Theoretical Computer Science 788 (2019) 39–52

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Accelerate the classification statistics in RFID systems

Jiapeng Huang, Zhenzao Wen, Linghe Kong *, Li Ge, Min-You Wu, Guihai Chen

Shanghai Jiao Tong University, 200240, Shanghai, PR China

A R T I C L E   I N F O

Article history:
Received 6 June 2018
Received in revised form 19 September 2018
Accepted 30 November 2018
Available online 5 December 2018
Communicated by P. Spirakis

Keywords:
RFID
Classification
Statistics

A B S T R A C T

Radio Frequency Identification (RFID) technology has been widely used in many applications such as logistics, warehouse management and animal identification. However, the dilemma of short time requirement and massive tags makes traditional one-by-one identification methods impractical. Meanwhile, existing off-the-shelf methods cannot count and classify RFID tags at the same time. In this paper, RFID classification statistics problem is defined as classifying the tags into distinct groups and counting the quantity of tags in each group by the reader. The issue of time efficiency is significant in classification statistics, especially when the number of tags is large. To address this problem, we propose a novel Twins Accelerating Gears (TAG) approach. One gear shortens the classification process in frequency domain through subcarrier allocation, when another gear accelerates the statistics process in time domain through geometric distribution based quantity estimation. TAG can handle classification and quantity estimation during one process while existing methods need to handle it separately. We give elaborate proof of the running time and quantity estimation value of the process in theory. Typically, the total time of TAG is $O(\log N)$ and TAG outperforms existing identification solutions about 99.8% time reduction on 1000 tags classified statistics.

1. Introduction

Radio Frequency Identification (RFID) technology has been widely used in cardinality counting [1–4], identity recognition [5–7], transportation and logistics [8–10] and RFID tracking [11,12] recently. For example, RFID tags can be attached to books in the library [13] and can be tracked in case these books are misplaced or missing.

Typically, RFID reader and RFID tags are two constituent parts of a RFID system [14]. RFID reader is a wireless device to collect information from tags and RFID tag is a small identifiable device attached to objects. The manufacturer can set a unique ID number for each tag either using given types (such as SSCC, SGTIN, GSRN, and GID) or using custom types. According to the EPC standard [15], this number can be divided into several segments to identify objects location, category, serial number and so on. The value of a segment is called classification ID (CID). Hence, the reader can classify tags into different groups according to the CID.

Classification statistics is a common task in real RFID applications. For example, in Walmart, staffs need to keep enough commodities at the right place. It will make works more convenient and efficient if you offer a fast counting and classifying technique. Staffs can periodically use RFID readers examining the number of remaining commodities within the range of RFID signals instead of checking commodities one by one. Meanwhile, if commodities were accidentally misplaced by

* Corresponding author.
E-mail address: linghe.kong@sjtu.edu.cn (L. Kong).

https://doi.org/10.1016/j.tcs.2018.11.031
0304-3975/

customers, staffs can easily arrange them into the right categories, where traditional estimation methods become intolerably slow in this situation. To the best of our knowledge, related research is still vacant in the literature. Thus, we define the RFID classification statistics problem as to classify the tags into groups and to obtain the cardinality of tags in every group. In classification statistics problem, time efficiency is the most significant and challenging issue, especially the tags are always large [16,17]. Hence, in this paper, we formulate and study the fast classification statistics problem in RFID systems and propose a Twin Accelerating Gears (TAG) approach.

From system aspect, TAG runs as follows: (i) the reader broadcasts a query message including the information of sub-carrier allocation and time synchronization, then listens to the answers from tags. (ii) Each tag picks the time slot with the probability following the geometric distribution, and answers one-bit Yes only once at assigned subcarrier. (iii) The reader collects all answers from tags. Under the signal processing of the composite answers, the classification statistics result can be estimated according to geometric distribution.

Hence, TAG reduces the total processing time by following advantages. (i) In time domain, tags select the time slot following the geometric distribution, N tags can finish answering in a short time $O(\log N)$. Meanwhile, tags answer only one-bit 'Yes' instead of their long ID number. (ii) In frequency domain, different subcarriers are allocated to map the classification IDs. Hence, tags in different groups can be counted simultaneously at different subcarriers, i.e., the one-bit answers from tags need not transmit one-by-one.

We give the total running time and quantity estimation value in theory. Compared with ALOHA [18,19] and tree approaches [16,20], performance evaluations show TAG saves more than 99.8% time and reduces the error ratio about 30%, when 1000 tags are uniformly distributed in 4 groups.

The rest of this paper is organized as follows. In Section 2, the related work is presented. In Section 3, the problem is formulated. The design principles of classification accelerating gear and statistics accelerating gear are introduced in Section 4 and Section 5 respectively. In Section 6, TAG realization is proposed. In Section 7, simulation is performed for evaluating the solutions. In Section 8, we conclude this work.

2. Related work

Many identification algorithms are combined with different technology or scenarios, e.g., PIP [21] and acoustic tag identification [22]. The identification algorithms usually can be divided into two categories: ALOHA-based scheme [18,19] and Tree-based scheme [16,23]. In ALOHA-based scheme, the reader broadcasts a request message to tags nearby. Each tag randomly picks a time slot to transmit its ID number after receiving the message. For anti-collision, the reader has to keep sending requests until every tag is identified at least once. Multi-reader RFID systems [24–26] is more complicated on how to effectively handle Reader–Tag collisions and Reader–Reader collisions between adjacent readers. Tree-based scheme adopts a binary-tree structure to collect IDs of an unknown set. Adaptive Binary Splitting protocol (ABS) [20] is the typical tree-based protocol ABS method. By employing this scheme, the reader splits the set of tags into two subsets and labels them by binary numbers in each round. The reader repeats such process until each subset has only one tag. Although identification methods can meet the requirement of classification statistics, the longtime consumption is inevitable and intolerable so far in some scenarios.

In order to improve time efficiency of the counting process, many researchers studied cardinality estimation methods. The first tag estimation algorithm is Unified Simple Estimator (USE) [27]. USE estimates the number of tags without collecting their ID numbers but their answers in a given length of successive time slots. Meanwhile, Lottery Frame (LoF) [2] estimates the tag numbers by utilizing the collision information. LoF arranges the collision slots in an ordered pattern, thus providing the scalability while saving the processing time and communication overhead. Physical layer based cardinality estimator (PLACE) [28], a slot state detection algorithm, extracts more information and infer integer states from the same slots in RFID communications. Recent RFID estimation methods [1,3,29] keep improving the accuracy, time efficiency, energy consumption in cardinality counting process and achieve good performance. These methods mainly focus on quick quantity estimation, but they pay no attention on classification.

Two existing works [30,31] are close to this paper. [30] studies the problem of joint cardinality estimation: Given any two tag sets in a large RFID system, the authors can estimate their union cardinality, intersection cardinality, and difference cardinalities. [31] proposes an approach called simultaneous estimation for multi-category RFID systems (SEM). SEM exploits the Manchester-coding mechanism to decode the combined signals, thereby simultaneously obtaining the reply status of tags from each category. However, the execution time of SEM is in second level via introducing coding method, while our TAG is in microsecond level. Hence, we compare the proposed TAG algorithm with other existing algorithms, as shown in Table 1. ALOHA and ABS can finish classification statistics process, but cannot meet the requirement of "fast". USE and LoF sharply reduce the time of quantity estimation to $O(\log N)$, however, these methods cannot classify the tags. Our TAG can achieve the RFID classification statistics with short time cost. We prove that the total time of TAG is $O(\log N)$ in the Section 5.

3. Problem formulation

3.1. System model

The RFID system consists of one reader and N tags. The reader prior knows the semantic of the ID numbers and the N tags are in the communication range of the reader, so the reader can collect the information from tags by wireless

Table 1
Comparison of approaches.

Approach	Classification	Statistics	Total time
ALOHA	Yes	Yes	$O(N^2)$
ABS	Yes	Yes	$O(NlogN)$
USE	No	Yes	$O(logN)$
LoF	No	Yes	$O(logN)$
TAG	Yes	Yes	$O(logN)$

communication. Each tag contains a unique ID number of K-bit (usually $K = 96$ or 64 [15], but not limited). In these K bits, W bits ($W \leq K$) are selected to build up the segment of interest, whose value is treated as the classification criterion of the tags.

The set of all M CIDs is denoted by $C = \{C_1, C_2, \cdots, C_M\}$, where each value C_m is a classification ID and $m = 1, 2, 3, \cdots, M$. Since the segment of interest has W bits, it can present totally M different values which means

$$M = 2^W. \tag{1}$$

Different tags with the same CID are classified into the same group. We denote the number of tags in the group with the same CID C_m as $N_{\{C_m\}}$. Note that $N_{\{C_m\}}$ is an integer and $\sum_{m=1}^{M} N_{\{C_m\}} = N$.

For example, 12 tags are attached to 1 basketball, 5 footballs and 6 badmintons. Assuming 2 bits in the 96-bit ID number are selected to construct this category segment. Hence, there are 4 different values 00, 01, 10, and 11 mapping to basketball, volleyball, football, and badminton respectively. Obviously, tags with the same value 11 are classified into the badminton group. In this example, $N = 12$, $K = 96$, $W = 2$, $M = 4$, the set of CIDs is $C = \{00, 01, 10, 11\}$, $N_{\{00\}} = 1$, $N_{\{01\}} = 0$, $N_{\{10\}} = 5$ and $N_{\{11\}} = 6$.

3.2. Problem statement

Definition 1. (Problem: Classification Statistics) Given (i) N tags: N is a non-negative integer; (ii) a reader: it knows which W bits in ID number constructing the segment of interest. Then, the classification statistics problem is defined (i) to divide N tags into M groups according to CIDs; (ii) to obtain the quantity of every classified group $N_{\{C_1\}}, N_{\{C_2\}}, \cdots, N_{\{C_M\}}$.

We use two metrics, the average time cost T_{ave} and the error ratio ε, to measure the performance of a solution for RFID classification statistics. A solution with low T_{ave} and low ε is expected.

Definition 2. (Metric: Average Time Cost) Given N tags, and the total time cost T_{total} for N tags classification statistics, the average time cost T_{ave} is defined as the time cost per tag:

$$T_{\text{ave}} = \frac{T_{\text{total}}}{N}. \tag{2}$$

In (2), T_{total} can either be measured directly or be calculated from the reader side as:

$$T_{\text{total}} = (n_{\text{r}} \cdot t_{\text{r}} + n_{\text{w}} \cdot t_{\text{w}}) \cdot t_{\mu}, \tag{3}$$

where t_{r} and t_{w} are the time slots for reading the tags one turn and the time slots for waiting the idle between twice readings respectively; n_{r} and n_{w} are the number of turns to read and the number of times to wait respectively in a classification statistics process; and t_{μ} is the time unit of every time slot.

Definition 3. (Metric: Error Ratio) Given the number of tags N, and the real quantity of tags in each classified group $N_{\{C_m\}}$, $m = 1, 2, ..., M$, the error ratio ε is computed as:

$$\varepsilon = \sum_{m=1}^{M} \frac{\left| N_{\{C_m\}} - \tilde{N}_{\{C_m\}} \right|}{N}, \tag{4}$$

where $\tilde{N}_{\{C_m\}}$ is the statistics number of tags in the classified group with CID C_m. This error ratio measures the classification error and statistics error by a unified metric in one equation.

4. Accelerating gear I: classification subcarrier allocation

The classification statistics is speeded up by Twin Accelerating Gears. The first gear considers only classifying the tags. It accelerates the classification process by subcarrier allocation.

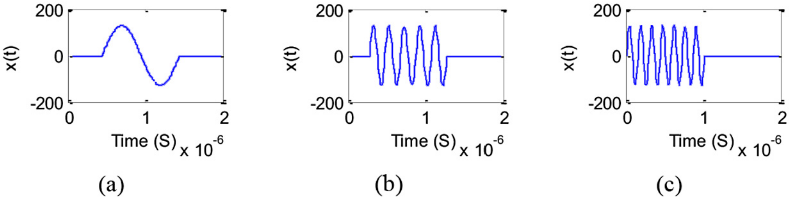

Fig. 1. (a) One basketball tag's answer signal. (b) One football tag's answer signal. (c) One badminton tag's answer signal.

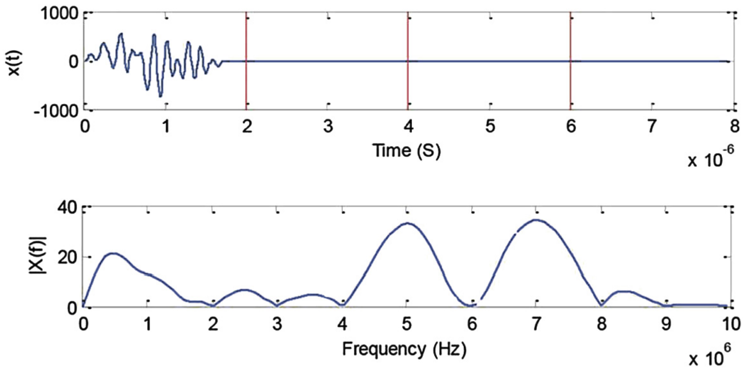

Fig. 2. The composite signal of 12 answers received by the reader with period 4 time slots and the decomposed sub-signal in frequency domain.

4.1. Classification design overview

Assume a channel has F available subcarriers for RFID communication, denoted by a set $S = \{S_1, S_2, ...S_F\}$. Given a set $C = \{C_1, C_2, ...C_M\}$ having M potential CIDs. We only consider the situation when $F \geq M$ here, the situation $F < M$ will be discussed in Section 4.3.

Firstly, select M subcarriers from S since $F \geq M$ and create a bijective function $f_b(C_m): C \to S$ between M CIDs and M subcarriers. This mapping relationship is called subcarrier allocation which maps every CID to exactly one subcarrier.

Secondly, the system runs as follows to classify tags by subcarrier allocation:

Step 1. RFID reader broadcasts a message including the subcarrier allocation information;
Step 2. After receiving the broadcast message, every tag immediately answers one-bit Yes once in the assigned subcarrier, which matches its own CID;
Step 3. The reader receives the composite signal of the answers from all tags. Performing this signal by Fast Fourier Transform (FFT), the frequency domain representation of the signal is obtained, which also provides the classification result.

Thirdly, we provide an example to explain this process. We use the same setting of the ball example aforementioned in Section 3.1. In addition, a baseband channel with 10 MHz bandwidth is given. Set $F = 5$, so the center frequency of these 5 subcarriers S_1, S_2, S_3, S_4 and S_5 are $1, 3, 5, 7,$ and 9 MHz. According to Step 1, the reader broadcasts the massage with the information about assigning the CID "00", "01", "10", and "11" to subcarrier S_1, S_2, S_3 and S_4, S_5 is null due to $F > M$. Assume S_1, S_2, S_3 and S_4 have been allocated to basketball, volleyball, football and badminton tags respectively. In Step 2, the basketball tag is assigned S_1, so it answers a one-bit "Yes". This "Yes" is modulated as a 1 MHz sine wave, whose period is 1 µs and amplitude is $-127 \sim 127$ unit; Each football and badminton tag is similar. The relationship is shown as (5):

$$
\begin{array}{l}
S_1 = 1\text{ MHz} \Leftrightarrow CID\text{“00”} \Leftrightarrow basketball \\
S_2 = 3\text{ MHz} \Leftrightarrow CID\text{“01”} \Leftrightarrow volleyball \\
S_3 = 5\text{ MHz} \Leftrightarrow CID\text{“10”} \Leftrightarrow football \\
S_4 = 7\text{ MHz} \Leftrightarrow CID\text{“11”} \Leftrightarrow badminton
\end{array}
\tag{5}
$$

Due to the different distances from the reader to tags, the answers arrive at the reader asynchronously. In Fig. 1, (a), (b) and (c) show the "Yes" signals of only one basketball, one football, and one badminton tag received by the reader; RFID reader gets the composite signal of 12 answers as shown in Fig. 2. We set the time slot in this example as 2 µs since the signal is completed received in 2 µs. It is impossible to decode a composite signal directly in time domain. However, according to Step 3, after doing the FFT on this signal in the 2 µs, only the subcarrier S_1, S_3 and S_4 have the obvious frequency components shown in Fig. 2. Hence, the classification result is got by the reader: these tags can be classified into

3 groups according to different ball categories, which are basketball group with CID “00”, football group with CID “10” and badminton group with CID “11”.

4.2. Classification time consumption comparison

We quantize the average time cost of our fast classification method and traditional methods by theoretical derivation:

Ideal identification method: According to (3), we calculate the total time cost of ideal identification method. Since the reader should read N tags one-by-one, $n_r = N$; each tag is K-bit, $t_r = K$; Assume the ideal case needs no time for waiting or anti-collision, $t_w = 0$; Then

$$T_{total} = (N \times K + 0) \times t_\mu = NKt_\mu. \tag{6}$$

Substituting (6) to (2), we get

$$T_{ave} = Kt_\mu. \tag{7}$$

Ideal tree-based method: The reader also needs read N tags, $n_r = N$; tree-based method adds $\log_2 N$ bit prefix to every tag for forming a binary tree, so $t_r = K + \log_2 N$; In ideal case, we also assume there is no waiting time, $t_{W=0}$. So

$$T_{total} = (N \times (K + \log_2 N) + 0) \times t_\mu = N(K + \log_2 N)t_\mu. \tag{8}$$

And then,

$$T_{ave} = (K + \log_2 N)t_\mu. \tag{9}$$

Our Method: Since all answers can be read at the same turn, $n_r = 1$; In addition, the channel is divided into F subcarriers, the transmission rate becomes $1/F$, then every bit need F time slots to be transmitted. Each Yes answer is one bit, so $t_r = F$; There is no wait time in our method, $t_W = 0$. We get

$$T_{total} = (1 \times F + 0) \times t_\mu = Ft_\mu. \tag{10}$$

Moreover,

$$T_{ave} = \frac{Ft_\mu}{N}. \tag{11}$$

For comparing the cost time between our method and ideal identification method, we measure the ratio of (11) and (7)

$$\text{ratio}(T_{ave}) = \frac{F}{NK} \times 100\%. \tag{12}$$

If F and K are given constants, (12) is $O(1/N)$. Thus, our solution dramatically reduces the classification time. In the ball example aforementioned, $N = 12$, $F = 5$ and $K = 96$. Hence, the ratio result is 0.43%. Even in the case of such a small scale, our method needs only 0.43% classification time compared with the optimal identification method. Note that (7) < (9). Thereby, in ideal case, both our method and identification are better than tree-based method on classification time.

4.3. Impact: the number of subcarriers

Since the number of subcarriers F depends on the physical feature of wireless channel and the limitation of RFID devices, F is usually a finite number. However, the number of potential CIDs M only relies on the length definition of classification IDs. It is possible that $F < M$. In order to break the bottleneck of finite number of F, we repeat the classification process $n_r = \lceil M/F \rceil$ turns. In every turn, F different CIDs are allocated, except the rest CIDs in last turn.

Intuitively, the number of subcarriers influences the performance of our method much. When $F < M$, there is no enough subcarriers allocated to CIDs and when $F \geq M$, there are redundant signals. The two situations may improve or reduce the running time. However, we prove that the running time of our method is independent of the number of subcarriers in ideal case in Theorem 1.

Theorem 1. *Given M, N and t_μ are fixed numbers and F is a variable. Using subcarrier allocation to classify N tags, the average running time is determined by the number of CIDs M.*

Proof. The proof of the situation $F \geq M$ is simple. According to (11),

$$T_{ave} = \frac{Ft_\mu}{N} \geq \frac{Mt_\mu}{N}. \tag{13}$$

Hence, we get the minimum value when $F = M$.

When $F < M$, we can get identical relation $M = s \cdot F + l$, where $s = \lfloor M/F \rfloor$ and $l = M - s \cdot F$. Note that we only need allocate l subcarriers in the last turn. Hence, according to (11), we can get

$$T_{\text{ave}} = \lfloor M/F \rfloor \cdot \frac{Ft_\mu}{N} + l \cdot \frac{t_\mu}{N} = \frac{(s \cdot F + l)t_\mu}{N} = \frac{Mt_\mu}{N}. \tag{14}$$

According to (13) and (14), the minimum average running time is determined by the number of CIDs M. □

Note that Theorem 1 assuming that M is known before experiments. Theorem 1 tells us the classification results can keep accuracy but the total time consumption will not change for the lack of subcarriers. However, in practice, the interval time between each turn may not be exact "0". If condition permits, we better set $F = M$ or appropriate value to reduce the repeat times. Hence, the adjustable F is set the same value of M in our method.

5. Accelerating gear II: geometric distribution based quantity estimation

Only classification is not enough in some applications, the second gear is introduced in this section. It accelerates the statistics by quantity estimation based on geometric distribution.

5.1. Statistics design overview

Firstly, we introduce some concepts in our solution. (i) 1/2 geometric distribution (GD): In this fast statistics method, the answer period is divided into T time slots. Each tag answers by selecting one time slot following the 1/2 GD. i.e., 1/2 probability to select the 1st time slot, 1/4 probability to select 2nd time slot,..., $(1/2)^T$ to select the T-th time slot. (ii) Time synchronization: The reader broadcasts the time synchronization flag, so that all tags know the beginning of any T-th time slot. iii) Signal decomposition: At any time slot, the part of the composite signal can be decomposed into M sub-signals through Band Pass Filter (BPF) for M non-null subcarriers. Decoding this part of sub-signal, there are three possible results: collision due to multi-answer; only one answer; or no answer. (iv) Bitmap: the bitmap B is a $M \times T$ matrix. The M rows distinguish the M non-null subcarriers in frequency domain, i.e. CIDs, and the T columns present the T time slots in time domain. The value in an element presents the answer states in a certain subcarrier in a certain time slot. The value can be "1" or "0" indicating the status of collision and one answer or no answer.

Secondly, in order to reduce the classification statistics time cost, the twin gears work together at the same time. Consequently, the three Steps are extended as follows:

STEP 1. The RFID reader broadcasts a message including the subcarrier allocation and time synchronization information;
STEP 2. After receiving the broadcast message, all tags are synchronized. Every tag selects a time slot following the 1/2 geometric distribution and answers one bit Yes once in the assigned subcarrier, which matches its own CID;
STEP 3. The reader receives the composite signal of the answers from all tags. At any time slot, the signal is decomposed into sub-signals of every non-null subcarrier through BPFs. The bitmap is created based on the decomposition results. The statistics result can be got by quantity estimation of the bitmap.

Thirdly, we continue to use the ball example for this process explanation. Due to 1/2 GD, assume that the basketball tag selects the 2nd time slot to answer in S_1; 3 football tags select the 1st time slot, 1 football tag selects the 2nd time slot, and 1 football tag selects the 3rd time slot to answer in S_3; 3 badminton tags select the 1st time slot, 2 badminton tags selects the 2nd time slot, and the final one selects the 3rd time slot to answer in S_4. The composite signal with all answers is received by the reader as shown in Fig. 3(a). This signal is decomposed by four BPFs. The center frequencies of these four BPFs are 1, 3, 5, and 7 MHz and their bandwidths are all 2 MHz. The decomposed sub-signals are shown in Fig. 3, (b), (c) and (d). We use the sub-signal in S_4 as an example to analyze the answer states. In Fig. 3(d), the sub-signals in 1st and 2nd time slots cannot be decoded due to irregular wave. It is considered the signals overlapping by collisions. So we set the states of these two time slots as "1"; in the 3rd time slots, the regular successive 7 sine waves can be found, which means the only one "Yes" answer. We set "1" for this state; there is no wave in the 4-th time slot, so the state is 0. According to the analysis of these sub-signals, the bitmap $B_{M\times T}$ can be built up:

$$B_{M\times T} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}.$$

Using the quantity estimation method (see Section 5.2) to analyze every row of $B_{M\times T}$, we can estimate 1 answer in S_1, 10 answers in S_3, and 10 answers in S_4. i.e., there are 1 basketball, 10 football and 10 badminton tags. Hence, the classification statistics result is achieved although the result here is not accurate.

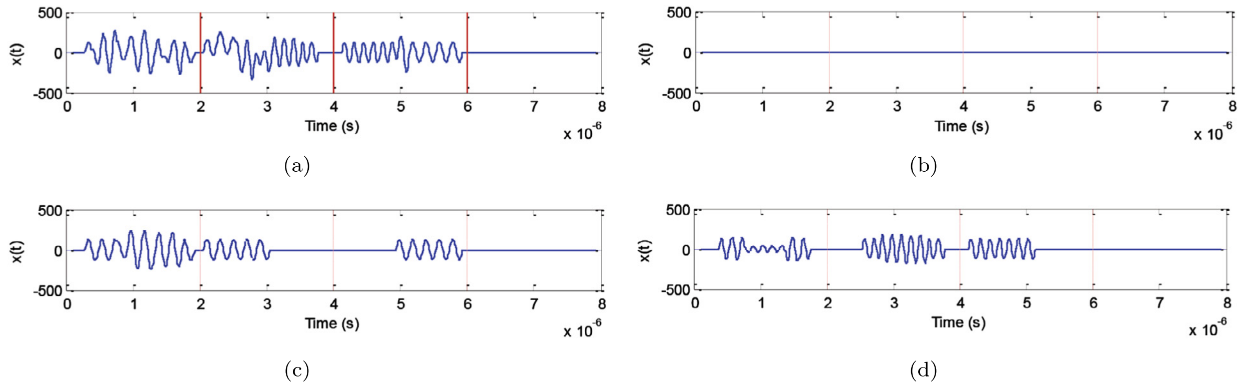

Fig. 3. (a) The composite signal received by the reader with period 4 time slots. (b) The decomposited sub-signal in subcarrier S_2. (c) The decomposited sub-signal in subcarrier S_3. (d) The decomposited sub-signal in subcarrier S_4.

5.2. Quantity estimation

In a bitmap, we only consider the rows having non-zero data. Each of these rows presents an existing classification. According to Definition 1, the goal of RFID classification statistics can be represented to obtain the quantity of tags in each existing classification. Hence, we design the quantity estimation method to apply on each of these rows.

For further analysis, we briefly introduce the notations here. The total number of "1" before the first "0" is denoted by N_m in a row corresponding to the CID: C_m. Then the total number of "1" after the first "0" in a row is denoted by N_m^0. $p_{n,k}$ and $q_{n,k}$ represent,

$$p_{n,k} = Pr(N_m = k), \quad q_{n,k} = Pr(N_m \geq k)$$

where $Pr(A)$ means the probability of an event A, n is the total number of tags.

Based on the analysis in [32], we get the following Lemma 2 and Lemma 3. Lemma 2 gives the exact equation of $q_{n,k}$ while Lemma 3 derives the exception and standard deviation of N_m.

Lemma 2. *The probability distribution of N_m is characterized by:*

$$q_{n,k} = \sum_{j=0}^{2^k} (-1)^{\nu(j)} (1 - \frac{j}{2^k})^n, \tag{15}$$

where $\nu(n)$ denotes the number of ones in the binary representation of n. e.g., $\nu(13) = \nu((1101)_2) = 3$.

Lemma 3. *Given N_m and N_m^0, the expected value of N_m satisfies:*

$$E[N_m] = \log_2\left(\varphi \cdot N_{\{C_m\}}\right) + P(\mu) + o(1). \tag{16}$$

And the standard deviation σ_X of N_m satisfies:

$$\sigma_X^2 = \sigma_c^2 + Q(\mu) + o(1), \tag{17}$$

where the constant $\varphi = 0.77351\ldots$ and $\sigma_c = 1.12127\ldots$. $P(\mu)$ and $Q(\mu)$ are two periodic functions of μ with mean value 0, period 1 and amplitude bounded by 10^{-5} respectively, $\mu = \log_2(N_{\{C_m\}})$.

Lemma 2 and Lemma 3 were proved in [32]. Omitting the term $P(\mu) + O(1)$, $N_{\{C_m\}}$ can be estimated by N_m and N_m^0, which are easily to be got from the bitmap.

Definition 4. (Method: Quantity Estimation) Given N_m and N_m^0, $\tilde{N}_{\{C_m\}}$ is an estimator of $N_{\{C_m\}}$, we have

$$\tilde{N}_{\{C_m\}} = \begin{cases} \left\lfloor \frac{1}{\varphi} \times 2^{N_m} \right\rfloor + N_m^0 & N_m \geq 1 \\ N_m^0 & N_m = 0 \end{cases}. \tag{18}$$

We provide some examples in Table 2 to show the quantity estimation results according to (18).

Table 2
Quantity estimation examples.

Row data in the bitmap	N_m	N_{m0}	$\tilde{N}_{\{C_m\}}$
0100	0	1	1
1110	3	0	10
1111 1111 1011	9	2	663
1111 1111 1110 111	11	3	2650

Theorem 4. *A bitmap with $\frac{3}{2}\log_2 N$ time slots is sufficient for the quantity estimation method using 1/2 geometric distribution answers, where N is the total number of all tags.*

Proof. Considering the case when $k = \frac{3}{2}\log_2(N) + \delta$, with $\delta \geq 0$. Hence, the value 1 before the first 0 must appear at least once. Hence,

$$Pr(N_m \geq k) = 1 - Pr(N_m < k) \leq 1 - (1 - \frac{1}{2^k})^N < 1 - \exp(-\frac{N}{2^k}). \tag{19}$$

In the range of values of k considered, $2^k = 2^\delta \cdot N^{\frac{3}{2}}$. Hence, the last expression is the order of $O(\frac{2^{-\delta}}{\sqrt{N}})$.

Denote the tail expectation begins at k as $E_{tail,k}$. Firstly, $E[N_m]$ has been proved existed in [32]. Then, we get the equation:

$$\begin{aligned} E[N_m] &= \sum_{k=1}^{\infty} k \cdot p_{N,k} = \sum_{k=1}^{\infty}(k-1)\cdot p_{N,k} + \sum_{k=1}^{\infty} p_{N,k} \\ &= \sum_{k=1}^{\infty} k \cdot p_{N,k+1} + q_{N,1} = \cdots = \sum_{k=1}^{\infty} q_{N,k} \end{aligned}. \tag{20}$$

Equation (20) implies the convergence of the sequence $\{q_{N,k}\}_0^\infty$, which ensures the correctness of the equation (21):

$$\begin{aligned} E_{tail,k_0} &= \sum_{k=k_0}^{\infty} k \cdot p_{N,k} = \sum_{k=k_0}^{\infty} k \cdot (q_{N,k} - q_{N,k+1}) \\ &= \sum_{k=k_0}^{\infty} k \cdot q_{N,k} - \sum_{k=k_0}^{\infty}(k+1)\cdot q_{N,k+1} + \sum_{k=k_0}^{\infty} q_{N,k+1} \\ &= \sum_{k=k_0}^{\infty} k \cdot q_{N,k} - \sum_{k=k_0+1}^{\infty} k \cdot q_{N,k} + \sum_{k=k_0+1}^{\infty} q_{N,k} \\ &= k_0 \cdot q_{N,k_0} + \sum_{k=k_0+1}^{\infty} q_{N,k} \end{aligned}. \tag{21}$$

Hence, (21) gives us an exact equation of tail exception. Considering the start point $k_0 = \lceil \frac{3}{2}\log_2(N) \rceil = \frac{3}{2}\log_2(N) + \delta_0$. Then,

$$E_{tail,k_0} = O(\frac{\frac{3}{2}\cdot \log_2 N \cdot 2^{-\delta_0}}{\sqrt{N}}) + O(\sum_{k=k_0+1}^{\infty} \frac{2^{-\delta}}{\sqrt{N}}) = O(\frac{\log_2 N}{\sqrt{N}}). \tag{22}$$

The value of E_{tail,k_0} drops faster since the probability $q_{N,k}$ is smaller than we prove. Equation (22) implies that we can ignore time slot after k_0. Therefore, $\left\lceil \frac{3}{2}\log_2 N \right\rceil$ time slots are sufficient for estimation. □

Theorem 4 shows that we can ignore the value bigger than $\left\lceil \frac{3}{2}\log_2 N \right\rceil$ even it appears in one experiment. Hence, using 1/2 GD, the proposed method can estimate the quantity in a short time. In practice, we can shorten $\left\lceil \frac{3}{2}\log_2 N \right\rceil$ into $\lceil \log_2 N \rceil + C_0$, where C_0 is a constant. For the reason that, according to Chebyshev's Inequality, no more than $1/C_0^2$ of the distribution's values can be more than C_0 standard deviations away from the mean. For example, if $C_0 = 10$, the probability approximately 1.3% which can be ignored in one time. With the growth of the total number of tags, we can take three-sigma rule, i.e., 68-95-99.7 rule.

We have calculated T_{ave} of only classification process in (14). Considering the classification and statistics process together, we re-calculate T_{ave} of the completed TAG. Compared with (14), n_{r} has no change, $n_{\text{r}} = \lceil M/F \rceil$. However, from Theorem 4, we know that $\left\lceil \frac{3}{2}\log_2 N \right\rceil$ time slots are demanded for quantity estimation, so $t_{\mathrm{r}} = \left(\left\lceil \frac{3}{2}\log_2 N \right\rceil\right) \cdot F$. Given N and t_μ, we extend (14) and get

$$T_{\mathrm{ave}} = Mt_\mu \times \left(\frac{\left\lceil \frac{3}{2}\log_2 N \right\rceil}{N}\right). \tag{23}$$

Tradeoff between time and accuracy: Although an error with approximately 1.12 can be acceptable for some applications, it is too high for some other applications. However, it is obvious that the proposed quantity estimation method is asymptotically unbiased. (The similar estimator has been proven to be unbiased in [32].) It means, if we make multiple independent estimations and compute the average result, the standard deviation will be significantly reduced.

If time allows, TAG can be repeated R rounds to reduce the error, i.e., after each round, the reader re-broadcasts an 8-Byte message for restarting TAG. In the r-th round, N_m and N_m^0 are denoted by $N_{m,r}$ and $N_{m,r}^0$ respectively, where $1 \le r \le R$. e.g., $\overline{N_m} = (1/R)\sum_{r=1}^{R} N_{m,r}$, $\overline{N_m^0} = (1/R)\sum_{r=1}^{R} N_{m,r}^0$. Thus, we rewrite (18) and get the quantity after R round estimation as

$$\tilde{N}_{\{C_m\}} = \begin{cases} \left\lfloor \frac{1}{\varphi} \times 2^{\overline{N_m}} + \overline{N_m^0} \right\rfloor & \forall N_{m,r} \ge 1 \\ N_{m,r}^0 & \exists N_{m,r} = 0 \end{cases}. \tag{24}$$

And the standard deviation after R round estimation is

$$\overline{\sigma} = \frac{\sigma_c}{\sqrt{R}}. \tag{25}$$

Let α be the error probability. We define that TAG is considered to achieve the accuracy requirement when $Pr\left(-\beta_1 N_{\{C_m\}} \le \tilde{N}_{\{C_m\}} - N_{\{C_m\}} \le \beta_2 N_{\{C_m\}}\right) \ge 1 - \alpha$.

Theorem 5. *Given α, β_1 and β_2, the defined accuracy requirement can be achieved if repeat R rounds TAG,*

$$R \ge max\left(\left[\frac{-\sigma_c\lambda}{log_2(1-\beta_1)}\right]^2, \left[\frac{\sigma_c\lambda}{log_2(1+\beta_2)}\right]^2\right), \tag{26}$$

where λ is obtained by solving $1-\alpha = erf\left(\lambda/\sqrt{2}\right)$, $erf(\cdot)$ is the Gaussian error function. Note that $0 < \beta_1 < 1$, $\beta_2 > 0$.

Proof. For convenience, we denote the exception of N_m as μ_0.

$$\begin{aligned} & Pr\left(-\beta_1 N_{\{C_m\}} \le \tilde{N}_{\{C_m\}} - N_{\{C_m\}} \le \beta_2 N_{\{C_m\}}\right) \ge 1-\alpha \\ \Longleftrightarrow & Pr\left(\frac{1-\beta_1}{\varphi} \cdot 2^{\mu_0} \le \frac{1}{\varphi} \cdot 2^{\overline{N_m}} \le \frac{1+\beta_2}{\varphi} \cdot 2^{\mu_0}\right) \ge 1-\alpha \\ \Longleftrightarrow & Pr\left(\frac{\log_2(1-\beta_1)}{\sigma_c/\sqrt{R}} \le \frac{\overline{N_m}-\mu_0}{\sigma_c/\sqrt{R}} \le \frac{\log_2(1+\beta_2)}{\sigma_c/\sqrt{R}}\right) \ge 1-\alpha \end{aligned}. \tag{27}$$

According to Lindeberg central limit theorem, the distribution of the standardized sums of $N_{m,r}$ converges towards the standard normal distribution $N(0,1)$. When $1-\alpha = erf\left(\lambda/\sqrt{2}\right)$, we get

$$\frac{\log_2(1-\beta_1)}{\sigma_c/\sqrt{R}} \le -\lambda, \quad \frac{\log_2(1+\beta_2)}{\sigma_c/\sqrt{R}} \ge \lambda. \tag{28}$$

Solve inequality (28), we get the conclusion. □

Notice the fact that $0 < \log_2(1+\beta) < -\log_2(1-\beta)$. Hence, if $\beta_1 = \beta_2 = \beta$, then (26) can be rewritten as:

$$R \ge \left(\frac{\sigma_c\lambda}{log_2(1+\beta)}\right)^2. \tag{29}$$

Inequality (26) implies that it is more difficult to control the upper bound than the lower bound. The nature of inequality (26) is that normally $E(2^x) \ne 2^{E(x)}$ since 2^x belongs to convex functions. Hence, in practice, we take the error ratio between $\left[-\frac{1}{3}, \frac{1}{2}\right]$, i.e., $\beta_1 = \frac{1}{3}$ and $\beta_2 = \frac{1}{2}$. Assuming that $1-\alpha = 0.90$, then $\lambda = 1.645$. Then we get $R \ge 9.950$. Hence, we repeat our process 10 times.

5.3. Adaptive estimation time

Most cardinality estimation methods [27,33] in RFID systems use fixed length of time slots T in bitmap. These methods require prior knowledge of the approximate number of tags N', where $O(N') = O(N)$, for deciding a length of time slots $T = f_T(N')$. Otherwise, without the prior knowledge of N', these methods lead to either time waste when $T >> f_T(N')$ or low accuracy when $T << f_T(N')$.

The proposed TAG can adapt the length of T without any prior knowledge of N'. According to 1/2 GD of the answers, in Theorem 4, we have proved when $T > \frac{3}{2}\log_2 N$, the value in those time slots will be always "0". Taking advantage of this feature, we design the Automatic Stop Flag (ASF) method to control the adaptive T.

ASF method set a stop flag by appearance of successive j – "0". In TAG, ASF runs respectively for every row in the bitmap. When all rows have the ASFs, the TAG process finishes automatically. e.g., if we adopt successive 3-0 as the ASF, when all rows have occurred "000", we consider that all N tags have answered. The TAG process is stopped automatically. Thus, the adaptive estimation time is achieved.

6. Twin accelerating gears realization

Base on the above analysis and theoretical derivation, we develop the TAG algorithm. TAG algorithm is divided into three parts: at tag side, at reader side for answer collection, at reader side for classification statistics respectively.

TAG algorithm running on tag side is simple as shown in Algorithm 1. After decoding the message from the reader, a tag can get the subcarrier allocation information $f_b(\cdot)$ and the synchronization information T_0. Substituting its own classification ID C_m into $f_b(\cdot)$, the tag get the assigned subcarrier S_f, and then, it selects a time slot τ by 1/2 geometric distribution. Finally, the tag transmits its answer at subcarrier S_f and time $\tau + T_0$. Algorithm 2 provides the pseudo code of TAG algorithm at reader side for bitmap construction. Above all, the reader broadcasts a message including the given $f_b(\cdot)$ and T_0. Then, from T_0, it begins to build up a bitmap $B_{M\times T}$ by answer collection. The value of element $B(m,\tau)$ is set "1" if answer collision and one answer in subcarrier S_f and time $\tau + T_0$; or "0" if no answer. In $B_{M\times T}$, the number of column T depends on the ASFs. In every time slot, each row is checked whether it has an ASF. When ASFs appear in all rows (the number of ASFs is M), the answer collection process is finished and $B_{M\times T}$ is got. For simplicity, Algorithm 2 only presents the one turn situation. It is enough when the cases of $F \geq M$. When $F < M$, this algorithm should be repeated $\lceil M/F \rceil$ turns.

In Algorithm 3, the classification statistics part of TAG algorithm at reader side is illustrated. First, this algorithm counts the number of "1" in every row. Then, the result of quantity estimation is got according to (18) and is stored in a column $\tilde{N}_{M\times 1}$. The value of every element $\tilde{N}(m,1)$ in $\tilde{N}_{M\times 1}$ is the estimated quantity of tags in the classification C_m.

If repeating R rounds of Algorithm 1, 2, and 3 and estimating the quantity as (24), we can obtain a more accurate result but the process time is prolonged R times.

Algorithm 1 Tag side.

Input: Message from the RFID reader. $f_b(\cdot): C \rightarrow S$: subcarrier allocation information, which is a bijective function; T_0: time synchronization information; $f_g d(1/2)$: a function to select a time slot following 1/2 geometric distribution;
Output: One-bit answer;
Procedure
  while TRUE **do**
    wait_message();
    if *wait_message*() == 1 **then**
      decode_message$(f_b(\cdot), T_0)$; ▷ get $f_b(\cdot)$ and T_0 from the message
      $S_f \leftarrow f_b(C_m)$; ▷ set the subcarrier according to own CID
      $\tau \leftarrow f_g d(1/2) + T_0$; ▷ transmit one-bit answer in S_f in τ
    end if
  end while
end Procedure

7. Performance evaluation

7.1. Experimental methodology and setting

We use Matlab to implement the simulation. The default parameters are set as follows: the total number of tags $N = 1000$; the number of classifications $M = 4$; the number of subcarriers $F = 5$; the length of ID number $K = 96$; the number of repeated round $R = 1$; Flag ASF = "00000".

The performance depends on the distribution of the tags in the classifications. We select two extreme situations: uniform and Max-1-0 distribution to test the efficiency of TAG.

1. **Uniform Distribution (UD):** The quantity of tags in every classified group is nearly the same. Hence, each group has $(\lfloor N/M \rfloor + 1)$ or $(\lfloor N/M \rfloor)$ tags.

Algorithm 2 Reader side for bitmap construction.

Input: $f_b(\cdot)$; T_0; ASF: a flag of "0" serial with given length, N_{ASF} is the number of ASF;
Output: $B_{M\times T}$: the bitmap of all answers in non-null subcarriers.

Procedure
  while TRUE **do**
    $broadcast\_message(f_b(\cdot), T_0)$;
    $\tau \leftarrow 1$;
    while $N_{\text{ASF}} < M$ **do** ▷ receive answers until all rows having ASFs
      $N_{\text{ASF}} \leftarrow 0$;
      for m = 1 to M **do**
        $B(m,\tau) \leftarrow decode\_answer(S_f, (\tau + T_0))$; ▷ Build up $B_{M\times T}$
        if $check\_ASF(B(m,)) == 1$ **then** ▷ check whether a row has ASF
          $N_{\text{ASF}} \leftarrow N_{\text{ASF}} + 1$;
        end if
      end for
      $\tau \leftarrow \tau + 1$;
    end while
  end while
end Procedure

Algorithm 3 Reader side for classification statistics.

Input: $B_{M\times T}$;
Output: $\tilde{N}_{M\times 1}$: a column vector storing estimation results;

Procedure
  while TRUE **do**
    for m = 1 to M **do**
      $N_1 \leftarrow count\_1(B(m,))$; ▷ count "1" before the first "0" in a row
      $N_0 \leftarrow count\_0(B(m,))$; ▷ count "1" after the first "0" in a row
      if $N_1 == 0$ **then**
        $\tilde{N}(m,1) \leftarrow N_0$;
      else
        $\tilde{N}(m,1) \leftarrow (1/\phi)2^{N_1} + N_0$;
      end if
    end for
  end while
end Procedure

2. **Max-1-0 Distribution (M10D):** One group has the maximal number of tags, another group has only 1 tag, and the other groups have no tag. e.g., one book is put wrong shelf in the library; one product is moved to another area in the supermarket M10D exists in these examples.

TAG is compared with ALOHA, ABS, USE, LoF and TAG10. Note that the TAG10 represents TAG with $R = 10$ rounds. In addition, USE and LoF cannot classify tags actually. For approximating, we assume that they can estimate the quantity of tags group-by-group. When any group is finished, the reader broadcasts an 8 Bytes message including the synchronization information and the next group CIDs.

7.2. Performance analysis

Varying Number of Tags: The maximum number of tags for large-scale RFID estimation in the literature [2,16,17] are mostly from $10^2 \sim 10^4$. Moreover, tags number at 10^4 level is sufficient in many scenarios such as supermarket inventory management. Hence, we first evaluate TAG and other approaches by varying the number of tags N from 1 to 10000. Note that Fig. 4 and Fig. 5 share the same legend with subfigure 4(a).

The log-scale graph Fig. 4 (a) presents the performance of total time cost (Definition 2) against N varying in UD, and Fig. 4 (b) plots it in M10D. We find that (i) TAG achieves the least time among all in both distributions. When $N = 1000$, TAG costs 187 us. Compared with 530 ms of ALOHA or 144 ms of ABS, TAG spends $\leq 0.02\%$ time of existing approaches to finish classification statistics; (ii) TAG, USE and LoF are in the same order, and they use much less time than ALOHA, ABS. Such results confirm the comparison in Table 2; (iii) Although in a same order, for a given N, USE and LoF need more time than TAG. Furthermore, TAG10 demands more time than USE and LoF. e.g., when $N = 10000$, TAG10 costs 2 ms, USE costs 289 us, LoF costs 292 us, and TAG costs only 195 us. This result implies that the parallel processing is faster than the serial one; (iv) TAG10 costs more time in M10D than UD. The reason is that the total time of TAG10 is decided by $max(N_m)$.

The performance of error ratio (Definition 3) against N in UD and M10D are shown in Fig. 4 (c) and (d) respectively. It is found that (i) ALOHA and ABS are always 0% owing to no error counting; (ii) ε of TAG10 is about 25% better than USE and LoF, especially much better in the small scale. Hence, TAG can use for statistics in all scale. In addition, if we repeat TAG more times leads to less standard deviation in statistic process.

Varying Number of Classifications: Performance evaluation is also carried out when M changes from 1 to 20.

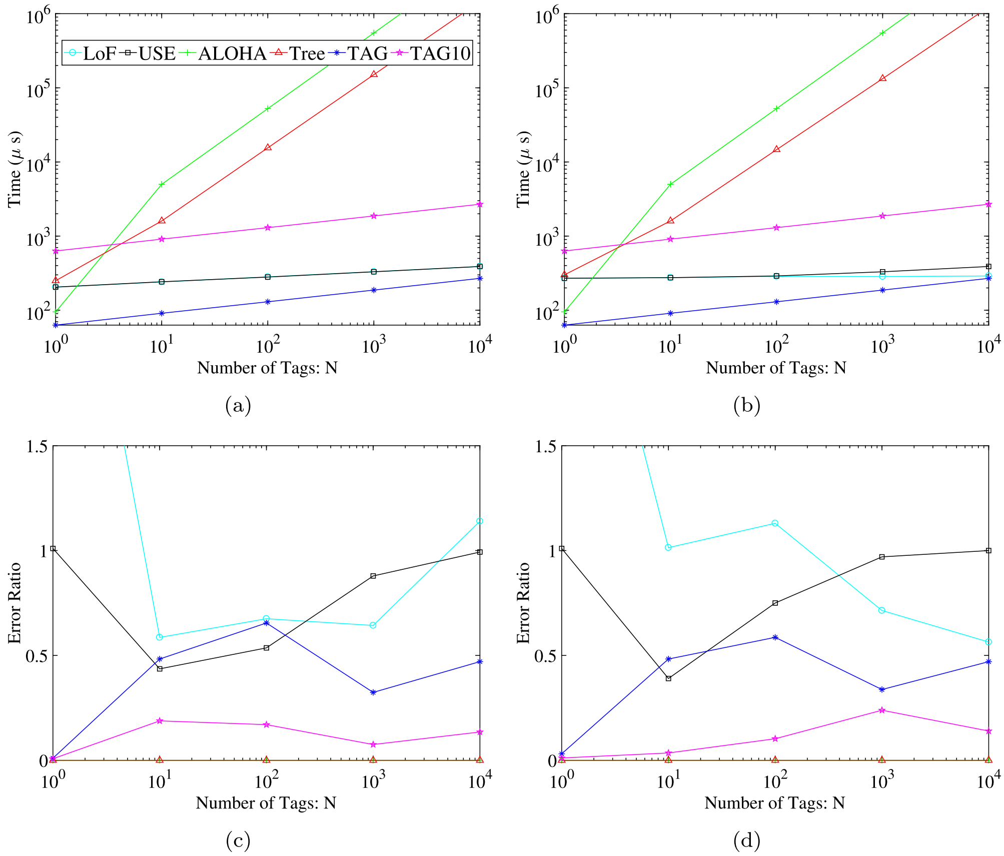

Fig. 4. (a) T_{total} against N under uniform distribution. (b) T_{total} against N under Max-1-0 distribution. (c) ε against N under uniform distribution. (d) ε against N under Max-1-0 distribution. Fig. 4 (a) and (b) are log-scale graphs.

Fig. 5 (a) and (b) illustrate the performance of T_{total} with varying M in two distributions. We observe that (i) there are almost no changes for ALOHA and ABS. T_{total} of them does not depend on M; (ii) T_{total} of the other four approaches increases when M increases; (iii) TAG10 still provides the good performance of T_{total}. Obviously, TAG10 are periodic waves in Fig. 5 (a) or (b). The jumps exist when $M = 5, 10, 15$, where are the multiples of $F = 5$. In these positions, TAG needs one more turn to allocate all CIDs to subcarriers.

Error ratios against M are exhibited in Fig. 5 (c) and (d). It can be seen that (i) TAG10 keeps performing better than USE and LoF; (ii) ε of the four approaches are more sensitive to N than to M when comparing with Fig. 4 (c) and (d).

8. Conclusion

In this paper, we have formulated a new problem classification statistics in RFID systems. We have also discovered the significance and challenges of time efficiency issue in this problem. However, nearly no existing approaches can solve this problem satisfactorily. To address this problem, we have proposed a novel TAG approach. TAG achieves the processing time in $O(logN)$ by accelerating the classification in frequency domain as well as the statistics in time domain. Theoretical analysis and simulation show the feasibility and high-performance of TAG.

Since the fast classification statistics of RFID system is a relatively new concept, several issues still remain to be studied. e.g., this paper studies the single reader case for RFID classification statistics. Using multiple readers by TAG to further improve the performance of RFID classification statistics efficiency and accuracy is an interesting topic. Moreover, we plan to extend the TAG method from simulation to real world applications in our future work.

Acknowledgements

The work is partly supported by China NSF grants (61672349, 61672353, 61472252, 61373155) and China 973 project (2014CB340303).

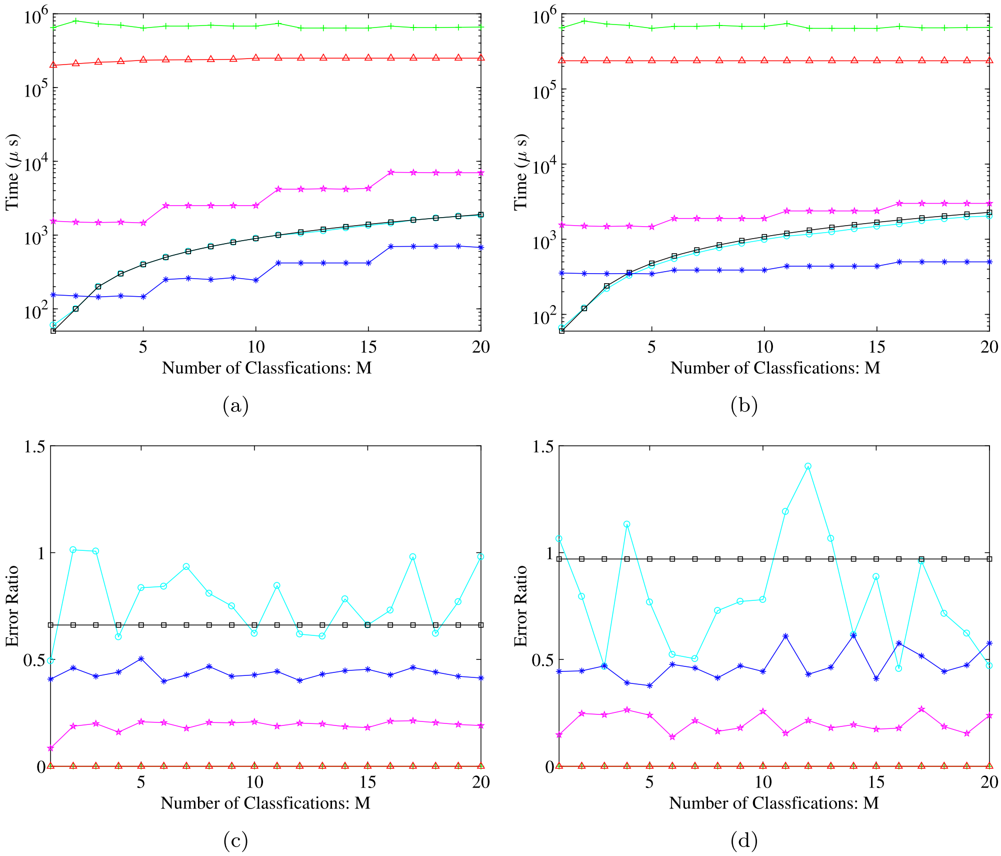

Fig. 5. (a) T_{total} against M under uniform distribution. (b) T_{total} against M under Max-1-0 distribution. (c) ε against M under uniform distribution. (d) ε against M under Max-1-0 distribution. Fig. 5 (a) and (b) are log-scale graphs.

References

[1] W. Gong, H. Liu, L. Chen, K. Liu, Y. Liu, Fast composite counting in RFID systems, IEEE/ACM Trans. Netw. 24 (5) (2016) 2756–2767.

[2] H. Liu, W. Gong, L. Chen, W. He, K. Liu, Y. Liu, Generic composite counting in RFID systems, in: IEEE 34th International Conference on Distributed Computing Systems (ICDCS), 2014, pp. 597–606.

[3] J. Park, C. Moon, I. Yeom, Y. Kim, Cardinality estimation using collective interference for large-scale RFID systems, J. Netw. Comput. Appl. 83 (2017) 101–110.

[4] X. Liu, B. Xiao, K. Li, A.X. Liu, J. Wu, X. Xie, H. Qi, RFID estimation with blocker tags, IEEE/ACM Trans. Netw. 25 (1) (2017) 224–237.

[5] U.A. Patel, R. Swaminarayan Priya, Development of a student attendance management system using rfid and face recognition: a review, Int. J. Adv. Res. Comput. Sci. Manag. Stud. 2 (8) (2014) 109–190.

[6] H. Li, P. Zhang, S. Al Moubayed, S.N. Patel, A.P. Sample, Id-match: a hybrid computer vision and rfid system for recognizing individuals in groups, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2016, pp. 4933–4944.

[7] S. Bhagat, V. Singh, N. Khajuria, B. Student, Atm security using iris recognition technology and RFID, Internat. J. Engrg. Sci. (2017) 11486.

[8] B. Williams, Intelligent Transport Systems Standards, Artech House, 2008.

[9] S.A. Ahson, M. Ilyas, RFID Handbook: Applications, Technology, Security, and Privacy, CRC Press, 2017.

[10] R.Y. Zhong, G.Q. Huang, S. Lan, Q. Dai, X. Chen, T. Zhang, A big data approach for logistics trajectory discovery from RFID-enabled production data, Int. J. Prod. Econ. 165 (2015) 260–272.

[11] F. Bibi, C. Guillaume, N. Gontard, B. Sorli, A review: rfid technology having sensing aptitudes for food industry and their contribution to tracking and monitoring of food products, Trends Food Sci. Technol. 62 (2017) 91–103.

[12] J. Chai, C. Wu, C. Zhao, H.-L. Chi, X. Wang, B.W.-K. Ling, K.L. Teo, Reference tag supported rfid tracking using robust support vector regression and kalman filter, Adv. Eng. Inform. 32 (2017) 1–10.

[13] J. Sing, N. Brar, C. Fong, The state of rfid applications in libraries, Infor. Technol. Libr. 25 (1) (2006) 24–32.

[14] R. Want, An introduction to RFID technology, IEEE Pervasive Computing 5 (1) (2006) 25–33.

[15] EPC tag data standard, https://www.gs1.org/sites/default/files/docs/epc/GS1_EPC_TDS_i1_10.pdf, March 2017.

[16] Y. Zheng, M. Li Pet, Probabilistic estimating tree for large-scale RFID estimation, IEEE Trans. Mob. Comput. 11 (11) (2012) 1763–1774.

[17] L. Arjona, H. Landaluce, A. Perallos, E. Onieva, Scalable RFID tag estimator with enhanced accuracy and low estimation time, IEEE Signal Process. Lett. 24 (7) (2017) 982–986.

[18] S.-R. Lee, S.-D. Joo, C.-W. Lee, An enhanced dynamic framed slotted aloha algorithm for RFID tag identification, in: The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, MobiQuitous 2005, 2005, pp. 166–172.

[19] L.G. Roberts, Aloha packet system with and without slots and capture, Comput. Commun. Rev. 5 (2) (1975) 28–42.

[20] J. Myung, W. Lee, Adaptive splitting protocols for RFID tag collision arbitration, in: Proceedings of the 7th ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2006, pp. 202–213.

[21] L. Kong, L. He, Y. Gu, M.-Y. Wu, T. He, A parallel identification protocol for RFID systems, in: Proceedings IEEE, INFOCOM, 2014, 2014, pp. 154–162.
[22] F.J. Álvarez, T. Aguilera, J.A. Paredes, J.A. Moreno, Acoustic tag identification based on noncoherent fsk detection with portable devices, IEEE Trans. Instrum. Meas. 67 (2) (2018) 270–278.
[23] Y. Hou, Y. Zheng, Phy assisted tree-based RFID identification, in: IEEE Conference on Computer Communications, INFOCOM 2017, IEEE, 2017, pp. 1–9.
[24] X. Tan, H. Wang, L. Fu, J. Wang, H. Min, D.W. Engels, Collision detection and signal recovery for uhf rfid systems, IEEE Trans. Autom. Sci. Eng. 15 (1) (2018) 239–250.
[25] X. Liu, B. Xiao, F. Zhu, S. Zhang, Let's work together: fast tag identification by interference elimination for multiple RFID readers, in: IEEE 24th International Conference on Network Protocols (ICNP), 2016, pp. 1–10.
[26] C. Zhao, C. Wu, J. Chai, X. Wang, X. Yang, J.-M. Lee, M.J. Kim, Decomposition-based multi-objective firefly algorithm for rfid network planning with uncertainty, Applied Soft Computing 55 (2017) 549–564.
[27] M. Kodialam, T. Nandagopal, Fast and reliable estimation schemes in RFID systems, in: Proceedings of the 12th Annual International Conference on Mobile Computing and Networking, 2006, pp. 322–333.
[28] Y. Hou, J. Ou, Y. Zheng, M. Li, Place: physical layer cardinality estimation for large-scale RFID systems, IEEE/ACM Trans. Netw. 24 (5) (2016) 2702–2714.
[29] L.-H. Zhang, T. Li, T.-J. Fan, Radio-frequency identification (rfid) adoption with inventory misplacement under retail competition, European J. Oper. Res. 270 (3) (2018) 1028–1043.
[30] Q. Xiao, S. Chen, M. Chen, Y. Zhou, Z. Cai, J. Luo, Adaptive joint estimation protocol for arbitrary pair of tag sets in a distributed RFID system, IEEE/ACM Trans. Netw. 25 (5) (2017) 2670–2685.
[31] X. Liu, K. Li, A.X. Liu, S. Guo, M. Shahzad, A.L. Wang, J. Wu, Multi-category RFID estimation, IEEE/ACM Trans. Netw. 25 (1) (2017) 264–277.
[32] P. Flajolet, G.N. Martin, Probabilistic counting algorithms for data base applications, J. Comput. System Sci. 31 (2) (1985) 182–209.
[33] B. Sheng, C.C. Tan, Q. Li, W. Mao, Finding popular categories for RFID tags, in: Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2008, pp. 159–168.