Computers, Materials & Continua CMC, vol.64, no.3, pp.1705-1724, 2020

A Cross Layer Protocol for Fast Identification of Blocked Tags in Large-Scale RFID Systems

Chu Chu[1], Zhong Huang[1], Rui Xu[1], Guangjun Wen[1,*] and Lilan Liu[2]

Abstract: Blocker tag attack is one of the denial-of-service (DoS) attacks that threatens the privacy and security of RFID systems. The attacker interferes with the blocked tag by simulating a fake tag with the same ID, thus causing a collision of message replies. In many practical scenarios, the number of blocked tags may vary, or even be small. For example, the attacker may only block the important customers or high-value items. To avoid the disclosure of privacy and economic losses, it is of great importance to fast pinpoint these blocked ones. However, existing works do not take into account the impact of the number of blocked tags on the execution time and suffer from incomplete identification of blocked tags, long identification time or privacy leakage. To overcome these limits, we propose a cross layer blocked tag identification protocol (CLBI). CLBI consists of multiple rounds, in which it enables multiple unblocked tags to select one time slot and concurrently verify them by using tag estimation in physical layer. Benefiting from the utilization of most collision slots, the execution time can be greatly reduced. Furthermore, for efficient identification of blocked tags under different proportions, we propose a hybrid protocol named adaptive cross layer blocked tag identification protocol (A-CLBI), which estimates the remaining blocked tag in each round and adjusts the identification strategy accordingly. Extensive simulations show that our protocol outperforms state-of-the-art blocked tags identification protocol.

Keywords: RFID, blocked tag, physical layer, estimation, identification.

1 Introduction

Radio Frequency Identification (RFID) technology is promoting the rapid development of the Internet of things [Han, Zheng, Wen et al. (2018); Wang, Gao, Yin et al. (2018); Wang, Gao, Liu et al. (2019); Ren, Liu, Ji et al. (2018); Medhane, Sangaiah, Hossain et al. (2020); Yin, Zhou, Zhang et al. (2017)]. The current researches on RFID have focused more on the problems in practical application scenarios. i.e., rapid inventory of moving goods [Chen and Feng (2019); Wang, Xie, Wang et al. (2019)], complete identification of missing goods [Yu, Chen and Wang (2019); Shahzad and Liu (2015); Chen, Wang, Xia et al. (2018); Liu, Li, Min et al. (2015)], inventory of different kinds of goods [Liu, Li, Jie

[1] School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, 611731, China.

[2] School of Engineering, University of British Columbia, Kelowna, V1V 1V7, Canada.

[*] Corresponding Author: Guangjun Wen. Email: wgj@uestc.edu.cn.

Received: 15 February 2020; Accepted: 19 April 2020.

et al. (2016)]. In recent years, the security problem in RFID system has also attracted extensive attention [Vahedi, Shah-Mansouri, Wong et al. (2011); Bu, Liu, Luo et al. (2013); Chen and Feng (2009); Liu, Xie, Zhao et al. (2018); Luo, Wen, Su et al. (2019)], among which the blocking attack can interfere with reader's identification of tags. Specifically, the blocker tag simulates a fake tag that has the same ID as the target tag. Once the reader wants to collect information about a target tag in a certain time slot, the fake tag will also respond to noise in this slot. If not be prevented, blocking attack can lead to incomplete information collection, thus prolonging identification time and causing energy waste.

While existing techniques, i.e., encryption algorithms and cryptography [Bolotnyy and Robins (2017)], can be used to prevent the blocking attack, they are not suitable for low-cost tags. Another direct method called Poll & Listen (P & L) [Liu, Xie, Zhao et al. (2018)] identifies each tag one by one using the select command that contains the tags' ID. However, P & L is time consuming and may reveal personal privacy. Consider that in a hospital, the reader needs to quickly collect the patient's physical conditions such as temperature, heartbeat and blood pressure and provides them to the doctor. If some patients' information cannot be timely fed back due to the existence of blocked tags, the diagnosis of the disease will be delayed. At the same time, doctors should protect patients' personal information from being disclosed. Therefore, it is crucial to confirm exactly which tags are blocked and take effective action accordingly.

So far, many anti-collision algorithms have achieved high throughput and time efficiency [Su, Sheng, Leung et al. (2019); Su, Sheng, Liu et al. (2020); Su, Chen, Sheng et al. (2020); Chen (2016)]. However, if there exist blocked tags, the reader cannot use the traditional Aloha method to separate the collision information. i.e., the fake tags and blocked tags always reply like a pair. Therefore, the ID of some specific tags cannot be determined. The tag identification protocols [Liu, Li, Min et al. (2015); Liu, Li, Min et al. (2014)], which construct the expected vector by using the IDs of the known tags and improve the proportion of singleton time slot in the vector by the use of bloom filter, multiple-hash, collision reconcilement and other methods. These protocols can be used to determine the existence of the blocking attack by comparing whether a singleton slot turns out to be collision during the actual execution. However, through extensive simulations, we find that these methods also suffer from performance degradation when the number of blocked tags changes.

Different from most MAC layer tag identification protocols that only consider the use of singleton slots [Liu, Li, Min et al. (2015); Su, Sheng, Liu et al. (2020); Liu, Qi, Li et al. (2015)], we aim to extract information from the physical layer to assist in the design of the MAC layer protocol, so as to improve the identification efficiency of the blocked tag. Follow this idea, a cross layer blocked tag identification protocol (CLBI) is proposed. CLBI improves the time efficiency by extracting the information from not only the singleton slots but also collision slots. Through the tag estimation in collision slots, multiple unblocked tags can be verified in only one slot. Moreover, through extensive simulations, we observe that the efficiency of CLBI decreases with the increase of the blocked tags, hence an adaptive cross layer blocked tag identification protocol (A-CLBI) is proposed to achieve time efficiency with different number of blocked tags. The main

contributions of this paper are summarized as follows:

1. We investigate the problem of blocked tags identification in RFID system and propose a cross layer method. Compare it with existing approaches, the proposed one can achieve higher time efficiency when the proportion of the blocked tags is small.
2. We estimate the number of blocked tags in each round and design an adaptive cross layer blocked tag identification protocol that is suitable for different proportion of blocked tags.
3. We theoretically analyze the optimization of parameters and calculate the optimal frame length of each round to maximize the time efficiency.
4. We conduct extensive simulation analysis to evaluate the performance of CLBI and A-CLBI under different working scenarios, and the results show that our approach outperforms the state-of-the-art.

The rest of the paper is organized as follows: Section II reviews the related work. Section III presents the system models used in this study and formulates the problem. In Section IV, we propose our approach to solve the problem. In Section V, the proposed A-CLBI and CLBI are evaluated and compared with the state-of-the-art protocols. Section VI concludes the paper.

2 Related work

Research based on blocked tags has attracted more and more attention [Vahedi, Shah-Mansouri, Wong et al. (2011); Bu, Liu, Luo et al. (2013); Liu, Xie, Zhao et al. (2018); Zanetti, Fellmann and Capkun (2010); Lehtonen, Michahelles and Fleisch (2009); Wang, Xiao, Bu et al. (2013)]. However, the current works still face three problems: (1) incomplete identification of blocked tags (2) long identification time (3) privacy leakage. In Vahedi et al. [Vahedi, Shah-Mansouri, Wong et al. (2011)], the Probabilistic Blocker Tag Detection (P-BTD) algorithm was proposed for both Binary Tree walking systems and Aloha system. P-BTD compares information extracted from the previous interrogation with the current one to determine whether a blocked tag exists. However, P-BTD cannot accurately and completely identify the specific blocked tag, thus cannot fundamentally eliminate the harm of blocking attack. Synchronized Secrets (SYNC) was implemented to identify the specific clone tag [Chen and Feng (2009)]. SYNC scans the tag multiple times and each time the reader writes a random number into the tag's memory and records it. When it scans this tag next time, it recognizes the clone ID if it gets a different random number. Liu et al. [Liu, Xie, Zhao et al. (2018)] proposed a hybrid approach which consists of the Aloha Filtering (AF) and Poll & Listen (P & L) protocol, the approach firstly uses AF to filter most unknown tags and unblocked tags, then it adopts P & L to polls IDs of the target tags one by one and monitors the channel to check the responses of these tags. Once receiving a collision response, the corresponding tag can be identified as a blocked tag by the reader. However, both SYNC and P & L are time consuming, which are not applicable to large-scale RFID system. Moreover, these two protocols transmit the tags' ID directly in the air, thus may lead to privacy disclosure. Wang et al. [Wang, Xiao, Bu et al. (2013)] proposed the Tree-based protocol for blocked tag identification. However, the tree-based mechanism is not supported by the C1G2

RFID standard. In Bu et al. [Bu, Liu, Luo et al. (2013)], Broadcast Friendly Cloned-Tag Identification (BID) protocol was proposed based on slotted Aloha mechanism. In BID, each tag uses its ID and hash function to select a slot to reply to the reader. Specifically, the reader first predicts and records the expected singleton slots in a frame, and then detects the actual reply of these singleton slots. Once an expected singleton slot turns out to be a collision slot in the execution phase, the blocked tag can be identified. However, this approach only extracts the information of single slots with nearly 73.2% of the slots being wasted. Therefore, BID is also time-consuming when the proportion of blocked tags is small.

3 Problem statement

3.1 System model

In this paper, we consider an RFID system, which consists of a back-end server, an RFID reader and N tags [Chen, Wang, Xia et al. (2018)]. Each tag has a unique 96-bit ID and is equipped with a hash function. All the IDs of N tags are prestored in the database of the back-end server, which communicates with the reader via a high data rate link. Hence, we consider them as an integral part and use the "reader" to represent them. Note that in a large-scale application scenario, multiple readers can be treated as one if they are well synchronized and coordinated. For simplicity, our protocol only considers a single-reader in this paper and can be extended to multiple readers.

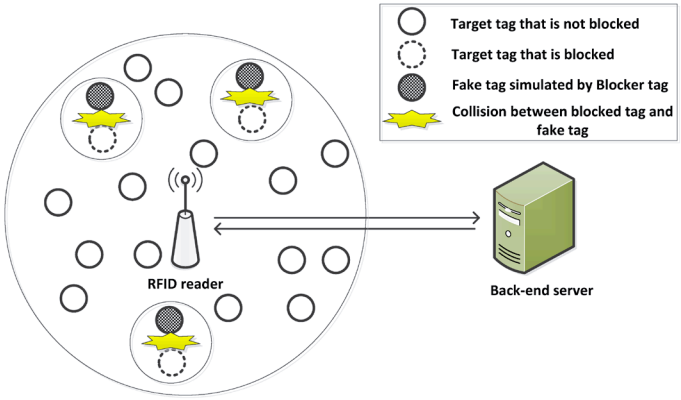

Figure 1: System model of the studied problem

3.2 Attack model

The attack model [Bu, Liu, Luo et al. (2013); Zanetti, Fellmann and Capkun (2010); Lehtonen, Michahelles and Fleisch (2009)] is considered in this paper. As illustrated in Fig. 1, some fake tags simulated by blocker tags are pre-configured in some valid tags with the same IDs in set N and we called these valid tags as blocked tags. When the reader intends to collect information of a certain set of tags, each tag uses its ID and hash function $H(\bullet)$ to calculate $H(ID,r) \bmod f$ as the selected slot to reply. However, the fake tag bundled with a blocked tag will also reply with a random number in the same

slot, resulting in the reader not being able to correctly receive any useful information about the blocked tag.

3.3 Problem definition

Since blocking attacks will lead to failure of information collection, which increases the waste of time and energy, in this paper, we focus on how to completely identify the blocked tags in an efficient way. As shown in Fig. 1, the set of all target tags that we want to verify is denoted as N, whose IDs are known by the reader in advance. The set of blocked tags in set N is represented by B. Obviously, $B \subseteq N$. Neither the IDs nor the number of blocked tags is known by the reader. Therefore, the problem can be summarized as follow: Given the target tags set N known by the reader, identify all blocked tags B in N with the minimum execution time.

Tab. 1 summaries the main notations used in the paper.

Table 1: Key notation

Symbols	Description
N	The set of all target tags in the system
B	The set of blocked tags in the system
$H(\bullet)$	The hash function with a uniform rand distribution
r_i	The random seed that is fresh in each round i
f_i	The length of the filter in each round i
V	Filter vector
T	The total execution time of the i th round
t_{long}	The long-response slot which is 10 bits
t_{tag}	The tag slot which is 96 bits
q	A variable given by N/f_i.
e	The natural constant which is approximately equal to 2.71828.

4 Cross layer design

4.1 Motivation

In order to extract useful information in collision slots, we combine the physical layer tag estimation algorithm to design the cross layer blocked tag identification protocol (CLBI). As we have described above, previous protocols determine a blocked tag by checking whether the expected single slot becomes an actual collision slot. We illustrate the basic idea of CLBI in Fig. 2, that is, if the number of tags in the expected slot is equal to the number actual slot, all tags in this slot can be identified as non-blocked tags. Moreover, through the subsequent optimization of frame length, the number of collision slots and single slots can be increased, thus reducing the waste of empty slots. Therefore, CLBI can identify more tags with fewer time slots, thereby fundamentally improving the time efficiency.

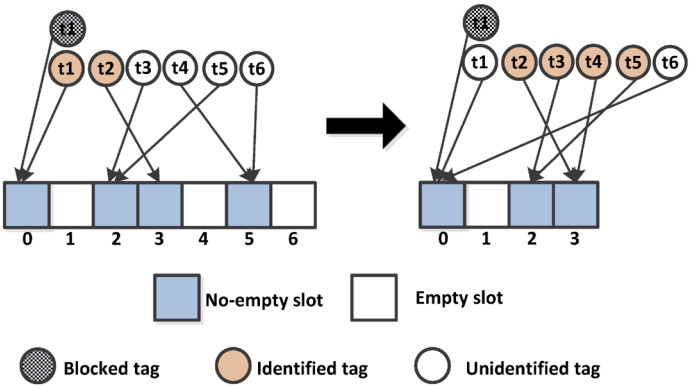

Figure 2: An example of identifying more tags with less time slots

4.2 Improved mean-shift algorithm

In the RFID system, the reader can down-converted the signal replied by multiple tags to the baseband and extract useful information. we can plot the base signal in the I/Q (in-phase amplitude/quadrature amplitude) plane [Angerer, Langwieser and Rupp (2010)]. Hence, the received signals are gathered around some center points, which form as several clusters. The number of collision tags can be obtained by counting the number of clusters. However, due to the presence of noise, it is difficult to determine the cluster boundary.

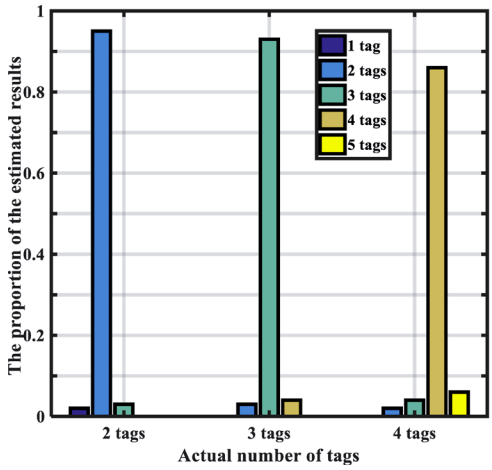

Figure 3: Estimation results of PEAC

In Huang et al. [Huang, Xu, Chu et al. (2019)], Physical Layer Estimation Algorithm based on Clustering (PEAC) was proposed to improve the Mean-Shift algorithm [Trieu and Maruyama (2011)]. Based on the feature that clusters are always located in pairs and symmetry to the center of all samples. PEAC adjusts the number of clusters, thus improving the accuracy of the estimation of tags in one slot. Fig. 3 plots the estimation accuracy of different tags. It can be observed that the accuracy of results is affected by the number of tags in a slot. In our design, the estimation results of tags in physical layer

indicate that whether there exist extra tags, i.e., fake tag replies in a time slot. Taking advantage of the effective information in the physical layer, we design the so-called cross layer blocked tag identification protocol that will be described in the next section.

4.3 The proposed solution

CLBI consists of three phases: slot pre-allocation phase, tag filtering phase and tag identification phase. In slot pre-allocation phase, the reader allocates each tag a time slot and constructs a vector-based frame to broadcast accordingly. Once receiving the frame, each tag then decides whether to reply or keep silent based on the value of its selected time slot in the frame during the tag filtering phase. Subsequently, the reader estimates the number of tags in each slot and determines the status of these tags. We describe the detailed protocol process as follows.

4.3.1 Slot pre-allocation phase

As shown in Fig. 4, in each round i, the reader first generates a random seed r_i and calculates the frame length f_i. Then it calculates $H(ID, r_i) \bmod f_i$ as the time slot for each tag to select. Since the reader knows the IDs of all tags, it can construct an f_i-bit vector, represented as V, according to the status of each slot. Note that time slots are classified into empty slots, singleton slots, or collision slots. To save the number of bits, we use Huffman coding to indicate the different states of the slot according to the following rules:

a) "0": the time slot is expected to be empty.
b) "10": the time slot is expected to be mapped by one tag.
c) "11": the time slot is expected to be mapped by multiple tags.

Note that if one time slot is selected by too many tags, the number of clusters will increase exponentially by 2, and the clusters will be closer to each other, which leads to an increase in detection errors. Therefore, we treat the time slot selected by more than 3 tags as "0", which is the same as the empty slot. An example is shown in Fig. 3, t_1, t_2, t_3, t_4, t_5, t_6 respectively select the first, second, thirdly, fifth slot. Based on the expected number of tags in each slot, the vector is encoded as "111110010". Then the reader splits the V into $\lceil V/96 \rceil$ segments and broadcasts them as well as the frame length f_i and random seed r_i to all the tags N.

4.3.2 Tag filtering phase

Upon receiving the parameters f_i, r_i and the vector V, each active tag also locates their associated slot j by calculating $H(ID, r_i) \bmod f_i$. In addition, it checks the value of jth bit in V and records the number of slots before the its selected slot with the value of "10" and "11", which are denoted as m and l, respectively. If the value happens to be "11", the tag will transmit a long response in $l_j + m_j + 1$th slot and wait for the next

command A from reader. If the value happens to be "10", the tag will also transmit a long

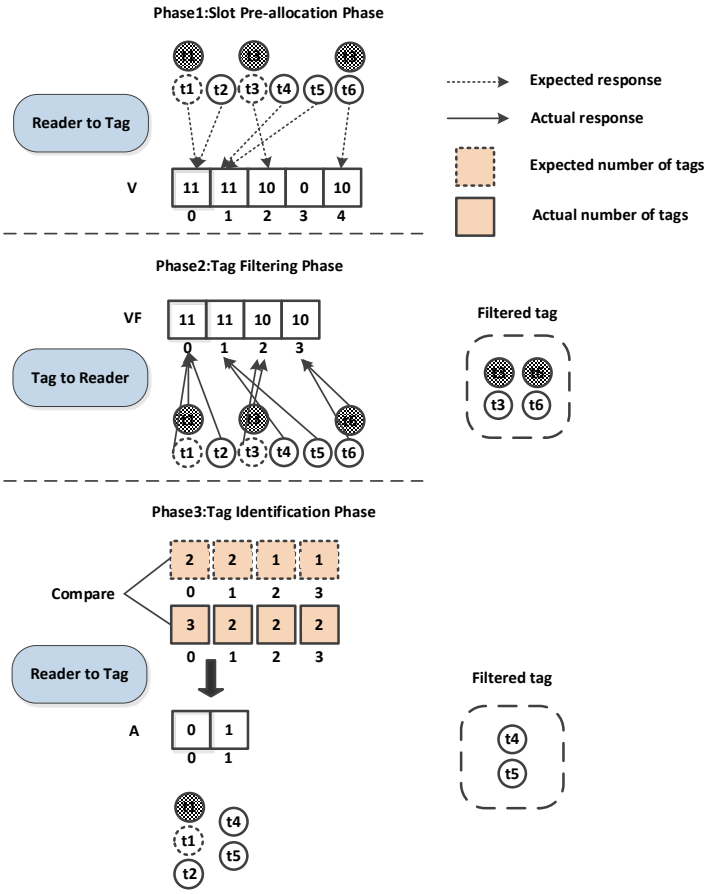

Figure 4: Illustration of the proposed CLBI protocol

response in $l_j + m_j + 1$ th slot but will not participate in the subsequent identification process. Otherwise, it will keep silent until next round begins. As shown in Fig. 3, t_1, t_2, t_3, t_4 and t_5 find the value of their associated slots in V are "11", then t_1, t_2 reply in the first slot i.e., m=0 and $l = 0$, and t_4, t_5 reply in the second slot, i.e., m=0 and $l = 1$. Similarly, t_3 replies in the third slot and t_6 replies in the fourth slot. However, t_3 and t_6 will not participate in the subsequent process.

4.3.3 Tag identification phase

In tag identification phase, the reader checks the actual replies of each slot from the active tags and constructs an actual frame denoted as VF. If the reader detects collision signals in a slot, it plots the I/Q constellation as described above, and then adopts the

PEAC algorithm to extract the number of tags in this slot. The expected number of tags and the actual number of tags in a slot are denoted as EN and AN, respectively. By comparing EN and AN, the reader can know whether any tags are blocked. Note that only expected singleton slots and expected collision slots are used in VF, so we have the following cases:

Case 1: if VF=10 and AN=1, the tag corresponding to the j th slot is not blocked.

Case 2: if VF=10 and $AN>1$, the tag corresponding to the j th slot is blocked.

Case 3: if VF=11 and $EN=AN$, the tags corresponding to the j th slot are not blocked.

Case 4: if VF=11 and $EN\neq AN$, it means that at least one of these tags is blocked, thus the tags corresponding to the j th slot cannot be identified.

Subsequently, the reader constructs a vector, represented as A, whose length is equal to the total number of "11" slots. According to these four cases: $A(j)=1$ in case 3, otherwise $A(j)=0$. Then, it also splits the A into $\lceil A/96 \rceil$ segments and broadcasts them to tags. Each remaining active tag checks the value of l_j th bit in A. If the value happens to be "1", it will not participate in the subsequent identification process. Otherwise, they will wait for the start of the next round. As shown in Fig. 4, the reader compares the expected and actual number of tags in each slot. Then it verifies that t_3, t_6 are blocked while t_4, t_5 are not blocked. Therefore, the reader constructs the vector A and encodes it as "01" to broadcast to all active tags. Consequently, t_4 and t_5 will not participate in the follow-up process and t_1, t_2 will wait for the start of the next round.

4.4 Parameter optimization

In this subsection, we determine the optimal frame size f_i in each round i to maximize the time efficiency of CLBI. We first assume that the current number of blocked tags B^* is known by us, and then we'll estimate it later.

We denote N^* as the candidate tags (fake tags are not included) at the beginning of each round. For arbitrary time slot, the probability that it is selected by only one tag in N^* is:

$$P_1 = C_{|N^*|}^1 \times \left(\frac{1}{f_i}\right) \times \left(1-\frac{1}{f_i}\right)^{|N^*|-1} \approx \frac{|N^*|}{f_i} \times e^{-\frac{|N^*|-1}{f_i}} \tag{1}$$

Note that when f_i is very large, $\left(1-\frac{1}{f_i}\right)^{|N^*|-1}$ is approximated to $e^{\frac{|N^*|-1}{f_i}}$. Similarly, we calculate the probability of 2-collision slots and 3-collision slots as follows:

$$P_2 = C_{|N^*|}^2 \times \left(\frac{1}{f_i}\right)^2 \times \left(1-\frac{1}{f_i}\right)^{|N^*|-2} \approx \frac{|N^*| \times |N^*-1|}{2f_i^2} \times e^{-\frac{|N^*|-2}{f_i}} \tag{2}$$

$$P_3 = C_{|N^*|}^3 \times \left(\frac{1}{f_i}\right)^3 \times \left(1-\frac{1}{f_i}\right)^{|N^*|-3} \approx \frac{|N^*| \times |N^*-1| \times |N^*-2|}{6f_i^3} \times e^{-\frac{|N^*|-3}{f_i}} \tag{3}$$

Hence, we obtain the expected number of useful slots in arbitrary round i, which are denoted as D, as follow:

$$D = f_i \times (P_1 + P_2 + P_3) \tag{4}$$

In the tag identification phase, a tag can be identified in two cases: a tag selects a single slot or multiple tags select the same slot (no more than three), and none of which are blocked. Hence, we denote P_{11}, P_{22} and P_{33} to represent the probability of each case and respectively calculate them as follow:

$$P_{11} = C_{|N^*|-|B^*|}^1 \times \left(\frac{1}{f_i}\right)^1 \times \left(1-\frac{1}{f_i}\right)^{|N^*|-1} \approx \frac{|N^*|-|B^*|}{f_{\mathrm{i}}} \times e^{-\frac{|N^*|-1}{f_i}} \tag{5}$$

$$P_{22} = C_{|N^*|-|B^*|}^2 \times \left(\frac{1}{f_i}\right)^2 \times \left(1-\frac{1}{f_i}\right)^{|N^*|-2} \approx \frac{(|N^*|-|B^*|) \times (|N^*|-|B^*|-1)}{2f_i^2} \times e^{-\frac{|N^*|-2}{f_i}} \tag{6}$$

$$P_{33} = C_{|N^*|-|B^*|}^3 \times \left(\frac{1}{f_i}\right)^3 \times \left(1-\frac{1}{f_i}\right)^{|N^*|-3} \approx \frac{(|N^*|-|B^*|) \times (|N^*|-|B^*|-1) \times (|N^*|-|B^*|-2)}{6f_i^3} \times e^{-\frac{|N^*|-3}{f_i}} \tag{7}$$

Let denote K_{total} as the total expected number of tags that can be identified in each round i, we can obtain:

$$K_{total} = (P_{11} + 2P_{22} + 3P_{33}) \tag{8}$$

Let T represent the total execution time of the i th round in CLBI. T includes three parts: the time for the reader to broadcast the parameters and the vector V, the time required for the tag to reply, and the time for the reader to broadcast the vector A. Hence, we get:

$$T = \left\lceil \frac{f_i+1}{96} \right\rceil \times t_{tag} + D \times \mathrm{t}_{\mathrm{long}} + \left\lceil \frac{f_i \times (P_2+P_3)}{96} \right\rceil \times t_{tag} \tag{9}$$

We denote the proportion of unblocked tags in each round as p and replace N/f_i with q. Combing Eqs. (2), (3), (5)-(7) and (9), the average time to identify a tag in the i th round, which is denoted as E, can be calculated as follow:

$$\begin{aligned} E &= \frac{T}{K} \\ &= \frac{\left\lceil \frac{f_i+1}{96} \right\rceil \times t_{tag} + D \times t_{long} + \left\lceil \frac{f_i \times (P_2+P_3)}{96} \right\rceil \times t_{tag}}{f_i \times (P_{11}+2P_{22}+3P_{33})} \\ &\approx \frac{\frac{t_{tag}}{96} + (q+\frac{1}{2}q^2+\frac{1}{6}q^3) \times e^{-q} \times t_{long} + \frac{\left(\frac{1}{2}q^2+\frac{1}{6}q^3\right) \times e^{-q} \times t_{tag}}{96}}{\left[q + p^2 \times q^2 + \frac{1}{2}p^3 \times q^3\right] \times e^{-q}} \end{aligned} \tag{11}$$

Note that we set $t_{tag} = 2.4$ ms and $t_{long} = 0.8$ ms which follow the settings in Bu et al. [Bu, Liu, Luo et al. (2013)], we calculate the derivative of E as follows:

$$\begin{aligned} E' = &\frac{(\frac{4}{5} + \frac{q}{40} - \frac{11}{80}q^3)}{\left[q + p^2q^2 + \frac{1}{2}p^3q^3\right]} - \frac{\left(\frac{4}{5}qe^{-q} + \frac{33}{80}q^2e^{-q} + \frac{11}{80}q^3e^{-q} + \frac{1}{40}\right)}{\left[q + p^2q^2 + \frac{1}{2}p^3q^3\right]^2 \times e^{-q}} \\ &\times \left(1 - q - p^2q^2 + \frac{3}{2}p^3q^2 - \frac{1}{2}p^3q^3 + 2p^2q\right) \end{aligned} \tag{12}$$

Given p, E' is a function of q. Let E'=0, we can get $q = q_0$. When $q < q_0$, $E' < 0$ and when $q > q_0$, $E' > 0$. For example, if p is set to 0.9, we can get q_0=0.225 to satisfy E'=0. Hence, the average time to identify a tag is approximately equal to 0.5033 ms.

4.5 Protocol optimization

In our designed CLBI, when there are a small number of blocked tags, only one time slot is needed to identify two or three tags thus achieves higher efficiency. However, as the CLBI is continuously executed, the proportion of blocked tags may increase. Another situation is that the number of blocked tags is high initially. Therefore, directly executing the CLBI may not be the most efficient way. To solve this problem, we further propose the adaptive cross layer blocked tag identification protocol (A-CLBI).

A-CLBI determines the execution strategy based on the predicted number of total blocked tags. If the proportion of blocked tags is small, A-CLBI chooses to execute the CLBI because CLBI can identify multiple tags within a time slot. On the contrary, if the proportion of blocked tags increases, the ES-BID method is more appropriate to solve the problem. Therefore, we denote this threshold of the blocked tag proportion as p_0.

According to Bu et al. [Bu, Liu, Luo et al. (2013)], the optimal frame length is set as equal to the number of candidate tags N^*. Therefore, the average time to identify a blocked tag is:

$$
\begin{aligned}
E_{ES-BID} &= \frac{T_{ES-BID}}{K_{ES-BID}} \\
&= \frac{\left\lceil \frac{f_i+1}{96} \right\rceil \times t_{tag} + f_i \times P_1 \times t_{long}}{f_i \times P_{11}} \\
&\approx \frac{\frac{t_{tag}}{96} + (q) \times e^{-q} \times t_{long}}{q \times e^{-q}} \\
&\approx 0.8680
\end{aligned}
\tag{13}
$$

Then, we set $E_{CLBI} = E_{ES-BID}$ and get:

$$
\frac{\frac{t_{tag}}{96} + (q + \frac{1}{2}q^2 + \frac{1}{6}q^3) \times e^{-q} \times t_{long} + \frac{\left(q + q^2 + \frac{1}{3}q^3\right) \times e^{-q} \times t_{tag}}{96}}{\left[q + p^2 \times q^2 + \frac{1}{2}p^3 \times q^3\right] \times e^{-q}} = 0.8680 \tag{14}
$$

Fig. 5 plots the average time spent by the two protocols to identify one tag. We can obtain the threshold p_0 as: p_0=0.2980 . Therefore, in A-CLBI protocol, when the proportion of blocked tags is less than p_0, it adopts CLBI method to quickly identify them. As the proportion of blocked tags exceeds p_0, the ES-BID approach is adopted.

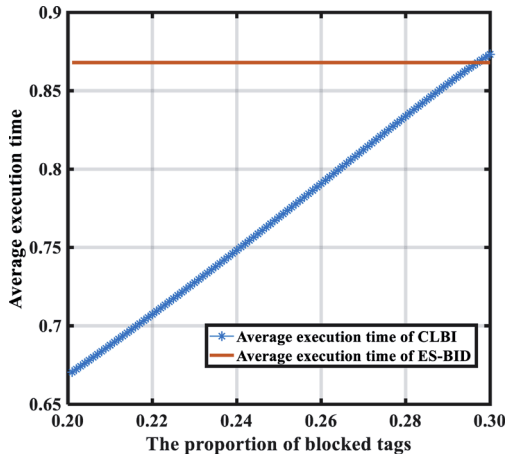

Figure 5: The average execution time for verifying a tag with respect to *p*

4.6 Cardinality estimation

In this section, we propose a quantity estimation method to fast estimate the number of blocked tags. Although there are many protocols [Chen, Zhou and Yu (2013); Liu, Xiao, Li et al. (2017)] that can accurately detect the number of tags, mixing these methods in the identification process will cause additional overhead and increases the protocol complexity. In the A-CLBI protocol, the number of slots with 3, 2, 1 tag selection can be obtained by using the physical layer information, which further assists the estimation of

the blocked tags.

Let N_j denote the expected number of slots selected by j tags and N_{jj} represent the expected number of slots in which the actual number of tags j is equal to the expected ones i.e., j= 1,2,3. Based on the difference between N_j and N_{jj}, we can estimate the remaining unidentified blocked tags.

According to Eqs. (1) and (2), we calculate the expected number of slots selected by 1 tag, 2 tags and 3 tags as:

$$\begin{aligned} E_1 &= f_i \times P_1 \approx \left|N^*\right| \times e^{-\frac{\left|N^*\right|-1}{f_i}} \\ E_2 &= f_i \times P_2 \approx \frac{\left|N^*\right| \times \left(\left|N^*\right|-1\right)}{2f_i} \times e^{-\frac{\left|N^*\right|-2}{f_i}} \\ E_3 &= f_i \times P_3 \approx \frac{\left|N^*\right| \times \left(\left|N^*\right|-1\right) \times \left(\left|N^*\right|-2\right)}{6f_i^2} \times e^{-\frac{\left|N^*\right|-3}{f_i}} \end{aligned} \tag{15}$$

Similarly, the expected number of E_{jj} can also be obtained as: $E_{11} = f_i \times P_{11}$, $E_{22} = f_i \times P_{22}$ and $E_{33} = f_i \times P_{33}$. Combing E_{N_j} and $E_{N_{jj}}$, we can obtain the expression of B^*, denoted as B_1, B_2, B_3, as follows:

$$\begin{aligned} \left|B_1\right| &= \left|N^*\right| - \frac{E_{N_{11}}}{E_{N_1}} \times \left|N^*\right| \\ \left|B_2\right| &= \left|N^*\right| - \sqrt{\frac{E_{N_{22}}}{E_{N_2}} \times \left|N^*\right|^2} \\ \left|B_3\right| &= \left|N^*\right| - \sqrt[3]{\frac{E_{N_{33}}}{E_{N_3}} \times \left|N^*\right|^3} \end{aligned} \tag{16}$$

Then, substituting N_j for E_{N_j} and N_{jj} for $E_{N_{jj}}$, we get the estimators of $\left|\widehat{B_1}\right|$, $\left|\widehat{B_2}\right|$, $\left|\widehat{B_3}\right|$.

When one frame is completely executed, the number of 1 slot, 2 slots and 3 slots may be different, thus resulting in different values of $\left|\widehat{B}\right|$. In order to improve the accuracy of the estimation, the reader first counts the number of N_{11}, N_{22} and N_{33}, and then set $N_{\max} = \max\left\{N_{11}, N_{22}, N_{33}\right\}$ to estimate the value of $\left|B^*\right|$ for each round. We denote K_B as the number blocked tag identified in each round and get K_B as follow:

$$\begin{aligned} K_{total} &= f_i \times C_{\left|B^*\right|}^1 \times \left(\frac{1}{f_i}\right) \times \left(1 - \frac{1}{f_i}\right)^{\left|N^*\right|-1} \\ &\approx \left|B^*\right| \times e^{-\frac{\left|N^*\right|-1}{f_i}} \end{aligned} \tag{17}$$

Therefore, we can calculate the proportion of blocked tags as:

$$p = \frac{\left|\widehat{B_i}\right| - K_B}{\left|N^*\right| - K_{total}} \tag{18}$$

5 Performance evaluation

The performance of A-CLBI is compared with state-of-the-art protocols, i.e., BID, ES-BID [Bu, Liu, Luo et al. (2013)], SWIPT [Liu, Li, Ming et al. (2015)] and Poll & Listen [Liu, Xie, Zhao et al. (2018)]. Note that SWIPT has been modified where each tag transmits a long response in a singleton time slot. We adopt the execution time of the protocols as the performance metric. The communication channel between the reader and the tags is considered as error-free. According to the setting in Bu et al. [Bu, Liu, Luo et al. (2013); Liu, Li, Ming et al. (2015); Liu, Xie, Zhao et al. (2018)], we set $t_{tag} = 2.4$ ms to transmit a tag ID or a segment (96 bits). A long response is required by the reader to distinguish between empty, singleton, and collision slots. Here, $t_{long} = 0.8$ ms. For reliability, two hundred experiments are carried out for each parameter groups and the results are averaged.

5.1 Impact of number of blocked tags

As shown in Fig. 6, we compare the A-CLBI with existing protocols when varying the number of blocked tags B from 0 to 10000, 0 to 100, and 500 to 2500. Here we set $|N|=10000$. It can be observed that the execution time of SFMTI, S-BID, ES-BID and P & L are stable. The reason is that all of these protocols are designed to use the singleton time slot to identify the corresponding tag. Therefore, these protocols do not consider the impact of the number of blocked tags on execution time. CLBI achieves better performance when the number of blocked tags is small. This is because CLBI enables multiple unblocked tags to hash to one slot and concurrently identify them. However, as the number of blocked tag increases, the performance of CLBI degrades. The reason is that most of the tags are blocked, resulting in the waste of 2-collision and 3-collision slots used to identify unblocked tags. A-CLBI achieves the best time efficiency when the number of blocked tags is small. Moreover, even B increases, A-CLBI can always get the optimal time efficiency by adaptively adjusting the size of the frame and the strategy to be used. For CLBI and A-CLBI, it is also shown in Fig. 6(b) that the execution time of each is close. Since the proportion of blocked tags changes from 0 to 0.01, CLBI can achieve desirable performance when there are few blocked tags. When the number of blocked tags increases as shown in Fig. 6(c), A-CLBI gains better performance.

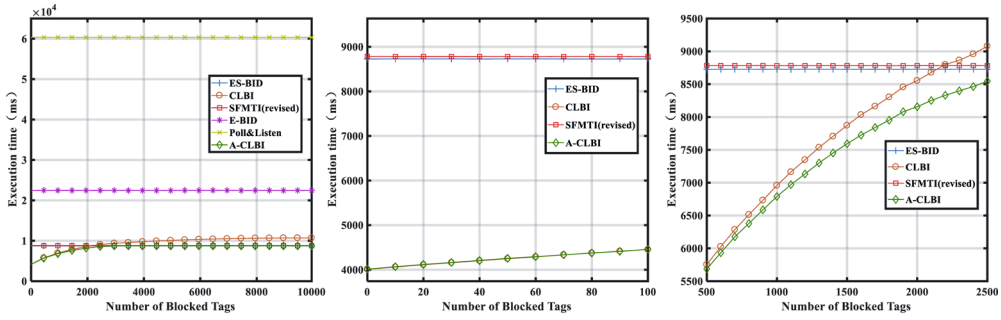

Figure 6: Execution time of different protocols when varying the number of blocked tags

5.2 Impact of number of target tags

The performance of A-CLBI is verified in Fig. 7 by varying the number of target tags N. Here we vary $|N|$ from 1000 to 10000 and $|B|=0.05|N|$ in (a) and $|B|=100$ in (b). It can be observed in Fig. 7(a) that the execution time of all these protocols increases with respect to the number of target tags. The reason is because all the protocols need more slots to identify more target tags. P & L consumes the most execution time since it identifies each tag by broadcasting its ID one by one. The execution time of SFMTI and ES-BID is close. A-CLBI achieves the optimal performance compared with other protocols. Moreover, when the proportion of blocked tags is fixed as shown in Fig. 7(b), the execution time of A-CLBI protocol grows slower than other protocols as the number of target tags increases.

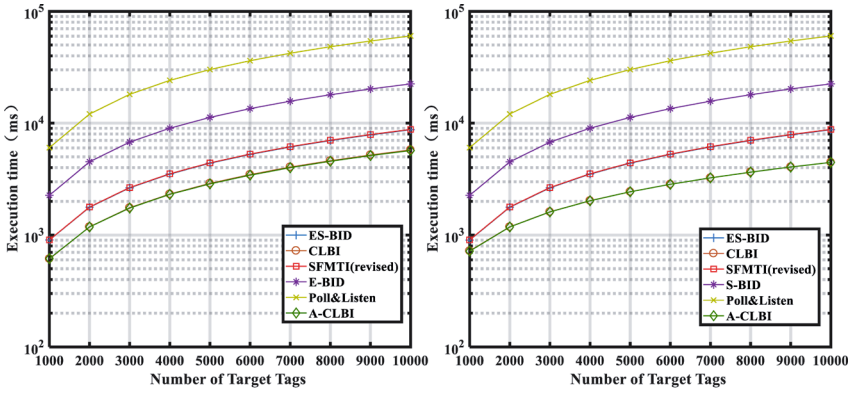

Figure 7: Execution time of different protocols when varying the number of target tags

5.3 Impact of different hashing strategies

In Fig. 8, we evaluate the impact of different hashing strategies. Here we set $|N|=10000$, then vary $|B|$ from 0 to 500 in (a) and change $|B|$ from 1000 to 10000 in (b). According to the design of CLBI and A-CLBI, part of collision slots (2, 3, 4, ...) can be used to fast filter the unblocked tags. Considering the error of physical layer estimation, we compare three hashing strategies for CLBI using 2-collision slots, 2 & 3 collision slots, 2-4 collision slots in terms of execution time. We represent each strategy as CLBI (S2), CLBI (S3) and CLBI (S4), respectively.

It can be observed that when $|B|/|N|$ is low, CLBI (S4) outperforms CLBI (S2) and CLBI (S3). This is because most target tags are not blocked, the reader can utilize more collision tags to identify multiple unblocked tags together. However, as the proportion of $|B|/|N|$ increases, the execution time of three strategies becomes close. Moreover, when $|B|/|N|$ is large, the performance of CLBI (S2) is superiors to that of CLBI (S3) and CLBI (S4). The reason is because more blocked tags cause more collision slots to be

wasted. A-CLBI is also compared with these three approaches. As $|B|/|N|$ increases, A-CLBI can adjust the frame size and protocol to get better performance.

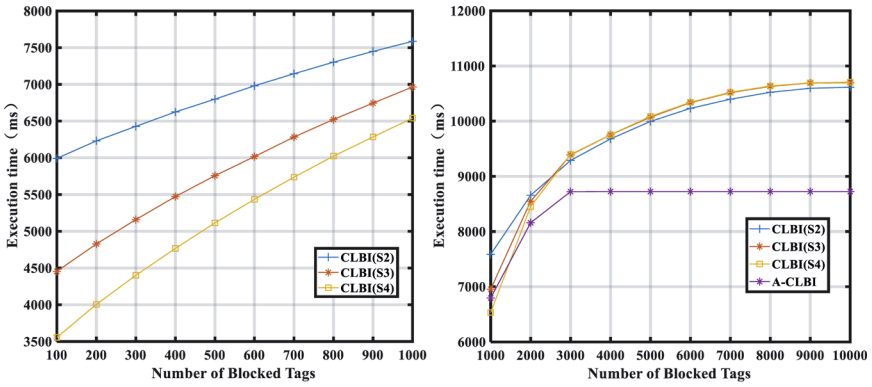

Figure 8: Execution time of different hashing strategies when varying the number of blocked tags

5.4 Impact of estimation error

In Fig. 9, we evaluate the impact of clustering error in the physical layer by comparing the execution time and the accuracy of the proposed CLBI and A-CLBI protocols. As we illustrated above, the execution time of CLBI grows faster than A-CLBI as the proportion of B increases, hence we set $|N|$=10000 and $|B|$ changes from 1000 to 2000.

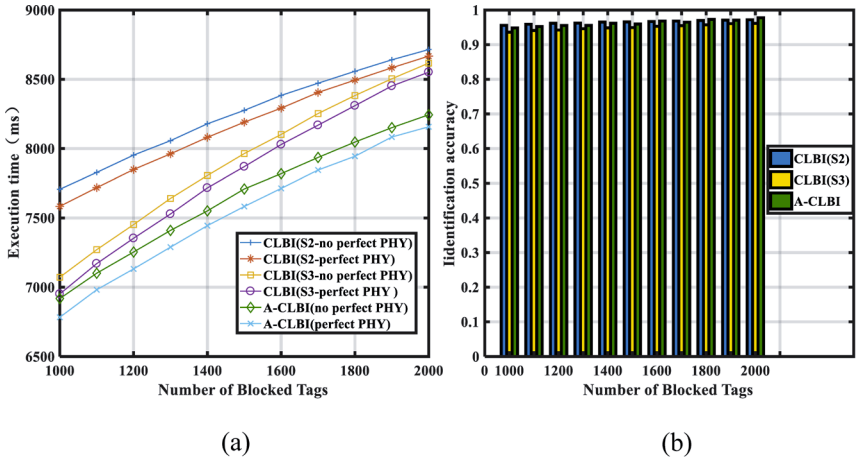

Figure 9: Impact of clustering error in physical layer

It can be observed in Fig. 9(a) that the execution time of both CLBI (S2), CLBI (S3) and A-CLBI is prolonged due to the clustering error. This is because some unblocked 2 or 3 collision tags cannot be verified timely due to the quantity estimation error in theirs slot,

hence they will participate in subsequent process of identification. CLBI also achieves the best time efficiency. The accuracy of three protocols is shown in Fig. 9(b). Due to the clustering error, some slots may be misclassified, i.e., 3 tags are estimated to be 2 tags. Therefore, the identification results will suffer from bias. The accuracy of CLBI (S3) is lower than CLBI(S2) because it is more difficult to successfully estimate more than three tags. The accuracy of A-CLBI is gradually higher than CLBI (S3) with the increase of B. The reason is that A-CLBI makes use of the singleton slot to identify blocked tags when $|B|/|N|$ is high, thus its overall accuracy is higher.

In Fig. 10, we evaluate the impact of the estimation error of $|B|$ by comparing the execution time of the proposed CLBI and A-CLBI protocols. Here we set $|N|$=10000 and $|B|$ changes from 100 to 3100. Note that the setting of the optimal value of f is influenced by the estimated accuracy of $|B|$, which further determines the execution time. It can be observed that the execution time of CLBI (perfect estimation) and A-CLBI (perfect estimation) is shorter than that of the imperfect ones. However, the execution time of the latter has only increased slightly. When the $|B|$ is 3000, CLBI (perfect estimation) and A-CLBI (perfect estimation) need to consume 9.32 s and 8.68 s. The execution time of CLBI (using CLE) and A-CLBI (using CLE) is 9.38 s and 8.78 s. The reason is because no matter $|B|/|N|$ is high or low, the estimation method in BLKI selects the maximum number of time slots among the singleton slot, 2-collision slots and 3-collision slots of the current frame, so as to increase the accuracy of the estimation. Moreover, CLBI and A-CLBI estimate the number of $|B|$ in each round rather than using other protocols. Therefore, the extra cost is reduced.

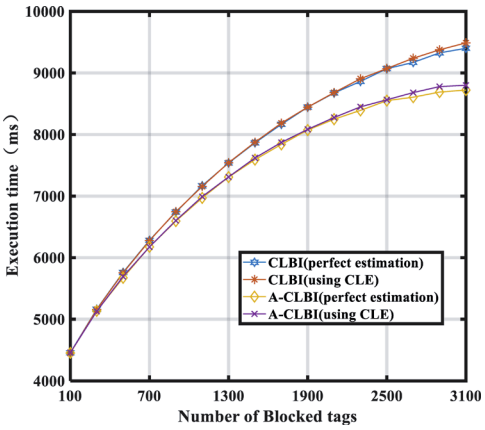

Figure 10: Impact of estimation error of |B|

5.5 Privacy protection of A-CLBI

In A-CLBI, each tag uses the hash function to select the slot to reply to the reader, hence the IDs or category IDs cannot be inferred. Moreover, the ID of each tag is not

directly transmitted in the wireless channel. Therefore, A-CLBI can protect the privacy of users compared with some existing protocols [Bu, Liu, Luo et al. (2013); Liu, Xie, Zhao et al. (2018)].

6 Conclusion

This paper investigates the problem of quickly and completely identifying the blocked tags with different cardinality, which is of great significance to the security of large-scale RFID systems. We first consider an RFID system with a small amount of blocked tags. Based on making full use of information in the physical layer, we propose a cross layer blocked tag identification protocol (CLBI). Different from previous works, CLBI estimates the tags in the physical layer, thereby utilizing collision slots to verify multiple tags simultaneously. Moreover, we propose an adaptive cross layer blocked tag identification protocol (A-CLBI) to apply to different proportions of blocked tags. We also propose a method to estimate the number of blocked tags without adding additional overhead. The extensive simulations have been conducted to demonstrate the superiority of the proposed protocols. According to the results, when the proportion of blocked tags is 0.01, our best protocol reduces the execution time by almost 50.5% , compared with the state-of-the-art.

Funding Statement: This work was supported in part by the National Natural Science Foundation of China under project contracts Nos. 61701082, 61701116, 61601093, 61971113 and 61901095, in part by National Key R & D Program under project Nos. 2018YFB1802102 and 2018AAA0103203, in part by Guangdong Provincial Research and Development Plan in Key Areas under project contract Nos. 2019B010141001 and 2019B010142001, in part by Sichuan Provincial Science and Technology Planning Program under project contracts Nos. 2018HH0034, 2019YFG0418, 2019YFG0120 and 2018JY0246, in part by the fundamental research funds for the Central Universities under project contract No. ZYGX2016J004, and in part by Science and Technology on Electronic Information Control Laboratory.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Angerer, C.; Langwieser, R.; Rupp, M. (2010): RFID reader receivers for physical layer collision recovery. *IEEE Transactions on Communications*, vol. 58, pp. 3526-3537.

Bolotnyy, L.; Robins, G. (2017): Physically unclonable function-based security and privacy in RFID systems. *IEEE International Conference on Pervasive Computing and Communications*, pp. 211-220.

Bu, K.; Liu, X.; Luo, J.; Xiao, B.; Wei, G. (2013): Unreconciled collisions uncover cloning attacks in anonymous RFID systems. *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 3, pp. 429-439.

Chen, H.; Wang, Z.; Xia, F.; Li, Y.; Shi, L. (2018): Efficiently and completely identifying missing key tags for anonymous RFID systems. *IEEE Internet of Things*

Journal, vol. 5, no. 4, pp. 2915-2926.

Chen, W. (2016): Optimal frame length analysis and an efficient anti-collision algorithm with early adjustment of frame length for RFID systems. *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3342-3348.

Chen, Y.; Feng, Q. (2009): Securing RFID systems by detecting tag cloning. *Pervasive Computing*, vol. 5538, pp. 291-308.

Chen, Y.; Feng, Q. (2019): A collision avoidance identification algorithm for mobile RFID device. *IEEE Transactions on Consumer Electronics*, vol. 65, no. 4, pp. 493-501.

Han, Y.; Zheng, W.; Wen, G.; Chu, C.; Su, J. et al. (2019): Multi-rate polling: Improve the performance of energy harvesting backscatter wireless networks. *Computers, Materials & Continua*, vol. 60, no. 2, pp. 795-812.

Huang, Z.; Xu, R.; Chu, C.; Li, Z.; Qiu, Y. et al. (2019): A novel cross layer anti-collision algorithm for slotted Aloha-based UHF RFID systems. *IEEE Access*, vol. 7, pp. 36207-36217.

Lehtonen, M.; Michahelles, F.; Fleisch, E. (2009): How to detect cloned tags in a reliable way from incomplete RFID traces. *IEEE International Conference on RFID*, pp. 257-264.

Li, T.; Chen, S.; Ling, Y. (2013): Efficient protocols for identifying the missing tags in a large RFID system. *IEEE/ACM Transaction on Networking*, vol. 21, no. 6, pp. 1974-1987.

Liu, X.; Li, K.; Jie, W.; Liu, A. X.; Xue, W. (2016): Top-k queries for multi-category RFID systems. *Proceeding of IEEE INFOCOM*, pp. 1-9.

Liu, X.; Li, K.; Min, G.; Shen, Y.; Liu, A. X. et al. (2015): Completely pinpointing the missing RFID tags in a time-efficient way. *IEEE Transaction on Communications*, vol. 64, no. 1, pp. 87-96.

Liu, X.; Li, K.; Min, G.; Shen, Y.; Qu, Y. (2014): A multiple hashing approach to complete identification of missing RFID tags. *IEEE Transaction on Communications*, vol. 62, no. 3, pp. 1046-1057.

Liu, X.; Qi, H.; Li, K.; Stojmenovic, I.; Liu, A. X. et al. (2015): Sampling bloom filter-based detection of unknown RFID tags. *IEEE Transaction on Communications*, vol. 63, no. 4, pp. 1432-1442.

Liu, X.; Xiao, B.; Li, K.; Liu, A. X.; Wu, J. et al. (2017): RFID estimation with blocker tags. *IEEE/ACM Transaction on Networking*, vol. 25, no. 1, pp. 224-237.

Liu, X.; Xie, X.; Zhao, X.; Wang, K.; Li, K. et al. (2018): Fast identification of blocked RFID tags. *IEEE Transaction on Mobile Computing*, vol. 17, pp. 2041-2054.

Luo, H.; Wen, G.; Su, J.; Huang, Z.; Inserra, D. (2019): Multi-hop distance-bounding for improving security and efficiency of ad-hoc networks. *IEEE Internet of Things Journal*, vol. 6, pp. 5312-5323.

Medhane, D. V.; Sangaiah, A. K.; Hossain, M. S.; Muhammad, G.; Wang, J. (2020): Blockchain-enabled distributed security framework for next generation IoT: an edge-cloud and software defined network integrated approach. *IEEE Internet of Things Journal*.

Ren, Y. J.; Liu, Y. P.; Ji, S.; Sangaiah, A. K.; Wang, J. (2018): Incentive mechanism

of data storage based on blockchain for wireless sensor networks. *Mobile Information Systems*, pp. 1-10.

Shahzad, M.; Liu, A. X. (2015): Expecting the unexpected: Fast and reliable detection of missing RFID tags in the wild. *Proceeding of IEEE INFOCOM*, pp. 1939-1947.

Su, J.; Chen. Y.; Sheng, Z.; Huang, Z.; Liu. A. (2020): From M-ary query to bit query: a new strategy for efficient large-scale RFID identification. *IEEE Transaction on Communications*, pp. 1-13.

Su, J.; Sheng, Z.; Leung, V. C. M.; Chen. Y. (2019): Energy efficient tag identification algorithms for RFID: survey, motivation and new design. *IEEE Wireless Communications*, vol. 26, no. 3, pp. 118-124.

Su, J.; Sheng, Z.; Liu, A.; Fu, Z.; Chen. Y. (2020): A time and energy saving based frame adjusting strategy (TES-FAS) tag identification algorithm for UHF-RFID systems. *IEEE Transaction on Wireless Communications*, pp. 1-13.

Su, J.; Sheng, Z.; Liu, A.; Han, Y.; Chen, Y. (2020): A group-based binary splitting algorithm for UHF RFID anti-collision systems. *IEEE Transaction on Communications*, vol. 68, no. 2, pp. 998-1012.

Trieu, D. B. K.; Maruyama, T. (2011): An implementation of the mean shift filter on FPGA. *International Conference on Field Programmable Logic and Applications*, pp. 219-224.

Vahedi, E.; Shah-Mansouri, V.; Wong, V. W. S.; Blake, I. F.; Ward, R. K. (2011): Probabilistic analysis of blocking attack in RFID systems. *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 803-817.

Wang, C.; Xie, L.; Wang, W.; Chen, Y.; Xue, T. et al. (2019): Probing into the physical layer: moving tag detection for large-scale RFID systems. *IEEE Transactions on Mobile Computing*, pp. 1.

Wang, F.; Xiao, B.; Bu, K.; Su, J. (2013): Detect and identify blocker tags in tree-based RFID systems. *IEEE International Conference on Communications*, pp. 2133-2137.

Wang, J.; Gao, Y.; Liu, W.; Sangaiah, A. K.; Kim, H. J. (2019): An intelligent data gathering schema with data fusion supported for mobile sink in wireless sensor networks. *International Journal of Distributed Sensor Networks*, vol. 15, no. 3.

Wang, J.; Gao, Y.; Yin, X.; Li, F.; Kim, H. J. (2018): An enhanced PEGASIS algorithm with mobile sink support for wireless sensor networks. *Wireless Communications and Mobile Computing*, pp. 1-9.

Yin, B.; Zhou, S. W.; Zhang, S. W.; Gu, K.; Yu, F. (2017): On efficient processing of continuous reverse skyline queries in wireless sensor networks. *KSII Transactions on Internet and Information Systems*, vol. 11, no. 4, pp. 1931-1953.

Yu, J.; Chen, L.; Wang, K. (2019): Finding needles in a haystack: Missing tag detection in large RFID systems. *IEEE Transaction on Communications*, vol. 65, no. 5, pp. 2036-2047.

Zanetti, D.; Fellmann, L.; Capkun, S. (2010): Privacy-preserving clone detection for RFID-enabled supply chains. *IEEE International Conference on RFID*, pp. 37-44.

Zhou, Z.; Chen, B; Yu, H. (2016): Understanding RFID counting protocols. *IEEE Transaction on Networking*, vol. 24, no. 1, pp. 312-327.