# Wise Counting: Fast and Efficient Batch Authentication for Large-Scale RFID Systems

Wei Gong[*†], Yunhao Liu[*], Amiya Nayak[†], Cheng Wang[†]
{gongwei,yunhao}@greenorbs.com, {nayak,cwan3}@uottawa.ca

[*]School of Software, TNLIST, Tsinghua University, China
[†]School of Electrical Engineering and Computer Science, University of Ottawa, Canada

## ABSTRACT

Radio Frequency Identification technology (RFID) is widely used in many applications, such as asset monitoring, e-passport and electronic payment, and is becoming one of the most effective solutions in cyber physical system. Since the identification alone does not provide any guarantee that tag corresponds to genuine identity, authentication of tag information is needed in most RFID systems. Meanwhile, as the number of tags is rapidly growing in recent years, per-tag based methods suffer from severely low efficiency and thus give way to probabilistic batch authentication. Most previous methods, however, share a common drawback from statistical perspective: they fail to explore correlation information, i.e., they do not comprehensively utilize all the information in authentication data structures. In addition, those schemes are not scalable well when multiple tag sets need to be verified simultaneously. In this paper, we propose a fast and efficient batch authentication scheme, Wise Counting (WIC), for large-scale RFID systems. We are the first to formally introduce the general batch authentication problem with multiple tag sets and give counterfeits estimation scheme with high efficiency. By employing a novel hierarchical authentication structure, we show that WIC is able to fast and efficiently authenticate both a single tag set and multiple tag sets in an easy, intuitive way. Through detailed theoretical analysis and extensive simulations, we validate the design of WIC and demonstrate its large superiority over state-of-the art approaches.

## Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Wireless Communication; H.4 [**Information Systems Applications**]: Miscellaneous

## Keywords

RFID tags, batch authentication, counterfeits estimation, hierarchical data structure

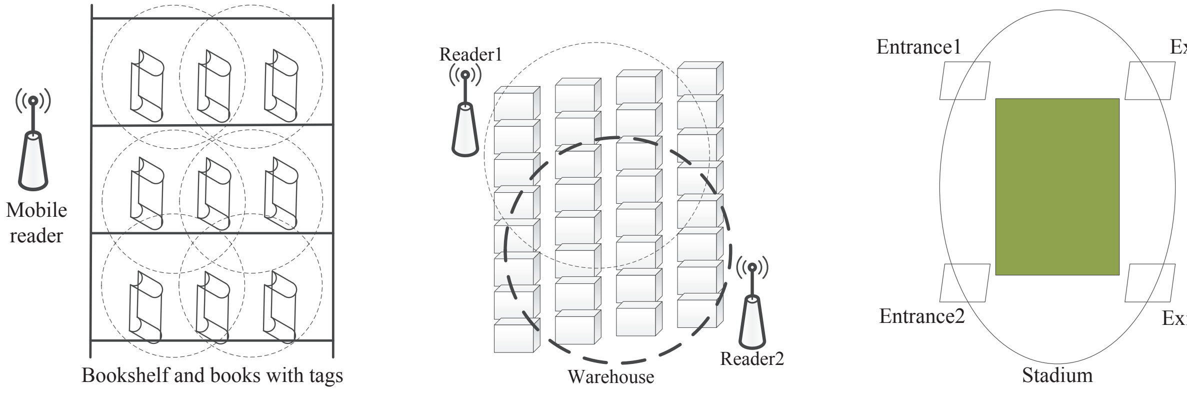
*MobiHoc'14,* August 11–14, 2014, Philadelphia, PA, USA.
Copyright 2014 ACM 978-1-4503-2620-9/14/08 ...$15.00.
http://dx.doi.org/10.1145/2632951.2632963 .

## 1. INTRODUCTION

Since *friend or foe* system was first developed in Second World War, Radio Frequency Identification (RFID) technology is widely used in many fields, such as asset monitoring[1][2][3], e-passport [4] and electronic payment [5][6], and is becoming one of the most effective solutions in cyber physical system [7][8][9]. A typical RFID system mainly consists of three parts: tag, reader and backend server. The tag, which is often attached to the object, contains a unique identification - Electronic Product Code (EPC) [10]. The reader is usually in charge of identification, i.e., querying the tag and collecting the information from it. The backend server that is connected to the reader can perform various operations, such as searching further object information according to specific EPC. But identification itself does not provide any means to certainly assure that the collected ID/EPC exactly corresponds to the genuine identity. Therefore, tag authentication (or verification) is indeed necessary.

Previously, much work focus on how to securely and effectively authenticate a single tag. Supposed that there are $N$ tags registered in backend server. By leveraging the one-way hash function, Hash Lock scheme is proposed to linearly search the backend database [11]. The authentication complexity of it is $\mathcal{O}(N)$. To achieve better performance, the search efficiency is improved to $\mathcal{O}(\log N)$ based on tree data structure, at the expense of storage spaces on tag [12][13]. Nevertheless, those per-tag based approaches suffer from severe scalability problem in batch authentication mode. As stated in [14], one reader can only authenticate 103 batches (each with 10,000 tags) per day. It is unlikely to accept such a low authentication efficiency in practice. Meanwhile, as RFID technology is used as one of the most important anti-counterfeiting measures for fast-moving consumer goods that are growing fast over years [15], the population of tags is expected to dramatically increase in the next few years. Thus one of top concerns for practical large-scale RFID systems is how to efficiently authenticate tags in batch.

The key advantage of batch authentication over per-tag based methods is the relaxation of the result, i.e., trading a small authentication failure probability for better authentication performance. So we also call it probabilistic batch authentication. There are two different but close-related probabilistic batch authentication problems: counterfeits detection and counterfeits estimation. In counterfeits detection problem, the result of authentication is binary. For example, in [14], positive result means detection of counterfeits. Negative result indicates that all tags in the batch are judged



(a) Continuous scanning (b) Inventory control (c) Flow monitoring

**Figure 1: Scenarios of authenticating multiple tag sets**

as legitimate with high probability. Obviously, this type of result is so coarse since it is unable to discriminate a batch with only one counterfeit tag from a batch with all fake tags. In fact, many RFID applications need to know more detailed information about counterfeits. Consequently, a fine-grained authentication scheme is proposed to approximate the count of counterfeit tags, addressing counterfeits estimation issue [16]. Although these two problems are separated, they are complementary to each other. Typically, the administrator of RFID system first performs counterfeits detection operation. If the result turns out to be positive, then counterfeits estimation algorithm can be used to obtain an accurate count of counterfeits. In this paper, we aim to efficiently solve counterfeits estimation problem with arbitrary accuracy requirements.

By reviewing former methods from statistical perspective, we observe two important issues that are neglected before. First, they view authentication synopses as separate entities and thus do not explore correlation information among them [14][16]. Second, the estimation is confined on singleton-observed synopses which account for a relative small fraction, wasting a large amount of other authentication synopses [16]. In sum, counterfeits estimators in previous work are not statistically efficient since they do not comprehensively utilize all the information in authentication structures, giving plenty of rooms for authentication efficiency optimization. More importantly, most previous methods are designed for authenticating one single batch at a time. But in practice, it would always be required to authenticate multiple tag sets/batches simultaneously.

In this paper, we propose a fast and efficient batch authentication scheme, Wise Counting (WIC), for large-scale RFID systems. First, we formally formulate the general counterfeits estimation problem for probabilistic batch authentication with multiple tag sets. Then, we introduce a novel hierarchical authentication structure, which combines uniform random hash and minimum order statistics. With the help of this building block data structure, WIC achieves $\mathcal{O}(\frac{N}{\varepsilon^2|C|}\log\log N)$ authentication complexity (communication complexity), where $|C|$ is the number of counterfeits from multiple tag sets and $\varepsilon$ is the desired relative error. Through detailed analysis and extensive simulated comparison, we show that WIC is significantly advantageous over previous methods.

The major contributions of this work are as follows.

1. We are the first to define batch authentication problem with multiple tag sets and provide a fast and efficient authentication scheme, WIC, which works well with one or more tag sets.

2. We design a novel hierarchical authentication structure to significantly reduce authentication and communication overheads. In particular, WIC wisely utilize all minimum order statistics of this data structure to obtain much more accurate counterfeits estimate, compared with previous methods. We also give a general and implementable communication protocol for both readers and tags using WIC.

3. We validate the design of WIC through both theoretical analysis and extensive simulations. We also show that WIC is effective and significantly outperforms state-of-the-art approaches.

# 2. PRELIMINARIES

## 2.1 Motivation

Previous work on batch authentication are designed for simple scenario in which only one tag set needs to be verified. Authenticating multiple tag sets, however, are required in many practical and complex scenarios to meet various demands. Considering the following three examples as shown in Figure 1:

1. **Continuous scanning** is a common and basic method in large-scale RFID system [17][18][19]. It uses multiple continuous scanning operations to cover a much larger area than a single scanning. Figure 1(a) shows that a mobile reader performs scanning several times to collect all book information on the bookshelf. In order to authenticate all books, we need to verify compound set as: $R = R_1 \cup R_2 \cup ... \cup R_n$, where $R_i$ is the set of scanned tags in $i$-th scanning.

2. **Inventory control** generally employs multiple static readers deployed at different locations, according to

density, size and orientation of objects in warehouse [20][21][22]. As illustrated in Figure 1(b), Reader1 and Reader2 are deployed to cover all goods in the warehouse. If we want to authenticate the goods that both covered by Reader1 and Reader2, then the compound set $R = R_1 \cap R_2$ needs to be tested, where $R_1$ and $R_2$ are the sets of tags scanned by Reader1 and Reader2 respectively.

3. **Flow monitoring** is essential for many large-scale events for security reason, such as rallies, conferences and sport games [23]. Figure 1(c) demonstrates an example in stadium with 2 entrances and 2 exits. A reader is equipped at each gate to monitor people flow passing by and the person who holds a genuine tagged-ticket is authorized to go in and out unlimited times. If administrator needs to estimate the number of unauthorized audiences in the stadium, the compound set $R = (R_1 \cup R_2) - (R_3 \cup R_4)$ needs to be authenticated, where $R_1$, $R_2$, $R_3$ and $R_4$ are the sets of tags scanned by Entrance1, Entrance1, Exit1 and Exit2, respectively.

## 2.2 Challenges

Estimating the number of counterfeits in multiple tag sets at the same time is an important but not trivial problem in large-scale RFID systems. Since multiple tag sets might have intersections as shown in Figure 1, it cannot be easily reduced to estimate the counterfeits number one set by one set. Of course, there is a naive way to authenticate multiple tag set: first aggregate all raw tag information on server from each scanning/set, then compute the compound set according to required set expression, finally authenticate the compound set. Although this naive method works, it violates several principles that we strive to. First, the authentication process should be fast and time-efficient, i.e., the communication overhead should be kept minimum. Second, the raw tag information transmission is better to be avoided since it may create potentials for infringements of RFID data privacy [24].

## 2.3 Problem Formulation

In a typical RFID system, the backend server often registers a number of tags which are said to be legal. Let $\mathcal{S}$ be the set of all legitimate tags on server and $N$ be the cardinality of $\mathcal{S}$. Usually, there are $\mathcal{T}_1$, $\mathcal{T}_2$, ..., $\mathcal{T}_l$ tag sets to be verified, where $l$ is an integer more than 1. In many scenarios (e.g., using multiple readers to cover a large area), the system administrator needs to know how many counterfeit tags in the compound expressions of $\mathcal{T}_i$. We therefore generalize this as

$$\mathcal{T}' = \mathcal{T}_1 \bigodot \mathcal{T}_2 \bigodot ... \bigodot \mathcal{T}_l,$$

where $\mathcal{T}'$ is the final compound set to be verified, and $\bigodot$ can be replaced by any one of $-$ (difference), $\cup$ (union), and $\cap$ (intersection).

Moreover, we define that the tag is counterfeit if the tag's information is not registered on the server. Otherwise, the tag is genuine. Similar to most prior work [14][16][25], *we do not discuss the problem the genuine tag is attached to counterfeit goods, vice versa.* Any problems related to tag misplacement/duplication are beyond the scope of this paper. Further, we assume that both genuine and counterfeit tags obey the same transmission protocol. By the definition of counterfeits, we use $\mathcal{C}$ to denote the set of counterfeits to be verified, thus

$$\mathcal{C} = \mathcal{T}' - \mathcal{S}.$$

**Table 1: Main Notations**

| Symbols | Descriptions |
|---|---|
| $\mathcal{S}$ | the set of registered tags. |
| $\mathcal{T}$ | the set of tags to be verified. |
| $\mathcal{C}$ | the set of counterfeits tags. |
| $\mathcal{G}$ | the set of genuines tags. |
| $N$ | the cardinality of tags registered on server |
| $n$ | the cardinality of tags to be verified |
| $M$ | the cardinality of $\mathcal{T} \cup \mathcal{S}$. |
| $f$ | the size of frame. |
| $c$ | the counterfeit ratio. |
| $\|\|$ | the cardinality of set. |
| ˆ | the estimate of original parameter. |
| $\mathcal{A}^{\mathcal{T}}$ | the HAS of set $\mathcal{T}$. |

The counterfeits estimation in probabilistic batch authentication is to obtain the approximate cardinality of $\mathcal{C}$, i.e., $|\mathcal{C}|$. In order to measure the accuracy of estimation result, we use two parameters: relative standard error $\varepsilon$ and failure probability $\delta$. Thereby the result of counterfeits estimation, $|\hat{\mathcal{C}}|$, must satisfy

$$\Pr\left[\left|\frac{|\mathcal{C}|}{|\hat{\mathcal{C}}|} - 1\right| \leq \varepsilon\right] \geq 1 - \delta. \quad (1)$$

For instance, if the exact number of counterfeits is 100, and $\varepsilon = 0.01, \delta = 0.01$, then the output estimate $|\hat{\mathcal{C}}|$ should be between 99 to 101 with probability 0.99. Similarly, we use $\mathcal{G}$ to denote the set of genuines in verified sets. The main notations are listed in Table 1.

## 2.4 Communication Model

We adopt the listen-before-talk mode and frame slotted ALHOA as the basic communication model, similar to [16][26]. In this model, the reader first queries tags with several predefined commands and parameters. Each tag then performs computations on board and decides whether and what to reply to reader based on preloaded transmission protocol. Besides EPC code, each tag has its own random hash function $\mathcal{H}$ or random number generator that takes several input parameters, e.g., frame size, EPC code [27].

# 3. WISE COUNTING OVERVIEW

Wise Counting is an efficient probabilistic batch authentication scheme that provides accurate counterfeits estimate, achieving optimal authentication efficiency.

Following a common practice in RFID authentication, the backend server has stored all the information of registered tags. When authentication starts, the backend server initializes parameters according to the user defined requirements, $\varepsilon$ and $\delta$. Then the server sends those parameters to readers through a secure line. The readers use those received parameters to gather authentication information from tags

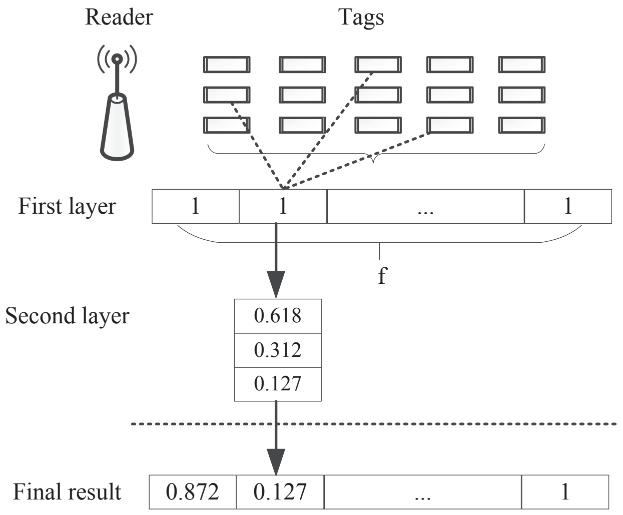


**Figure 2: Hierarchical authentication structure generation.**

to be verified. After aggregating all the authentication information from readers, the server performs verification to derive counterfeits estimate.

With a novel Hierarchical Authentication Structure (HAS) as later shown in Section 4, WIC goes through the following main steps:

1. The reader queries the tags in the operation range and builds its own HAS ($\mathcal{A}^{\mathcal{T}_i}$) for $\mathcal{T}_i$.
2. The server aggregates HASes from readers and derives observed counterfeits ratio based on the compound set expression.
3. The server estimates the cardinality of $\mathcal{T}'\cup\mathcal{S}$ ($|\widehat{\mathcal{T}'\cup\mathcal{S}}|$). Since $\mathcal{C}$ is a portion of $\mathcal{T}'\cup\mathcal{S}$, then we combine the observed counterfeits ratio and $|\widehat{\mathcal{T}'\cup\mathcal{S}}|$ to deduce the counterfeits estimate, $|\hat{\mathcal{C}}|$.

The next few sections elaborate on the above steps and provide detailed analysis.

## 4. DATA STRUCTURE

In this section, we describe how to generate hierarchical authentication structure from the tag set and then give the basic estimation formula.

Here we employ uniform random hash and minimum order statistics to produce HAS. As shown in Figure 2, initially the reader issues the query command to tags. Each tag then decides when (in which time slot) and what to reply according to received parameters, such as frame size ($f$) and random seed. The reply procedure is based on slotted ALOHA model and is virtually divided into two layers. In the first layer, each tag chooses replied time slot based on a uniform random hash function, similar to much prior work [14][25][26]. At the reader side, all slots are marked as 1. In the second layer, each tag replies a uniform random number, $R(EPC)$, between [0,1]. At the same time, the reader receives all the random numbers from tag(s) and records only the minimum number. In order to understand the principle of HAS, we are better to set aside possible collisions and just assume that the reader is able to extract minimum number from many tags' responses in one time slot. The details about the generation of replied random number on tag and reception of minimum random number on reader are given in Section 6. The final HAS consists of $f$ numbers. If the number is 1, it stands for no response in this time slot. Otherwise, the number is the minimum of all responses in this time slot. The pseudocodes of HAS generation are given in Algorithm 1 for the tag and Algorithm 2 for the reader.

**Algorithm 1** HAS generation algorithm for the tag

```
Receive a probing message from reader, containing frame
   size $f$ and random seed $r$.
Compute reply slot number $sn = \mathcal{H}(EPC, f, r)$.
while TRUE do
   wait-for-slot-start().
   if $sn == 0$ then
      generate random number $R(EPC)$.
      RepMinRndProc($R$).
   else
      $sn \leftarrow sn - 1$, keep silent.
   end if
end while
```

**Algorithm 2** HAS generation algorithm for the reader

```
Initialize $\mathcal{A}^{\mathcal{T}}[i] \leftarrow 1.0$ $(1 \leq i \leq f)$;
Broadcast a request to tags, including frame size $f$ and
   random seed $r$.
for $i = 1$ to $f$ do
   issue-slot-start() command.
   wait-for-tags-response().
   if there is no response in this slot then
      $\mathcal{A}^{\mathcal{T}}[i] \leftarrow 1$.
   else
      $\mathcal{A}^{\mathcal{T}}[i] \leftarrow$RecMinRndProc().
   end if
end for
```

Let $n$ be the number of tags in tested set ($\mathcal{T}$) and $\lambda = \frac{n}{f}$ be the mean number of tags' responses in each slot. We generate authentication synopsis $\mathcal{A}^{\mathcal{T}}[i]$ where $i = 1, 2, ..., f$ using Algorithm 1 and 2. We then study the characteristic of $\mathcal{A}^{\mathcal{T}}[i]$ and find that if $f$ is sufficiently large such that $(1-\frac{1}{f})^f \approx \frac{1}{e}$, then $\mathcal{A}^{\mathcal{T}}[i]$ approximates an independent sample from right truncated exponential distribution with rate parameter $\lambda$, where the cumulative distribution function is as follow [1]

$$F(x;\lambda) \approx \begin{cases} 1 - e^{-\lambda x} & 0 \leq x \leq 1 \\ 0 & x < 0 \text{ or } x > 1 \end{cases}$$

and the correlation coefficient ($\rho$) of $\mathcal{A}^{\mathcal{T}}[i]$ and $\mathcal{A}^{\mathcal{T}}[j]$ is

$$\rho(\mathcal{A}^{\mathcal{T}}[i], \mathcal{A}^{\mathcal{T}}[j]) \approx -\frac{1}{n}, i \neq j.$$

Therefore, we are able to easily obtain the MLE estimator of $\lambda$. Since the dependence between $\mathcal{A}^{\mathcal{T}}[i]$ and $\mathcal{A}^{\mathcal{T}}[j]$ is so weak, we ignore it and get the simplified likelihood function

[1] Due to limited space, we omit derivations/proofs here and the rest of this paper.

of $\lambda$ as

$$\mathcal{L}(\lambda) = e^{-\lambda \sum_{i=1}^{f} \mathbf{1}(\mathcal{A}^{\mathcal{T}}[i])} \prod_{\mathcal{A}^{\mathcal{T}}[i]<1} \lambda e^{-\lambda \mathcal{A}^{\mathcal{T}}[i]}.$$

where $\mathbf{1}(x)$ is an indicator function. If $x = 1$, $\mathbf{1}(x) = 1$. Otherwise, $\mathbf{1}(x) = 0$. The MLE of $\lambda$ and thereby estimated number of tags in $\mathcal{A}^{\mathcal{T}}$ are derived as

$$\hat{\lambda} = \frac{\sum_{i=1}^{f} \mathbf{1}'(\mathcal{A}^{\mathcal{T}}[i])}{\sum_{i=1}^{f} \mathcal{A}^{\mathcal{T}}[i]}, \hat{n} = f\hat{\lambda}. \tag{2}$$

where $\mathbf{1}'(x)$ is an indicator function. If $x < 1$, $\mathbf{1}'(x) = 1$. Otherwise, $\mathbf{1}'(x) = 0$.

In order to understand the simplicity of the estimation procedure, let's see a toy example. Suppose that we test a tag set ($n = 100$) and get a HAS of size 7: 0.0768, 0.0183, 1, 0.0307, 0.0125, 0.1642, 0.1217. The estimated result $\hat{n} = 7 * \frac{6}{0.0768+0.0183+0.0307+0.0125+0.1642+0.1217} = 99$.

# 5. ESTIMATING COUNTERFEITS AND GENUINES

After acquiring basic knowledge about our authentication data structure, now we first demonstrate our method to estimate the number of counterfeits and genuines in single set scenario. Later we introduce several techniques to extend estimation over compound sets.

## 5.1 Authenticating Single RFID Set

We first study an important characteristic of the probabilistic distribution for truncated exponential random variable.

Given $X_1 \sim Exp(\lambda_1)$, $X_2 \sim Exp(\lambda_2)$, and their corresponding right truncated distribution $X_1{}' = \min(X_1, 1)$, $X_2{}' = \min(X_2, 1)$. If $X_1$ is independent of $X_2$, then we have $\min(X_1{}', X_2{}') \sim \min(Exp(\lambda_1 + \lambda_2), 1)$. Further, we derive

$$\Pr[X_i{}' = \min(X_1{}', X_2{}')] = e^{-(\lambda_1+\lambda_2)} + \frac{\lambda_i}{\lambda_1 + \lambda_2}(1 - e^{-(\lambda_1+\lambda_2)}) \tag{3}$$

, where $i = 1, 2$.

This characteristic tells us that if two exponential random variables are independent of each other, then minimum order of their truncated version is easily deduced.

Now we have only one set to be verified, $\mathcal{T}$, and aim to estimate $|\mathcal{T} - \mathcal{S}|$. The direct application of above equation is to divide the union, $\mathcal{T} \cup \mathcal{S}$, into three independent parts: $\mathcal{T} - \mathcal{S}$, $\mathcal{T} \cap \mathcal{S}$, and $\mathcal{S} - \mathcal{T}$. We use $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{B}_3$ to denote virtual HASes for those three independent sets. Therefore we have [2]

$$\begin{aligned} \Pr[\mathcal{A}^{\mathcal{T}} = \mathcal{A}^{\mathcal{S}}] &= \Pr[\min(\mathcal{B}_1, \mathcal{B}_2) = \min(\mathcal{B}_2, \mathcal{B}_3)] \\ &= \Pr[\mathcal{B}_2 = \min(\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3)] \\ &\approx e^{-\frac{M}{f}} + (1 - e^{-\frac{M}{f}}) \frac{|\mathcal{T} \cap \mathcal{S}|}{M}. \end{aligned}$$

where $M$ is the cardinality of $\mathcal{T} \cup \mathcal{S}$. We thus derive the estimation equations for the number of counterfeits and genuines in single set scenario as follows.

THEOREM 1. *For a large $f$ such that $(1 - \frac{1}{f})^f \approx \frac{1}{e}$, then*

$$|\hat{\mathcal{C}}| = |\widehat{\mathcal{T} - \mathcal{S}}| = \Pr'[\mathcal{A}^{\mathcal{T}} < \mathcal{A}^{\mathcal{S}}] \cdot \frac{\hat{M}}{(1 - e^{-\frac{\hat{M}}{f}})} \tag{4}$$

$$|\hat{\mathcal{G}}| = |\widehat{\mathcal{T} \cap \mathcal{S}}| = (\Pr'[\mathcal{A}^{\mathcal{T}} = \mathcal{A}^{\mathcal{S}}] - e^{-\frac{\hat{M}}{f}}) \cdot \frac{\hat{M}}{(1 - e^{-\frac{\hat{M}}{f}})} \tag{5}$$

*, where $\Pr'$ denotes observed probability in $\mathcal{A}^{\mathcal{T}}$ and $\mathcal{A}^{\mathcal{S}}$.*

The common unknown parameter in above two equations is $\hat{M}$. Fortunately, the HAS of $\mathcal{T} \cup \mathcal{S}$, $\mathcal{A}^{\mathcal{T} \cup \mathcal{S}}$, is just the slot-wise $\min(\mathcal{A}^{\mathcal{T}}[i], \mathcal{A}^{\mathcal{S}}[i])$. It is easy to obtain $\hat{M}$ by applying equation 2 on $\mathcal{A}^{\mathcal{T} \cup \mathcal{S}}$. Further, if $\hat{M}$ is large enough that $e^{-\frac{\hat{M}}{f}}$ is close to 0, we derive simplified theorem 1 as

$$|\hat{\mathcal{C}}| = \Pr'[\mathcal{A}^{\mathcal{T}} < \mathcal{A}^{\mathcal{S}}] \cdot \hat{M} \tag{6}$$

$$|\hat{\mathcal{G}}| = \Pr'[\mathcal{A}^{\mathcal{T}} = \mathcal{A}^{\mathcal{S}}] \cdot \hat{M} \tag{7}$$

## 5.2 Extension to Multiple RFID Sets

In many practical situations, we need to verify how many counterfeits are in the compound set

$$\mathcal{T}' = \mathcal{T}_1 \bigodot \mathcal{T}_2 \bigodot ... \bigodot \mathcal{T}_l$$

, where $l \geq 1$, i.e., estimating $|\mathcal{C}| = |\mathcal{T}' - S|$. Still, let $M$ be the cardinality of union, $\mathcal{T}' \cup \mathcal{S}$. The estimation procedure is composed of three steps. We will illustrate each step with a concrete example. Consider $\mathcal{T}' = (\mathcal{T}_1 \cap \mathcal{T}_2) \cup \mathcal{T}_3$.

**Step one:** Use relation $|A - B| = |A| - |A \cap B|$ to remove all set difference operator(s). In our example,

$$|(\mathcal{T}_1 \cap \mathcal{T}_2) \cup \mathcal{T}_3 - S| = |(\mathcal{T}_1 \cap \mathcal{T}_2) \cup \mathcal{T}_3| - |((\mathcal{T}_1 \cap \mathcal{T}_2) \cup \mathcal{T}_3) \cap S|$$

**Step two:** Use relation $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ to replace union of set intersection by the intersection of set union. In our example,

$$|(\mathcal{T}_1 \cap \mathcal{T}_2) \cup \mathcal{T}_3| = |(\mathcal{T}_1 \cup \mathcal{T}_3) \cap (\mathcal{T}_2 \cup \mathcal{T}_3)|$$

and

$$|((\mathcal{T}_1 \cap \mathcal{T}_2) \cup \mathcal{T}_3) \cap S| = |(\mathcal{T}_1 \cup \mathcal{T}_3) \cap (\mathcal{T}_2 \cup \mathcal{T}_3) \cap \mathcal{S}|$$

.

**Step three:** Let $c$ be the ratio of counterfeits to the all sets union, i.e., $c = \frac{|\mathcal{C}|}{M}$. In order to get this counterfeits ratio, we first simply extended the theorem 1 to get the following equation:

$$\Pr[\mathcal{A}^{\mathcal{T}_0} = \mathcal{A}^{\mathcal{T}_1} = ... = \mathcal{A}^{\mathcal{T}_d} = \mathcal{A}_{\cup}] \approx e^{-\frac{M}{f}} + \frac{(1 - e^{-\frac{M}{f}})}{M} |\bigcap_{j=0}^{d} \mathcal{T}_j|$$

, where $\mathcal{A}_{\cup}$ is the HAS of $\bigcup_{i=0}^{l} \mathcal{T}_i$, $\mathcal{T}_0 = \mathcal{S}$, $|\bigcup_{i=0}^{l} \mathcal{T}_i| = M$ and $1 \leq d \leq l$. [3]

In our example, we thus have

$$c = P1 - P2 \tag{8}$$

[2] We omit [i] for each HAS in the equation to simplify the presentation in the following.

[3] Here we assume $e^{-\frac{M}{f}}$ is close to 0 and $f$ is large enough such that $(1 - \frac{1}{f})^f \approx \frac{1}{e}$. In case it is not negligible, we then invert equations theorem 1 to derive the closed-form solution.

**Algorithm 3** Reply minimal random value for tag: **RepMinRndProc**()

```
generate a random number $R(EPC)$ from unit exponential distribution EXP.
Save integer part of $R$ into $j = \log\log N$ bits, as $I[0] \sim I[j-1]$, position 0 is the highest bit.
Truncate decimal part of $R$ into $k = 10$ bits, as $I[j] \sim I[j+k-1]$, position j is the highest bit.
Initialize $T[i] \leftarrow 0$, $1 \le i < \log\log N + 10$.
$i \leftarrow 0$ and wait-for-subslot-start().
for $i = 0$ to $j + k - 1$ do
    wait-for-subslot-continue().
    while TRUE do
        Receive bit ’b’ and $T[i] \leftarrow b$.
        if $b == 1$ then
            if prefix $i$ bits of $T$ is equal to $I$ then
                instantly issue response and break while loop.
            else
                keep silent.
            end if
        end if
        if $b == 0$ then
            break while loop.
        end if
    end while
end for
```

where $P1 = \Pr[\min(\mathcal{A}^{\mathcal{T}_1}, \mathcal{A}^{\mathcal{T}_3}) = \min(\mathcal{A}^{\mathcal{T}_2}, \mathcal{A}^{\mathcal{T}_3}) = \mathcal{A}_{\cup}]$, $P2 = \Pr[\min(\mathcal{A}^{\mathcal{T}_1}, \mathcal{A}^{\mathcal{T}_3}) = \min(\mathcal{A}^{\mathcal{T}_2}, \mathcal{A}^{\mathcal{T}_3}) = \mathcal{A}^{\mathcal{T}_0} = \mathcal{A}_{\cup}]$.

Finally, by the definition of $c$, the estimated number of counterfeits for multiple tag sets is

$$|\hat{\mathcal{C}}| = \hat{M} \cdot c'. \tag{9}$$

where $\hat{M}$ is obtained by simply applying equation 2 on $\mathcal{A}_{\cup}$ and $c'$ is the observed ratio in equation 8.

## 6. THE EFFICIENCY OF WISE COUNTING

In this section, we analyze the time complexity of WIC and propose an optimization based on the exponential pseudorandom number generator.

### 6.1 Time Complexity

For an arbitrary set expression of tag sets to be verified, the relative standard error of WIC can be derived as :

$$\mathrm{RSE}^2_{wic} = \mathrm{E}(\frac{|\hat{\mathcal{C}}|_{WIC}}{|\mathcal{C}|} - 1)^2 \approx (fc)^{-1}.$$

Therefore if $\varepsilon$ is the required relative error, the size of frame should be $f \approx \varepsilon^{-2}c^{-1}$ for HAS generation in the first layer. [4]

For the second layer in HAS generation, we need to derive the number of subslots for transmitting the minimal random number. From section 4, we know that $\mathcal{A}^{\mathcal{T}}[i]$ follows a truncated exponential distribution: $\mathcal{A}^{\mathcal{T}}[i] \sim \min(\frac{\mathrm{EXP}}{\lambda}, 1)$, where EXP is the unit exponential variable.

Since we know there may be $M$ tags' responses in a single slot at worse and $\lambda \sim \mathcal{O}(M)$. This indicates that $\mathcal{O}(\log M)$ bits need to be transmitted in a single slot. In the following, we reduce this transmit time by employing an exponential pseudorandom number generator to replace the uniform random number generator described in Section 4.

**Algorithm 4** Receive minimal random value for reader: **RecMinRndProc**()

```
Initialize $T[i] \leftarrow 0$, $1 \le i < \log\log N + 10$.
issue-subslot-start() command.
for $i = 0$ to $\log\log N + 9$ do
    issue-subslot-continue() command
    Broadcast bit ’1’ and wait-for-tags-response().
    if there is no response in this subslot then
        $T[i] \leftarrow 0$. Broadcast bit ’0’.
    else
        $T[i] \leftarrow 1$.
    end if
end for
combine $T$’s integer and decimal part into a float $m$.
return minimal value $2^{-m}$.
```

Now we take log operation on $\mathcal{A}^{\mathcal{T}}[i]$ to have: $\log \mathcal{A}^{\mathcal{T}}[i] \sim \min(-\log\lambda + \log \mathrm{EXP}, 0)$. This motivates us to use $(\log\log M + \alpha)$ bits for transmission, where $-\log\log M$ represents integer part and $\alpha$ denotes decimal part. Hence the total time complexity of WIC is $\varepsilon^{-2}c^{-1}(\log\log M + \alpha)$.

We observed that $\alpha = 10$ is enough in practice. For instance, if we need to verified $\mathcal{T}_1$, then the observed probability for $\Pr[\mathcal{T}_0 = \mathcal{T}_1 = \mathcal{T}_{\cup}]$ is at most $2^{-\alpha+1} = 0.002$ shifted from ground truth. This difference is negligible for many applications. Actually the value of $\alpha$ is tunable according to various accuracy requirements. Algorithm 3 and 4 present the implementation of how to reply and receive the minimal value for both tag and reader. It is worth noting that we make a slight difference here. Instead of finding the minimal value from many uniform distributed variables, we are trying to acquire the maximal value among exponential distributed variables. This conversion stems from

$$\min(1, \mathcal{U}_0, \mathcal{U}_1, ...) \Rightarrow \max(0, -\log\mathcal{U}_0, -\log\mathcal{U}_1, ...),$$

where $\mathcal{U}_i$ is uniform random number in [0,1] and hence $-\log\mathcal{U}_i$ follows unit exponential distribution.

## 7. EVALUATION

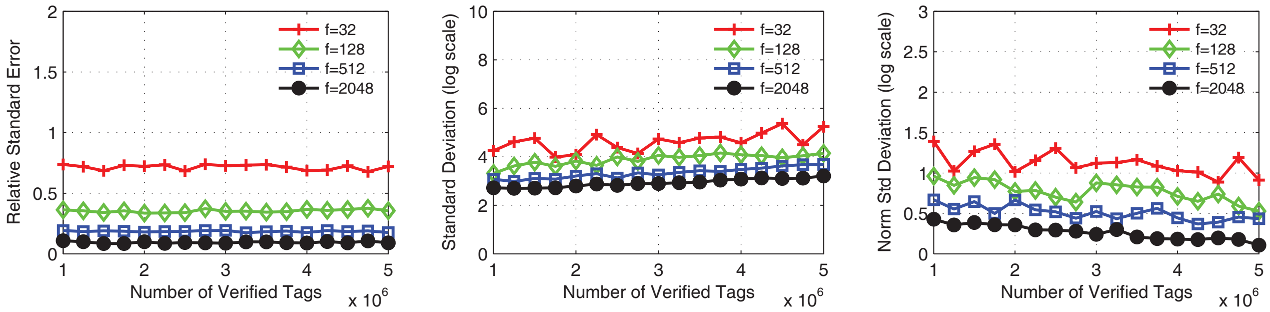
In this section, we evaluate the performance of WIC and compare it with the state-of-the-art method INC [16] under extensive simulations.

### 7.1 Setup and Metrics

The simulations are conducted on a PC with Intel i7 CPU at 3.2GHz and 8GB RAM. We write our own simulator adopting C# programming language. We register 10,000,000 tags on server using MS SQLServer 2000 database. We take 500 runs for each experiment and report the average.

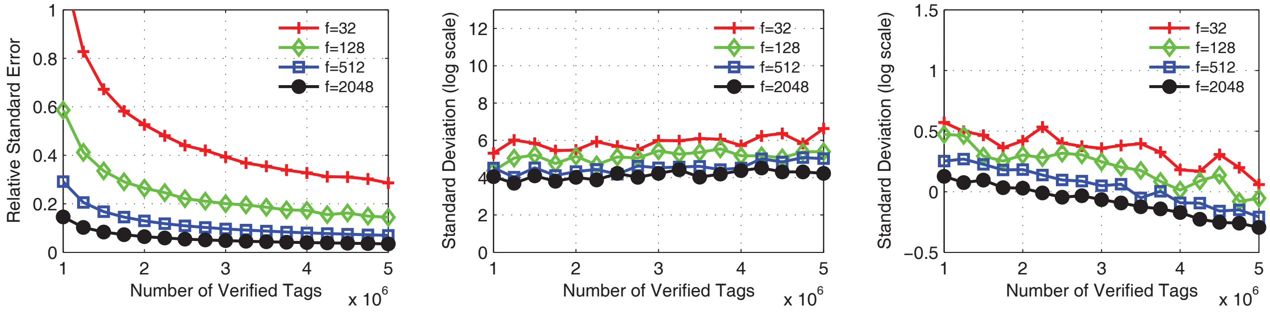
**Estimation Accuracy:** To evaluate the accuracy of WIC estimator, we use three standard metrics: Relative Standard Error (RSE), $|\frac{\hat{\theta}}{\theta} - 1|$; Standard Deviation (SD), $\sigma = \sqrt{\mathrm{E}[(\hat{\theta} - \theta)^2]}$; Norm Std Deviation (NSD), $\sigma' = \frac{\sigma}{\mathrm{E}[\theta]}$, where $\hat{\theta}$ is the estimate of original value $\theta$. Without loss of generality, we vary the size of verified tag set ($\mathcal{T}_1$) from 1,000,000 to 5,000,000, but fix counterfeits ratio at 0.07 which is the

[4] For simplicity, we do not include the analysis of $\delta$ in the rest part of this paper. Since it is well-known [28] that if we are able to achieve a constant $\delta$, we can repeat the process $\mathcal{O}(\log \frac{1}{\delta'})$ times to achieve arbitrary $\delta'$.



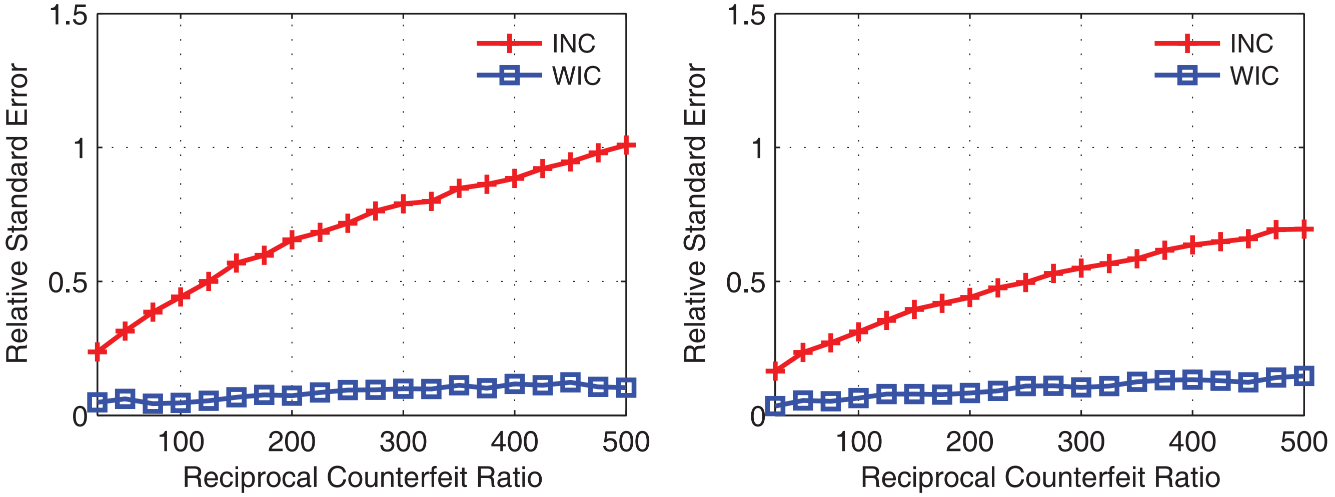
(a) Relative Standard Error. (b) Standard Deviation. (c) Norm Std Deviation.

**Figure 3: The accuracy of $|\hat{\mathcal{C}}|$ vs $|\mathcal{T}_1|$, when $|\mathcal{S}| = N = 10,000,000$, $c = 0.07$.**



(a) Relative Standard Error. (b) Standard Deviation. (c) Norm Std Deviation.

**Figure 4: The accuracy of $|\hat{\mathcal{G}}|$ vs $|\mathcal{T}_1|$, when $|\mathcal{S}| = N = 10,000,000$, $c = 0.07$.**

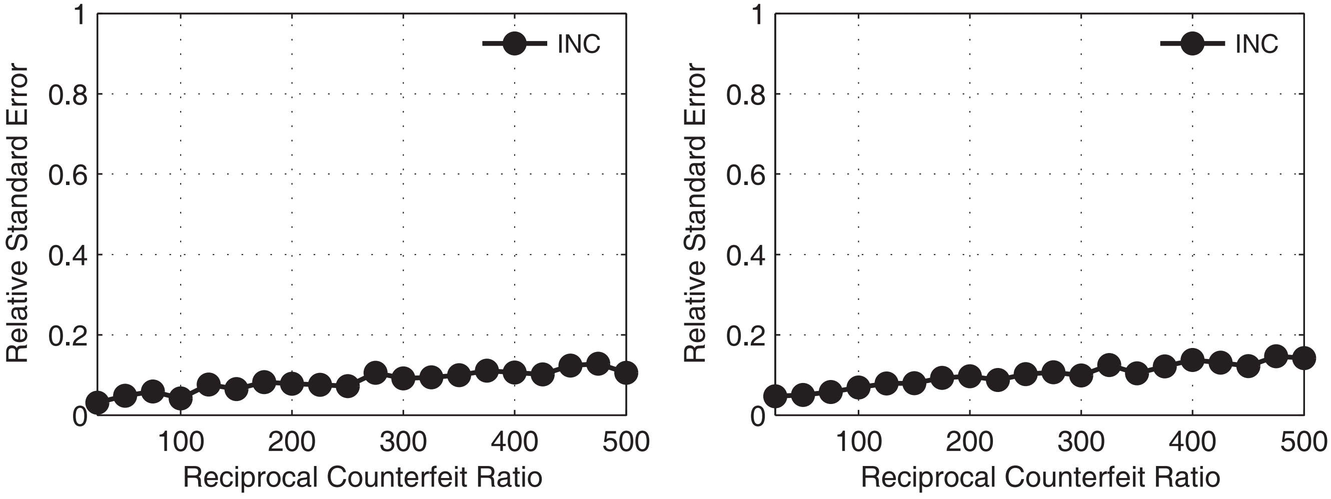


(a) Fixed total transmission cost at 20KB. (b) Fixed frame size at 51200.

**Figure 5: The accuracy $|\mathcal{T}_1 - \mathcal{S}|$ vs Reciprocal Counterfeit Ratio, when $|\mathcal{S}| = 10,000,000$, $|\mathcal{T}_1| = 10,000,000$ .**

reported statistics in [29]. As stated in Section 4, $f$ is the size of frame in the first layer.

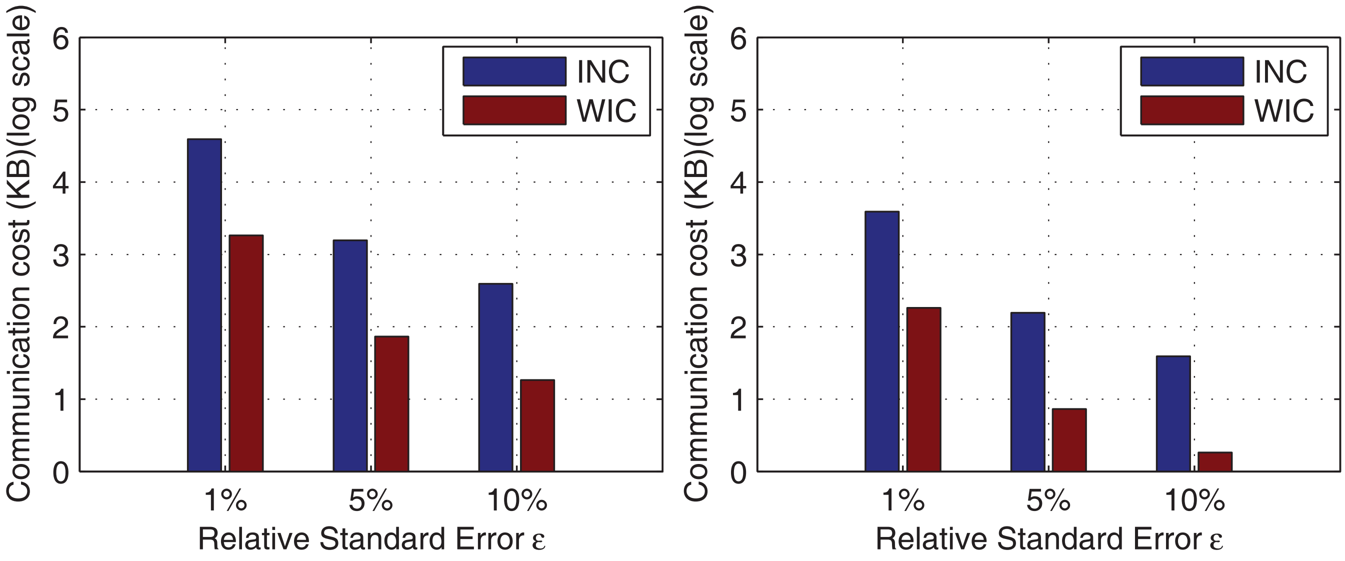
Figure 3 shows the result of estimator $|\hat{\mathcal{C}}|$. From those figures, we make several observations. First, the increase of frame size improves RSE of $|\hat{\mathcal{C}}|$. In particular, when the number of tags is 1,000,000, the RSE is 0.73 with f=32 and drops quickly to 0.10 with f=2048. Second, RSE remains almost unchanged as the number of verified tags grows. This nice probability confirms our former complexity analysis: the RSE is mainly related to frame size $f$ and counterfeit ratio $c$, not the number of tags to be verified. It thus provides us scalable authentication scheme for large-scale RFID



(a) Fixed total transmission cost at 20KB.

(b) Fixed frame size at 51200.

**Figure 6: The accuracy $|((\mathcal{T}_1 \cup \mathcal{T}_2) \cap \mathcal{T}_3) - \mathcal{S}|$ vs Reciprocal Counterfeit Ratio, when $|\mathcal{S}| = 10,000,000$, $|\mathcal{T}_1 \cup \mathcal{T}_2 \cup \mathcal{T}_3| = 10,000,000$ .**



(a) with different relative error $\varepsilon$ and $c = 0.1$.

(b) with different relative error $\varepsilon$ and $c = 0.01$.

**Figure 7: Communication cost between reader and tags (in *log scale*), when $|\mathcal{S}| = 10,000,000$, $|\mathcal{T}_1| = 50,000,000$.**

system. Third, SD and NSD of $|\hat{\mathcal{C}}|$ are also greatly reduced by larger frame size.

Similar trend can also be seen in Figure 4. But there is a major difference. Contrast to RSE of $|\hat{\mathcal{C}}|$, the RSE of $|\hat{\mathcal{G}}|$ is diminishing as $|\mathcal{T}_1|$ increases . This trend stems from the increasing genuines ratio due to increased $|\mathcal{T}_1|$, since the number of counterfeits is static.

**Single Set Authentication:** Here we study how counterfeits ratio affects accuracy of estimators. We vary the reciprocal counterfeit ratio and keep the total transmission size and frame size fixed. As shown in Figure 5a, we use total transmission size budget for 20KB. The result shows that our WIC is always advantageous over INC. For example, when reciprocal counterfeit ratio is 100, the RSE of WIC is 0.06 while RSE of INC is 0.44. The low performance of INC is mainly due to that they fail to examine the correlation information. The secondary factor may be more bits per slot used than WIC. This is confirmed by result in Figure 5b. We fix the frame size at 51200 to compare those methods, ignoring different lengths of time slot. The performance of INC is still worse than WIC, although the gap of RSE between INC and WIC is smaller than that in Figure 5a.

**Multiple Sets Authentication:** The settings for multiple set authentication are all same as in single set. The only difference is we need to authenticate a compound set expression: $|((\mathcal{T}_1 \cup \mathcal{T}_2) \cap \mathcal{T}_3) - \mathcal{S}|$. Note that for multiple sets INC does not have solution so far, so it is not included for comparison here. Figure 6a reports the RSE as a function of reciprocal counterfeits ratio. It shows that our WIC is still able to give accurate counterfetis estimate as in the single set situation, which is consistent with our asymptotic results in Section 5.2.

**Communication Cost:** We compare three methods given $\varepsilon$ changing from 1% to 15% and the counterfeit ratio at 0.01 and 0.1 in single set situation. The communication cost results of three schemes are depicted in Figure 7. We assume perfect communication channel between reader and tags. The slot length for INC is 32 bits. Again, we can see that WIC achieves much lower communication cost than INC. For instance, total transmission cost of WIC is merely 4.7% of INC when $c = 0.01, \varepsilon = 0.01$. From a different perspective, above comparison results indicate that given a certain time budget, the authentication accuracy of WIC will still much better than INC.

# 8. RELATED WORK

To authenticate a batch of tags, there are two main categories of methods: deterministic and probabilistic. For small-scale RFID systems, deterministic methods that typically combine anti-collision protocol with per-tag authentication scheme work quite well [30]. The anti-collision protocols includes ALOHA-based EPCGlobal C1G2 standard [10] and tree-based ISO 18000-6 standard [31]. To securely verify a single tags, various schemes are designed according to different purposes, e.g., system anonymity [32], anticloning [33], hash-based authentication [11], key management using tree structure [12][13]. The best known efficiency $\mathcal{O}(1)$ is achieved by Lu et al. by employing a weak privacy model [34]. Those per-tag based approaches, however, suffer from severe scalability issue when the population of tags rapidly increases, since the complexity of deterministic scheme should be linear with the number of tags.

The first probabilistic batch authentication for large-scale RFID systems, SEBA, is proposed in [14]. By using single echo from each tag, SEBA is able to successfully detect counterfeits with high probability if the fraction of counterfeits exceeds predefined threshold. Since the result of SEBA only indicates whether there is counterfeit and is lack of further information, a fine-grained and scalable batch authentication scheme, INC, is introduced to accurately approximate the count of counterfeits [16]. From statistical perspective, we discover most previous methods fail to explore information in authentication data structures, leaving the correlation information and a great deal of authentication synopses unused. In contrast, WIC is statistically efficient in sense of Cramer-RAO lower bound. Moreover, those former methods can only efficiently authenticate one tag set whereas WIC still performs well with multiple tag sets.

Besides above probabilistic batch authentication schemes, a number of probabilistic cardinality estimation approaches are proposed. Kodialam et al. introduce the first probabilistic RFID estimation algorithm using linear counting. Unified Probabilistic Estimator (UPE) and Unified Probabilistic Estimator (UPE) are proposed to deal with different frame-size constraints [26]. An asymptotically unbiased estimator, EZB, uses the number of zero slot in the frame to track the dynamics of tag population in both time and spatial domain [25]. Han et al. propose to use first non-empty slot observation to quickly estimate the cardinality [35]. Qian et al. propose LOF, which borrows the idea of FM-Sketch to geometrically approximate the count of tags [36]. Zheng et al. advance the estimation efficiency from $\mathcal{O}(\log n)$ by LOF to $\mathcal{O}(\log\log n)$, where $n$ is the number of tags to be estimated [37]. Average Run based Tag Estimation (ART) method uses a new statistical entity, average run length, to achieves 7x faster then UPE [38]. The first energy-efficient estimation scheme is proposed by Li et al [39]. Gong et al propose the first arbitrarily accurate approximation scheme for cardinality estimation using $t$-wise independent hash functions [40]. A comprehensive understanding of RFID counting protocols can be found in [41], which emphasizes the use of two-phase schemes. Although these algorithms can effectively derive the estimate of total number of tags in interested region, they are unable to distinguish counterfeit tags from genuine tags. They are thus not appropriately applicable in RFID authentication systems.

# 9. CONCLUDING REMARKS

In this paper, we propose a fast and efficient batch authentication scheme for large-scale RFID systems with optimal efficiency. First, we formally defined the general counterfeits estimation problem with multiple tag sets. Then we introduce a novel hierarchical authentication structure to efficiently encode tag sets. By using this statistically efficient data structure, WIC achieves $\mathcal{O}(\frac{N}{\varepsilon^2|C|}\log\log N)$ authentication complexity. By theoretical analysis and extensive simulation comparisons, we show that WIC is significantly advantageous over previous state-of-the-art methods. We are currently working on adapting WIC to off-the-shelf commercial tags. In the future, we also plan to provide systematic implementations of WIC using software defined radios as we already have some initial designs.

# 10. ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for valuable and insightful comments. This work is supported in part by the NSFC Young Scholar No.61303196, NSFC Distinguished Young Scholars No.61125202 and Natural Sciences and Engineering Research Council (NSERC) of Canada grant No.CRDPJ-445199-12. We also acknowledge the support from the Open RFID Lab (ORL) project [42].

# 11. REFERENCES

[1] Frost & Sullivan. Global RFID Healthcare and Pharmaceutical Market. *Industry Report*, 2011.

[2] Frost & Sullivan. Global RFID market. *Industry Report*, 2011.

[3] How RFID Is Transforming VA Hospital Operations. *www.rfidjournal.com/article/view/9819*, 2012.

[4] J. Hoepman, E. Hubbers, B. Jacobs, M. Oostdijk, and R. Schreur. Crossing Borders: Security and Privacy Issues of the European e-Passport. pages 152–167, 2006.

[5] J. Collins. Hong Kongạ́s Airport to Tag Bags. *RFID Journal*, 2004.

[6] M. Pelino, C. Mines, J. Warner, and S. Musto. M2M Connectivity Helps Telcos Offset Declining Traditional Services. *Forrester Research*, 2011.

[7] R. Want. RFID–A Key to Automating Everything. *Scientific American*, 290(1):56–65, 2004.

[8] D. Ma, C. Qian, W. Li, J. Han, and J. Zhao. GenePrint: Generic and Accurate Physical-layer Identification for UHF RFID Tags. In *Proc. of IEEE ICNP*, 2013.

[9] J. Han, C. Qian, D. Ma, X. Wang, J. Zhao, P. Zhang, W. Xi, and Z. Jiang. Twins: Device-free Object Tracking using Passive Tags. In *Proc. of IEEE INFOCOM*, 2014.

[10] EPCglobal Radio-Frequency Identity Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860 MHz-960MHz, 2008.

[11] S. Weis, S. Sarma, R. Rivest, and D. Engels. Security and Privacy Aspects of Low-Cost Radio Frequency Identification Systems. *Security in pervasive computing*, pages 50–59, 2004.

[12] L. Lu, J. Han, L. Hu, Y. Liu, and L.M. Ni. Dynamic Key-Updating: Privacy-Preserving Authentication for RFID Systems. In *Proc. of IEEE PERCOM*, 2007.

[13] L. Lu, J. Han, R. Xiao, and Y. Liu. ACTION: Breaking the Privacy Barrier for RFID Systems. In *Proc. of IEEE INFOCOM*, 2009.

[14] L. Yang, J. Han, Y. Qi, and Y. Liu. Identification-free Batch Authentication for RFID Tags. In *Proc. of IEEE ICNP*, 2010.

[15] The spread of counterfeiting: Knock-offs catch on. *The Economist*, 2010.

[16] W. Gong, K. Liu, X. Miao, Q. Ma, Z. Yang, and Y. Liu. Informative Counting Fine-grained Batch Authentication for Large-Scale RFID Systems. In *Proc. of ACM MOBIHOC*, 2013.

[17] B. Sheng, Q. Li, and W. Mao. Efficient Continuous Scanning in RFID Systems. In *Proc. of IEEE INFOCOM*, 2010.

[18] L. Xie, Q. Li, X. Chen, S. Lu, and D. Chen. Continuous Scanning with Mobile Reader in RFID Systems: an Experimental Study. In *Proc. of ACM MOBIHOC*, 2013.

[19] H. Liu, W. Gong, X. Miao, K. Liu, and W. He. Towards Adaptive Continuous Scanning in Large-Scale RFID Systems. In *Proc. of IEEE INFOCOM*, 2014.

[20] Y. Zheng and M. Li. P-MTI: Physical-layer Missing Tag Identification via Compressive Sensing. In *Proc. of IEEE INFOCOM*, 2013.

[21] F. Gandino, R. Ferrero, B. Montrucchio, and M. Rebaudengo. DCNS: An Adaptable High Throughput RFID Reader-to-Reader Anticollision Protocol . *Parallel and Distributed Systems, IEEE Transactions on*, 24(5):893–905, 2013.

[22] Y. Zheng and M. Li. Read bulk data from computational rfids. In *Proc. of IEEE INFOCOM*, 2014.

[23] H. Liu, W. Gong, L. Chen, W. He, K. Liu, and Y. Liu. Generic Composite Counting in RFID Systems. In *Proc. of IEEE ICDCS*, 2014.

[24] S.L. Garfinkel, A. Juels, and R. Pappu. RFID privacy: an overview of problems and proposed solutions. *Security Privacy, IEEE*, 3(3):34–43, May 2005.

[25] M. Kodialam, T. Nandagopal, and W.C. Lau. Anonymous Tracking Using RFID Tags. In *Proceedings of IEEE INFOCOM*, 2007.

[26] M. Kodialam and T. Nandagopal. Fast and Reliable Estimation Schemes in RFID Systems. In *Proc. of ACM MOBICOM*, 2006.

[27] H. Zhang, J. Gummeson, B. Ransford, and K. Fu. Moo: A Batteryless Computational RFID and Sensing Platform. *Tech Report UMASS*, 2011.

[28] Rajeev Motwani. *Randomized Algorithms*. Cambridge university press, 1995.

[29] ICC Counterfeiting Intelligence Bureau. Countering Counterfeiting: A Guide to Protecting and Enforcing Intellectual Property Rights. 1997.

[30] T.F. La Porta, G. Maselli, and C. Petrioli. Anticollision Protocols for Single-Reader RFID Systems: Temporal Analysis and Optimization. *IEEE Transactions on Mobile Computing*, 10(2):267 –279, 2011.

[31] Information technology Radio frequency identification for item management Part 6: Parameters for air interface communications at 860 MHz to 960 MHz, 2010.

[32] Y.K. Lee, L. Batina, and I. Verbauwhede. Untraceable RFID authentication protocols: Revision of EC-RAC. In *Proc. of IEEE RFID*, 2009.

[33] L. Bolotnyy and G. Robins. Physically Unclonable Function-Based Security and Privacy in RFID Systems. In *Proc. of IEEE PERCOM*, 2007.

[34] L. Lu, Y. Liu, and X. Li. Refresh: Weak Privacy Model for RFID Systems. In *Proc. of IEEE INFOCOM*, 2010.

[35] H. Han, B. Sheng, C.C. Tan, Q. Li, W. Mao, and S. Lu. Counting RFID Tags Efficiently and Anonymously. In *Proceedings of IEEE INFOCOM*, 2010.

[36] C. Qian, H. Ngan, Y. Liu, and L.M. Ni. Cardinality Estimation for Large-Scale RFID Systems. *IEEE Transactions on Parallel and Distributed Systems*, 22(9):1441–1454, 2011.

[37] Y. Zheng and M. Li. PET: Probabilistic Estimating Tree for Large-Scale RFID Estimation. *IEEE Transactions on Mobile Computing*, 11(11):1763–1774, 2012.

[38] M. Shahzad and A. Liu. Every Bit Counts - Fast and Scalable RFID Estimation. In *Proc. of ACM MOBICOM*, 2012.

[39] T. Li, S. S. Wu, S. Chen, and M. C. K. Yang. Generalized Energy-Efficient Algorithms for the RFID Estimation Problem. *IEEE/ACM Transactions on Networking*, 20(6):1978 –1990, 2012.

[40] W. Gong, K. Liu, X. Miao, and H. Liu. Arbitrarily Accurate Approximation Scheme for Large-Scale RFID Cardinality Estimation. In *IEEE INFOCOM*, 2014.

[41] B. Chen, Z. Zhou, and H. Yu. Understanding RFID Counting Protocols. In *Proc. of ACM MOBICOM*, 2013.

[42] Open rfid lab, http://pdcc.ntu.edu.sg/wands/orl.