymbol{\mu}_{2,m}, \mu_{3,m}) = 0, \quad (51a)$$
$$\mu_{1,m}^{(i)} f_{Q_m^{(1)}}^{(i)} = 0, \mu_{2,m}^{(i)} g_{Q_m^{(1)}}^{(i)} = 0, \quad \forall i \in \mathcal{I}, \quad (51b)$$

$$f^{(i)}_{Q^{(1)}_m}, g^{(i)}_{Q^{(1)}_m} \le 0, \ \forall i \in \mathcal{I}, \tag{51c}$$
$$h_{\mathbf{Q}^{(1)}_m} = 0, \tag{51d}$$
$$\mu^{(i)}_{1,m}, \mu^{(i)}_{2,m} \ge 0, \ \forall i \in \mathcal{I}, \tag{51e}$$

where $f^{(i)}_{Q^{(1)}_m} \triangleq -Q^{(1)}_{i,m}$, $g^{(i)}_{Q^{(1)}_m} \triangleq Q^{(1)}_{i,m} - 1$, $h_{\mathbf{Q}^{(1)}_m} \triangleq \sum_{i\in\mathcal{I}} \Delta_{i,m} Q^{(1)}_{i,m} - P_s(1-\alpha)^{N-n_m}\alpha^{n_m}$, $\boldsymbol{\mu}_{1,m}$ is a vector stacking $\{\mu^{(i)}_{1,m}\}_{i\in\mathcal{I}}$, $\boldsymbol{\mu}_{2,m}$ is a vector stacking $\{\mu^{(i)}_{2,m}\}_{i\in\mathcal{I}}$, and

$$F_1(\mathbf{Q}^{(1)}_m, \boldsymbol{\mu}_{1,m}, \boldsymbol{\mu}_{2,m}, \mu_{3,m}) \triangleq D^*_1\left(\mathbf{Q}^{(1)}_m\right) + \sum_{i\in\mathcal{I}} \mu^{(i)}_{1,m} f^{(i)}_{Q^{(1)}_m} + \sum_{i\in\mathcal{I}} \mu^{(i)}_{2,m} g^{(i)}_{Q^{(1)}_m} + \mu_{3,m} h_{\mathbf{Q}^{(1)}_m}. \tag{52}$$

Since for any $m \in \mathcal{M}_\Delta$ and any $i \in \mathcal{I}$, $f^{(i)}_{Q^{(1)}_m} = -Q^{(1)}_{i,m} \le 0$, $g^{(i)}_{Q^{(1)}_m} = Q^{(1)}_{i,m} - 1 \le 0$, and $f^{(i)}_{Q^{(1)}_m} + g^{(i)}_{Q^{(1)}_m} = -1$, we know from (51b) that,

$$\mu^{(i)}_{1,m}\mu^{(i)}_{2,m} = 0, \forall i \in \mathcal{I}, \forall m \in \mathcal{M}_\Delta. \tag{53}$$

Let $\mu^{*(i)}_{1,m}$, $\mu^{*(i)}_{2,m}$, and $\mu^*_{3,m}$ be the optimal Lagrange multipliers. For any $m \in \mathcal{M}_\Delta$ and any $i \in \mathcal{I}$, we know from (51b), (51c), (51e), and (53) that the optimal solution $Q^{*(1)}_{i,m}$ to (51) satisifies

$$\begin{cases} Q^{*(1)}_{i,m} = 0, \ \text{if } \mu^{*(i)}_{1,m} > 0 \text{ and } \mu^{*(i)}_{2,m} = 0, & (54)\\ Q^{*(1)}_{i,m} = 1, \ \text{if } \mu^{*(i)}_{1,m} = 0 \text{ and } \mu^{*(i)}_{2,m} > 0, & (55)\\ Q^{*(1)}_{i,m} \in [0,1], \ \text{if } \mu^{*(i)}_{1,m} = \mu^{*(i)}_{2,m} = 0. & (56) \end{cases}$$

For any $i \in \{i \in \mathcal{I} | B^*_i = 0\}$, we know from Claim 1 that if $B^*_i = 0$, then $A^*_i = 0$. Moreover, we know from the definition of A^*_i in Theorem 1 that $A^*_i = 0$ implies that $Q^{*(1)}_{i,m} = 0$ for all $m \in \mathcal{M}_\Delta$.

From (31) and (51a), by taking the partial derivative of (52) with respect to $Q^{(1)}_{i,m}$ for any $m \in \mathcal{M}_\Delta$ and any $i \in \{i \in \mathcal{I} | B^*_i > 0\}$ and setting it to zero, we can obtain that $\forall m \in \mathcal{M}_\Delta$ and $\forall i \in \{i \in \mathcal{I} | B^*_i > 0\}$, the optimal solution $Q^{*(1)}_{i,m}$ satisfies

$$\ln \frac{\Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}} \Delta_{i,m'} Q^{*(1)}_{i,m'} + \Delta_{i,m} Q^{*(1)}_{i,m} + \sum_{m'\in\mathcal{M}_\Phi} \Phi_{i,m'} Q^{*(2)}_{i,m'}}{\Psi_i + \sum_{m'\in\mathcal{M}_\Gamma} \Gamma_{i,m'} P^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Lambda} \Lambda_{i,m'} P^{*(2)}_{i,m'}} \times \Delta_{i,m} + \Delta_{i,m} - \mu^{*(i)}_{1,m} + \mu^{*(i)}_{2,m} + \mu^*_{3,m}\Delta_{i,m} = 0, \tag{57}$$

which is equivalent to

$$\ln \frac{\Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}} \Delta_{i,m'} Q^{*(1)}_{i,m'} + \Delta_{i,m} Q^{*(1)}_{i,m} + \sum_{m'\in\mathcal{M}_\Phi} \Phi_{i,m'} Q^{*(2)}_{i,m'}}{\Psi_i + \sum_{m'\in\mathcal{M}_\Gamma} \Gamma_{i,m'} P^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Lambda} \Lambda_{i,m'} P^{*(2)}_{i,m'}} = \frac{-\Delta_{i,m} + \mu^{*(i)}_{1,m} - \mu^{*(i)}_{2,m} - \mu^*_{3,m}\Delta_{i,m}}{\Delta_{i,m}}. \tag{58}$$

Next, we will develop $Q^{*(1)}_{i,m}$ for any $m \in \mathcal{M}_\Delta$, any $i \in \{i \in \mathcal{I} | B^*_i > 0\}$, and different value of $\mu^*_{3,m}$ by using (58). We first define

$$\zeta^{*(1)}_{\mathbf{q},m} \triangleq \exp\left\{-\mu^*_{3,m} - 1\right\}, \tag{59}$$

which only depends on $\mu^*_{3,m}$.

Claim 2: For any $m \in \mathcal{M}_\Delta$ and any $i \in \{i \in \mathcal{I} | B^*_i > 0\}$, if $\zeta^{*(1)}_{\mathbf{q},m} B^*_i \le \Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}} \Delta_{i,m'} Q^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Phi} \Phi_{i,m'} Q^{*(2)}_{i,m'}$, then in order to make (58) hold, $Q^{*(1)}_{i,m} = 0$.

Proof of Claim 2: For any $m \in \mathcal{M}_\Delta$ and any $i \in \{i \in \mathcal{I} | B^*_i > 0\}$, if $\zeta^{*(1)}_{\mathbf{q},m} B^*_i \le \Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}} \Delta_{i,m'} Q^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Phi} \Phi_{i,m'} Q^{*(2)}_{i,m'}$, then by using the definition of B^*_i in Theorem 1, we can obtain

$$\frac{\Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}} \Delta_{i,m'} Q^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Phi} \Phi_{i,m'} Q^{*(2)}_{i,m'}}{\Psi_i + \sum_{m'\in\mathcal{M}_\Gamma} \Gamma_{i,m'} P^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Lambda} \Lambda_{i,m'} P^{*(2)}_{i,m'}} \ge \zeta^{*(1)}_{\mathbf{q},m}. \tag{60}$$

Noting that $\Delta_{i,m} > 0$ and $Q^{*(1)}_{i,m} \ge 0$, we can obtain from (60) that

$$\begin{aligned} &\frac{\Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}} \Delta_{i,m'} Q^{*(1)}_{i,m'} + \Delta_{i,m} Q^{*(1)}_{i,m} + \sum_{m'\in\mathcal{M}_\Phi} \Phi_{i,m'} Q^{*(2)}_{i,m'}}{\Psi_i + \sum_{m'\in\mathcal{M}_\Gamma} \Gamma_{i,m'} P^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Lambda} \Lambda_{i,m'} P^{*(2)}_{i,m'}} \\ &\ge \frac{\Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}} \Delta_{i,m'} Q^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Phi} \Phi_{i,m'} Q^{*(2)}_{i,m'}}{\Psi_i + \sum_{m'\in\mathcal{M}_\Gamma} \Gamma_{i,m'} P^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Lambda} \Lambda_{i,m'} P^{*(2)}_{i,m'}} \\ &\ge \zeta^{*(1)}_{\mathbf{q},m}. \end{aligned} \tag{61}$$

Comparing (61) with (58), we can obtain that in order to make (58) hold,

$$\frac{\mu^{*(i)}_{1,m} - \mu^{*(i)}_{2,m}}{\Delta_{i,m}} \ge 0, \tag{62}$$

which yields $\mu^{*(i)}_{1,m} \ge 0$ and $\mu^{*(i)}_{2,m} = 0$ due to the fact that there are only three combinations of the signs of $\mu^{*(i)}_{1,m}$ and $\mu^{*(i)}_{2,m}$ which are illustrated in (54), (55), and (56). If $\mu^{*(i)}_{1,m} > 0$ and $\mu^{*(i)}_{2,m} = 0$, then from (54), we can obtain $Q^{*(1)}_{i,m} = 0$. If $\mu^{*(i)}_{1,m} = 0$ and $\mu^{*(i)}_{2,m} = 0$, then (58) is equivalent to

$$\frac{\Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}} \Delta_{i,m'} Q^{*(1)}_{i,m'} + \Delta_{i,m} Q^{*(1)}_{i,m} + \sum_{m'\in\mathcal{M}_\Phi} \Phi_{i,m'} Q^{*(2)}_{i,m'}}{\Psi_i + \sum_{m'\in\mathcal{M}_\Gamma} \Gamma_{i,m'} P^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Lambda} \Lambda_{i,m'} P^{*(2)}_{i,m'}} = \zeta^{*(1)}_{\mathbf{q},m}. \tag{63}$$

From (61) and (63), we have

$$\begin{aligned} &\frac{\Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}} \Delta_{i,m'} Q^{*(1)}_{i,m'} + \Delta_{i,m} Q^{*(1)}_{i,m} + \sum_{m'\in\mathcal{M}_\Phi} \Phi_{i,m'} Q^{*(2)}_{i,m'}}{\Psi_i + \sum_{m'\in\mathcal{M}_\Gamma} \Gamma_{i,m'} P^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Lambda} \Lambda_{i,m'} P^{*(2)}_{i,m'}} \\ &= \frac{\Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}} \Delta_{i,m'} Q^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Phi} \Phi_{i,m'} Q^{*(2)}_{i,m'}}{\Psi_i + \sum_{m'\in\mathcal{M}_\Gamma} \Gamma_{i,m'} P^{*(1)}_{i,m'} + \sum_{m'\in\mathcal{M}_\Lambda} \Lambda_{i,m'} P^{*(2)}_{i,m'}}, \end{aligned} \tag{64}$$

which implies that $Q_{i,m}^{*(1)} = 0$. The proof of Claim 2 is completed. ■

Claim 3: For any $m \in \mathcal{M}_\Delta$ *and any* $i \in \{i \in \mathcal{I}|B_i^* > 0\}$, *if* $\zeta_{\boldsymbol{q},m}^{*(1)} B_i^* \geq \Theta_i + \sum_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}$, *then in order to make (58) hold,* $Q_{i,m}^{*(1)} = 1$.

Proof of Claim 3: For any $m \in \mathcal{M}_\Delta$ and any $i \in \{i \in \mathcal{I}|B_i^* > 0\}$, if $\zeta_{\mathbf{q},m}^{*(1)} B_i^* \geq \Theta_i + \sum_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}$, then by using the definition of B_i^* in Theorem 1, we can obtain

$$\frac{\Theta_i + \sum\limits_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum\limits_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}}{\Psi_i + \sum\limits_{m' \in \mathcal{M}_\Gamma} \Gamma_{i,m'} P_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Lambda} \Lambda_{i,m'} P_{i,m'}^{*(2)}} \leq \zeta_{\mathbf{q},m}^{*(1)}. \tag{65}$$

Noting that $\Delta_{i,m} > 0$ and $Q_{i,m}^{*(1)} \leq 1$, we can obtain from (65) that

$$\begin{aligned} &\frac{\Theta_i + \sum\limits_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} Q_{i,m}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}}{\Psi_i + \sum\limits_{m' \in \mathcal{M}_\Gamma} \Gamma_{i,m'} P_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Lambda} \Lambda_{i,m'} P_{i,m'}^{*(2)}} \\ &\leq \frac{\Theta_i + \sum\limits_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum\limits_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}}{\Psi_i + \sum\limits_{m' \in \mathcal{M}_\Gamma} \Gamma_{i,m'} P_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Lambda} \Lambda_{i,m'} P_{i,m'}^{*(2)}} \\ &\leq \zeta_{\mathbf{q},m}^{*(1)}. \end{aligned} \tag{66}$$

Comparing (66) with (58), we can obtain that in order to make (58) hold,

$$\frac{\mu_{2,m}^{*(i)} - \mu_{1,m}^{*(i)}}{\Delta_{i,m}} \geq 0, \tag{67}$$

which yields $\mu_{1,m}^{*(i)} = 0$ and $\mu_{2,m}^{*(i)} \geq 0$ due to the fact that there are only three combinations of the signs of $\mu_{1,m}^{*(i)}$ and $\mu_{2,m}^{*(i)}$ which are illustrated in (54), (55), and (56). If $\mu_{1,m}^{*(i)} = 0$ and $\mu_{2,m}^{*(i)} > 0$, then from (54), we can obtain $Q_{i,m}^{*(1)} = 1$. If $\mu_{1,m}^{*(i)} = 0$ and $\mu_{2,m}^{*(i)} = 0$, then (58) is equivalent to

$$\frac{\Theta_i + \sum\limits_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} Q_{i,m}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}}{\Psi_i + \sum\limits_{m' \in \mathcal{M}_\Gamma} \Gamma_{i,m'} P_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Lambda} \Lambda_{i,m'} P_{i,m'}^{*(2)}} = \zeta_{\mathbf{q},m}^{*(1)}. \tag{68}$$

From (66) and (68), we have

$$\begin{aligned} &\frac{\Theta_i + \sum\limits_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} Q_{i,m}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}}{\Psi_i + \sum\limits_{m' \in \mathcal{M}_\Gamma} \Gamma_{i,m'} P_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Lambda} \Lambda_{i,m'} P_{i,m'}^{*(2)}} \\ &= \frac{\Theta_i + \sum\limits_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum\limits_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}}{\Psi_i + \sum\limits_{m' \in \mathcal{M}_\Gamma} \Gamma_{i,m'} P_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Lambda} \Lambda_{i,m'} P_{i,m'}^{*(2)}}, \end{aligned} \tag{69}$$

which implies that $Q_{i,m}^{*(1)} = 1$. The proof of Claim 3 is completed. ■

Claim 4: For any $m \in \mathcal{M}_\Delta$ *and any* $i \in \{i \in \mathcal{I}|B_i^* > 0\}$, *if* $\zeta_{\boldsymbol{q},m}^{*(1)} B_i^* \in (\Theta_i + \sum_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \sum_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}, \Theta_i + \sum_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)})$, *then in order to make (58) hold,* $Q_{i,m}^{*(1)} = (\zeta_{\boldsymbol{q},m}^{*(1)} B_i^* - \Theta_i - \sum_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} - \sum_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)})/\Delta_{i,m}$.

Proof of Claim 4: First, we show that for any $m \in \mathcal{M}_\Delta$ and any $i \in \{i \in \mathcal{I}|B_i^* > 0\}$, in order to make (58) hold, $\mu_{1,m}^{*(i)} = \mu_{2,m}^{*(i)} = 0$ for the case that $\zeta_{\mathbf{q},m}^{*(1)} B_i^* \in (\Theta_i + \sum_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \sum_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}, \Theta_i + \sum_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)})$.

Note that there are only three combinations of the signs of $\mu_{1,m}^{*(i)}$ and $\mu_{2,m}^{*(i)}$ which are illustrated in (54), (55), and (56). If $\mu_{1,m}^{*(i)} > 0$ and $\mu_{2,m}^{*(i)} = 0$, then from (54) we know that $Q_{i,m}^{*(1)} = 0$. Hence, (58) is equivalent to

$$\ln \frac{\Theta_i + \sum\limits_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}}{\Psi_i + \sum\limits_{m' \in \mathcal{M}_\Gamma} \Gamma_{i,m'} P_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Lambda} \Lambda_{i,m'} P_{i,m'}^{*(2)}} = \frac{-\Delta_{i,m} + \mu_{1,m}^{*(i)} - \mu_{3,m}^{*} \Delta_{i,m}}{\Delta_{i,m}}, \tag{70}$$

which yields

$$\frac{\Theta_i + \sum\limits_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}}{\Psi_i + \sum\limits_{m' \in \mathcal{M}_\Gamma} \Gamma_{i,m'} P_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Lambda} \Lambda_{i,m'} P_{i,m'}^{*(2)}} > \zeta_{\mathbf{q},m}^{*(1)}. \tag{71}$$

However, for the case that $\zeta_{\mathbf{q},m}^{*(1)} B_i^* \in (\Theta_i + \sum_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \sum_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}, \Theta_i + \sum_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)})$, we have $\zeta_{\mathbf{q},m}^{*(1)} B_i^* > \Theta_i + \sum_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \sum_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}$ which implies that

$$\frac{\Theta_i + \sum\limits_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}}{\Psi_i + \sum\limits_{m' \in \mathcal{M}_\Gamma} \Gamma_{i,m'} P_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Lambda} \Lambda_{i,m'} P_{i,m'}^{*(2)}} < \zeta_{\mathbf{q},m}^{*(1)}, \tag{72}$$

which contradicts (71). Hence, $\mu_{1,m}^{*(i)} > 0$ and $\mu_{2,m}^{*(i)} = 0$ cannot make (58) hold for this case. If $\mu_{1,m}^{*(i)} = 0$ and $\mu_{2,m}^{*(i)} > 0$, then from (55) we know that $Q_{i,m}^{*(1)} = 1$. Hence, (58) is equivalent to

$$\ln \frac{\Theta_i + \sum\limits_{m' \in \mathcal{M}_\Delta \setminus \{m\}} \Delta_{i,m'} Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum\limits_{m' \in \mathcal{M}_\Phi} \Phi_{i,m'} Q_{i,m'}^{*(2)}}{\Psi_i + \sum\limits_{m' \in \mathcal{M}_\Gamma} \Gamma_{i,m'} P_{i,m'}^{*(1)} + \sum\limits_{m' \in \mathcal{M}_\Lambda} \Lambda_{i,m'} P_{i,m'}^{*(2)}} = \frac{-\Delta_{i,m} - \mu_{2,m}^{*(i)} - \mu_{3,m}^{*} \Delta_{i,m}}{\Delta_{i,m}}, \tag{73}$$

which yields

$$\frac{\Theta_i+\sum\limits_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)}+\Delta_{i,m}+\sum\limits_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)}}{\Psi_i+\sum\limits_{m'\in\mathcal{M}_\Gamma}\Gamma_{i,m'}P_{i,m'}^{*(1)}+\sum\limits_{m'\in\mathcal{M}_\Lambda}\Lambda_{i,m'}P_{i,m'}^{*(2)}}<\zeta_{\mathbf{q},m}^{*(1)}. \quad (74)$$

However, the condition $\zeta_{\mathbf{q},m}^{*(1)}B_i^* < \Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)}$ implies that

$$\frac{\Theta_i+\sum\limits_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)}+\Delta_{i,m}+\sum\limits_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)}}{\Psi_i+\sum\limits_{m'\in\mathcal{M}_\Gamma}\Gamma_{i,m'}P_{i,m'}^{*(1)}+\sum\limits_{m'\in\mathcal{M}_\Lambda}\Lambda_{i,m'}P_{i,m'}^{*(2)}}>\zeta_{\mathbf{q},m}^{*(1)}, \quad (75)$$

which contradicts (74). Hence, $\mu_{1,m}^{*(i)}=0$ and $\mu_{2,m}^{*(i)}>0$ cannot make (58) hold for this case. For the last case that $\mu_{1,m}^{*(i)}=0$ and $\mu_{2,m}^{*(i)}=0$, we know from (56) that $Q_{i,m}^{*(1)}\in[0,1]$, and (58) is equivalent to

$$\ln\frac{\Theta_i+\sum\limits_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)}+\Delta_{i,m}Q_{i,m}^{*(1)}+\sum\limits_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)}}{\Psi_i+\sum\limits_{m'\in\mathcal{M}_\Gamma}\Gamma_{i,m'}P_{i,m'}^{*(1)}+\sum\limits_{m'\in\mathcal{M}_\Lambda}\Lambda_{i,m'}P_{i,m'}^{*(2)}}=\frac{-\Delta_{i,m}-\mu_{3,m}^*\Delta_{i,m}}{\Delta_{i,m}}, \quad (76)$$

which yields

$$Q_{i,m}^{*(1)}=\frac{\begin{pmatrix}\zeta_{\mathbf{q},m}^{*(1)}B_i^*-\Theta_i-\sum_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)}\\-\sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)}\end{pmatrix}}{\Delta_{i,m}}. \quad (77)$$

Since $\zeta_{\mathbf{q},m}^{*(1)}B_i^* \in (\Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)} + \sum_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)}, \Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)})$, we can know from (77) that $Q_{i,m}^{*(1)}\in(0,1)$ which satisfies the constraint $Q_{i,m}^{*(1)}\in[0,1]$. Therefore, for any $m\in\mathcal{M}_\Delta$ and any $i\in\{i\in\mathcal{I}|B_i^*>0\}$, in order to make (58) hold, $\mu_{1,m}^{*(i)}=\mu_{2,m}^{*(i)}=0$ if $\zeta_{\mathbf{q},m}^{*(1)}B_i^*\in(\Theta_i+\sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)}+\sum_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)}, \Theta_i + \sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)} + \Delta_{i,m} + \sum_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)})$. From (58) and $\mu_{1,m}^{*(i)}=\mu_{2,m}^{*(i)}=0$, we can get (77) which completes the proof of Claim 4. ■

Note that (47) holds for any $m\in\mathcal{M}_\Delta$. Hence from Claim 1, 2, 3, and 4, the expression for $Q_{i,m}^{*(1)}$ can be summarized as that $\forall i\in\mathcal{I}$ and $\forall m\in\mathcal{M}_\Delta$,

$$Q_{i,m}^{*(1)}=\min\left\{\left[\frac{\begin{pmatrix}\zeta_{\mathbf{q},m}^{*(1)}B_i^*-\Theta_i-\sum_{m'\in\mathcal{M}_\Phi}\Phi_{i,m'}Q_{i,m'}^{*(2)}\\-\sum_{m'\in\mathcal{M}_\Delta\setminus\{m\}}\Delta_{i,m'}Q_{i,m'}^{*(1)}\end{pmatrix}}{\Delta_{i,m}}\right]^+,1\right\}. \quad (78)$$

Note that (78) only satisfies the constraint $Q_{i,m}^{*(1)}\in[0,1]$. In order to satisfy the constraint in (51d), $\zeta_{\mathbf{q},m}^{*(1)}$ should ensure $\sum_{i\in\mathcal{I}}\Delta_{i,m}Q_{i,m}^{*(1)}=P_s(1-\alpha)^{N-n_m}\alpha^{n_m}$. It's worth mentioning that we know from (78) that $\zeta_{\mathbf{q},m}^{*(1)}$ must be positive so that the constraint in (51d) can be satisfied.

The processes of solving (48), (49), and (50) are similar to the process of solving (47). Similar to (78), the solutions $Q_{i,m,}^{*(2)}$, $P_{i,m,}^{*(1)}$, and $P_{i,m,}^{*(2)}$ to the optimization problems in (48), (49), and (50) can be expressed as (42), (39), and (40), respectively, which completes the proof.

REFERENCES

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[2] 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 Things J.*, vol. 4, no. 5, pp. 1125–1142, Oct. 2017.

[3] A. Mosenia and N. K. Jha, "A comprehensive study of security of Internet-of-Things," *IEEE Trans. Emerg. Topics Comput.*, vol. 5, no. 4, pp. 586–602, Oct. 2017.

[4] X. Pan, Y. Shen, and J. Zhang, "IoUT based underwater target localization in the presence of time synchronization attacks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3958–3973, Jun. 2021.

[5] S. Marano, V. Matta, and L. Tong, "Distributed detection in the presence of Byzantine attacks," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 16–29, Jan. 2009.

[6] J. Zhang and X. Wang, "Asymptotically optimal stochastic encryption for quantized sequential detection in the presence of eavesdroppers," *IEEE Trans. Inf. Theory*, vol. 66, no. 3, pp. 1530–1548, Mar. 2020.

[7] A. Vempaty, L. Tong, and P. K. Varshney, "Distributed inference with Byzantine data: State-of-the-art review on data falsification attacks," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 65–75, Sep. 2013.

[8] J. Zhang, R. S. Blum, X. Lu, and D. Conus, "Asymptotically optimum distributed estimation in the presence of attacks," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1086–1101, Mar. 2015.

[9] V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriyeh, A. Dehghantanha, and G. Srivastava, "Federated-learning-based anomaly detection for IoT security attacks," *IEEE Internet Things J.*, vol. 9, no. 4, pp. 2545–2554, Feb. 2022.

[10] J. Zhang, X. Wang, R. S. Blum, and L. M. Kaplan, "Attack detection in sensor network target localization systems with quantized data," *IEEE Trans. Signal Process.*, vol. 66, no. 8, pp. 2070–2085, Apr. 2018.

[11] D. Puthal, N. Malik, S. P. Mohanty, E. Kougianos, and C. Yang, "The blockchain as a decentralized security framework [future directions]," *IEEE Consum. Electron. Mag.*, vol. 7, no. 2, pp. 18–21, Mar. 2018.

[12] T. T. A. Dinh, R. Liu, M. Zhang, G. Chen, B. C. Ooi, and J. Wang, "Untangling blockchain: A data processing view of blockchain systems," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 7, pp. 1366–1385, Jul. 2018.

[13] P. Zhuang, T. Zamir, and H. Liang, "Blockchain for cybersecurity in smart grid: A comprehensive survey," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 3–19, Jan. 2021.

[14] M. A. Ferrag and L. Maglaras, "DeepCoin: A novel deep learning and blockchain-based energy exchange framework for smart grids," *IEEE Trans. Eng. Manag.*, vol. 67, no. 4, pp. 1285–1297, Nov. 2020.

[15] H. Liu et al., "Blockchain and federated learning for collaborative intrusion detection in vehicular edge computing," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6073–6084, Jun. 2021.

[16] V. Hassija, V. Gupta, S. Garg, and V. Chamola, "Traffic jam probability estimation based on blockchain and deep neural networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 3919–3928, Jul. 2021.

[17] X. Jiang, F. R. Yu, T. Song, Z. Ma, Y. Song, and D. Zhu, "Blockchain-enabled cross-domain object detection for autonomous driving: A model sharing approach," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 3681–3692, May 2020.

[18] Q. Yang and H. Wang, "Privacy-preserving transactive energy management for IoT-aided smart homes via blockchain," *IEEE Internet Things J.*, vol. 8, no. 14, pp. 11463–11475, Jul. 2021.

[19] J. Zhang, R. S. Blum, L. M. Kaplan, and X. Lu, "Functional forms of optimum spoofing attacks for vector parameter estimation in quantized sensor networks," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 705–720, Feb. 2017.

[20] J. Zhang, R. S. Blum, and H. V. Poor, "Approaches to secure inference in the Internet of Things: Performance bounds, algorithms, and effective attacks on IoT sensor networks," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 50–63, Sep. 2018.

[21] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," Decentralized Bus. Rev., Bitcoin, White Paper, Oct. 2008, p. 21260.

[22] E. Zaghloul, T. Li, M. W. Mutka, and J. Ren, "Bitcoin and blockchain: Security and privacy," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10288–10313, Oct. 2020.

[23] G. Karame, E. Androulaki, and S. Capkun, "Two Bitcoins at the price of one? Double-spending attacks on fast payments in Bitcoin," *IACR Cryptol. ePrint Arch.*, vol. 2012, no. 248, 2012.

[24] B. Hu, C. Zhou, Y. Tian, Y. Qin, and X. Junping, "A collaborative intrusion detection approach using blockchain for multimicrogrid systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 8, pp. 1720–1730, Aug. 2019.

[25] M. U. Hassan, M. H. Rehmani, and J. Chen, "Anomaly detection in blockchain networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 1, pp. 289–318, 1st Quart., 2023.

[26] X. Jiang, F. R. Yu, T. Song, and V. C. M. Leung, "Edge intelligence for object detection in blockchain-based Internet of Vehicles: Convergence of symbolic and connectionist AI," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 49–55, Aug. 2021.

[27] Y. Yang, Z. Guan, Z. Wan, J. Weng, H. H. Pang, and R. H. Deng, "PriScore: Blockchain-based self-tallying election system supporting score voting," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 4705–4720, 2021.

[28] F. P. Hjálmarsson, G. K. Hreioarsson, M. Hamdaqa, and G. Hjálmtýsson, "Blockchain-based e-voting system," in *Proc. IEEE 11th Int. Conf. Cloud Comput. (CLOUD)*, Jul. 2018, pp. 983–986.

[29] A. Lei, H. Cruickshank, Y. Cao, P. Asuquo, C. P. A. Ogah, and Z. Sun, "Blockchain-based dynamic key management for heterogeneous intelligent transportation systems," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1832–1843, Dec. 2017.

[30] S.-C. Cha, J.-F. Chen, C. Su, and K.-H. Yeh, "A blockchain connected gateway for BLE-based devices in the Internet of Things," *IEEE Access*, vol. 6, pp. 24639–24649, 2018.

[31] F. Lombardi, L. Aniello, S. De Angelis, A. Margheri, and V. Sassone, "A blockchain-based infrastructure for reliable and cost-effective IoT-aided smart grids," in *Proc. Living Internet Things, Cybersecur. IoT*, 2018, pp. 1–6.

[32] X. Ling, J. Wang, T. Bouchoucha, B. C. Levy, and Z. Ding, "Blockchain radio access network (B-RAN): Towards decentralized secure radio access paradigm," *IEEE Access*, vol. 7, pp. 9714–9723, 2019.

[33] L. Li et al., "CreditCoin: A privacy-preserving blockchain-based incentive announcement network for communications of smart vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 7, pp. 2204–2220, Jul. 2018.

[34] L. Zhou, L. Wang, Y. Sun, and P. Lv, "BeeKeeper: A blockchain-based IoT system with secure storage and homomorphic computation," *IEEE Access*, vol. 6, pp. 43472–43488, 2018.

[35] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. Berlin, Germany: Springer-Verlag, 1994.

[36] G. Shan, B. Zhao, J. R. Clavin, H. Zhang, and S. Duan, "Poligraph: Intrusion-tolerant and distributed fake news detection system," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 28–41, 2022.

[37] A. S. Rawat, P. Anand, H. Chen, and P. K. Varshney, "Collaborative spectrum sensing in the presence of Byzantine attacks in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 774–786, Feb. 2011.

[38] B. Kailkhura, S. Brahma, B. Dulek, Y. S. Han, and P. K. Varshney, "Distributed detection in tree networks: Byzantines and mitigation techniques," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 7, pp. 1499–1512, Jul. 2015.

[39] A. M. Antonopoulos, *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*. Sebastopol, CA, USA: O'Reilly Media, 2014.

[40] M. Pilkington, "Blockchain technology: Principles and applications," in *Research Handbook on Digital Transformations*. Cheltenham, U.K.: Edward Elgar Publishing, 2016.

[41] D. Stiawan, M. Y. Idris, R. F. Malik, S. Nurmaini, N. Alsharif, and R. Budiarto, "Investigating brute force attack patterns in IoT network," *J. Electr. Comput. Eng.*, vol. 2019, pp. 1–13, Apr. 2019.

[42] G. O. Karame, E. Androulaki, M. Roeschlin, A. Gervais, and S. Capkun, "Misbehavior in Bitcoin: A study of double-spending and accountability," *ACM Trans. Inf. Syst. Secur.*, vol. 18, no. 1, pp. 1–32, Jun. 2015.

[43] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, "Demystifying IoT security: An exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2702–2733, 3rd Quart., 2019.

[44] N. Kaur and S. K. Sood, "An energy-efficient architecture for the Internet of Things (IoT)," *IEEE Syst. J.*, vol. 11, no. 2, pp. 796–805, Jun. 2017.

[45] A. Ahmed, S. Abdullah, M. Bukhsh, I. Ahmad, and Z. Mushtaq, "An energy-efficient data aggregation mechanism for IoT secured by blockchain," *IEEE Access*, vol. 10, pp. 11404–11419, 2022.

[46] Y. Jiang and J. Zhang, "Vulnerability of finitely-long blockchains in securing data," 2023, *arXiv:2304.09965*.

[47] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

[48] J. Arora, *Introduction to Optimum Design*. Amsterdam, The Netherlands: Elsevier, 2004.

[49] Z. Q. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *J. Optim. Theory Appl.*, vol. 72, no. 1, pp. 7–35, Jan. 1992.

[50] I. Daubechies, M. Fornasier, and I. Loris, "Accelerated projected gradient method for linear inverse problems with sparsity constraints," *J. Fourier Anal. Appl.*, vol. 14, nos. 5–6, pp. 764–792, Dec. 2008.

[51] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Yiming Jiang received the B.Eng. degree in digital media technology and the M.Eng. degree in control science and engineering from Huazhong University of Science and Technology, Wuhan, China, in 2016 and 2019, respectively. He is currently working toward the Ph.D. degree in electrical engineering at Missouri University of Science and Technology, Rolla, MO, USA. His current research interests include blockchain, signal processing, and machine learning.

Jiangfan Zhang (Member, IEEE) received the B.Eng. degree in communication engineering from Huazhong University of Science and Technology, Wuhan, China, in 2008, the M.Eng. degree in information and communication engineering from Zhejiang University, Hangzhou, China, 2011, and the Ph.D. degree in electrical engineering from Lehigh University, Bethlehem, PA, USA, in 2016. From 2016 to 2018, he was a Post-Doctoral Research Scientist with the Department of Electrical Engineering, Columbia University, New York, NY, USA. Since 2018, he has been with the Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO, USA, where he is currently an Assistant Professor. His research interests include signal processing, machine learning, and their applications to cybersecurity, cyber-physical systems, the Internet of Things, smart grids, and blockchain. He was a recipient of the Dean's Doctoral Student Assistantship, the Gotshall Fellowship, and the P. C. Rossin Doctoral Fellow at Lehigh University.