Utilizing RF Interference to Enable Private Estimation in RFID Systems

1Lei Yang 2Zhuo Liu 1Jinsong Han 1Qingsong Yao 1Yong Qi 4Ying Chen 3Cheng Wang 4Xiao Zhong

1Xi’an Jiaotong University, China 2Hong Kong University of Science and Technology 3Tongji University, China 4IMB Research China

Abstract

Counting or estimating the number of tags is crucial for RFID system. Researchers have proposed several fast cardinality estimation schemes to estimate the quantity of a batch of tags within a short time frame. Existing estimation schemes scarcely consider the privacy issue. Without effective protection, the adversary can utilize the responding signals to estimate the number of tags as accurate as the valid reader. To address this issue, we propose a novel privacy-preserving estimation scheme, termed as MEAS, which provides an active RF countermeasure against the estimation from invalid readers. MEAS comprises of two components, an Estimation Interference Device (EID) and two well-designed Interference Blanking Estimators (IBE). EID is deployed with the tags to actively generate interfering signals, which introduce sufficiently large estimation errors to invalid or malicious readers. Using a secret interference factor shared with EID, a valid reader can perform accurate estimation via two IBEs. Our theoretical analysis and simulation results show the effectiveness of MEAS. Meanwhile, MEAS can also maintain a high estimation accuracy using IBEs.

Keywords - RFID, Cardinality Estimation, Privacy, Interference, Entropy

1. Introduction

Radio Frequency Identification (RFID) has been widely used in a variety of applications, such as logistics and supply chain management [1], access control [2], theft prevention [3], and movement tracking [4, 5] etc. An RFID system usually comprises of a number of RFID readers and tags. The RF based communication enables RFID systems to identify objects without keeping the tags within sight or wire-connected. This merit allows RFID to become a promising labeling technology replacing traditional Barcode technique in automatic identification and localization applications.

In common RFID systems, counting the number of tags within a region is important. Usually fast RFID identification schemes require a estimation of the actual number of tags, namely cardinality estimation. Cardinality estimation facilitates to set the optimal frame size for probabilistic identification schemes and compute the expected number of slots for deterministic identification schemes. Utilizing such an estimation, the reader can accelerate the singleton process [6] and improve the identification efficiency in RFID systems.

The cardinality estimation, however, is often viewed as private information, while existing methods fail to fully address the privacy issue. For instance, by merely eavesdropping in one frame, a malicious reader can accurately guess the quantity of products in a warehouse [1]. Even worse, current RFID tags, especially those passive tags, are designed to automatically reply the interrogation from readers. The competitors can estimate the sale amount just though observing the change of tag number. In this case, the enterprise may not know that the private information has been compromised if an attacker performs an active interrogation on tags. Many privacy-preserving solutions are proposed to protect privacy from attacks at the application level. They focus on identifying individual tags via deterministic identification [7-10], which is time consuming and unsuitable for cardinality estimation in terms of efficiency.

We propose a novel interference scheme to achieve privacy-preserving cardinality estimation. Our scheme, termed as Maximum Entropy based Adaptively Shaking scheme (MEAS), includes two components: (1) Estimation Interference Device (EID) deployed with the tags, and (2) Interference Blanking Estimators (IBE) deployed in valid readers. After each estimation, EID sends a secret interference factor to valid readers through a secure channel. Using the factor, a valid reader can conduct accurate estimation via two IBEs. The entire interference procedure of MEAS
is similar to encrypting the responses of tags so that any invalid reader cannot make accurate estimations due to the lack of the key, say the secret interference factor. The merit of MEAS is twofold. First, by well designing the interference factor and model, MEAS can prevent the malicious readers from estimating the cardinality of tags accurately. Second, MEAS maintains high estimation accuracy for legitimate readers via two carefully designed IBEs, ZE+ and CE+. With the rigorous analysis and extensive simulations, we demonstrate that MEAS provides effective protection on cardinality estimation of tags. MEAS does not require any modification on current tags and EID can be easily implemented in off-the-shelf products.

The rest of this paper is organized as follows. In Section 2, we briefly introduce the background of number estimation of tags. In Section 3, we present the MEAS scheme. In Section 4, we further optimize MEAS by identifying efficient interference factor. In Section 5, we propose interference blanking estimators enabling valid readers to accurately estimate. We present the performance evaluation in Section 6 and discuss related works in Section 0 and conclude this paper in Section 8.

2. Background

In our model, the RFID system comprises of a reader and a number of tags. In this paper, we adopt the framed-slotted ALOHA scheme to conduct the probabilistic estimation. Suppose $t$ tags involve in the estimation phase. At the beginning of the estimation phase, the reader sends a request with a frame of size $f$ slots. Upon the request, tags randomly select a slot uniformly at random, and reply in that slot. We assume the reader guarantee the slot synchronization by energizing probe/request. For facilitating theoretical analysis, we define indicator random variable $X_i$ to represent the event of NO tag replying in the $ith$ slot. If there is at least one tag transmitting the reply signal in the $ith$ slot, $X_i = 0$. otherwise, $X_i = 1$. If $X_i = 1$, the slot is termed as idle slot.

Similarly, we define a random variable $Y_i$ to represent the event that there is only one tag’s reply in the $ith$ slot, and $Z_i$ to indicate the event that there is a collision in the $ith$ slot (more than one tag reply). Correspondingly, if $Y_i = 1$ (or $Z_i = 1$ ), the slot is single slot (or collision slot). Note that $X_i + Y_i + Z_i = 1$ for any slot $i$.

Let $N_0 = \sum_{i=1}^{f} X_i$ denote that total number of idle slots, $N_t = \sum_{i=1}^{f} Y_i$ denote that total number of single slots, and $N_c = f - N_0 - N_t$ denote the total number of collision slots. Let $n_0$, $n_t$, and $n_c$ represent the values of $N_0$, $N_t$ and $N_c$ that are observed by the reader in a particular instance. Based on the expected values and the observed values, the reader is able to estimate $t$. We can employ three estimators proposed by M. Kodialam et al. [11]: Zero Estimator (ZE), Singleton Estimator (SE), and Collision Estimator (CE), shown in Table 1. Due to the monotonicity [11], only the ZE and CE are useful in the estimation. In practice, we can use these two estimators together to improve the accuracy of estimation [11]. We call this methodology as Joint Estimator (JE), e.g. USE. The results in [11] show that the estimation accuracy of two estimators is attractive. Note existing estimation approaches do not consider the privacy issue. Thus, the adversary can make estimation on the tags’ cardinality as accurate as the valid reader can.

3. Interference Scheme

In this section, we elaborate the design of MEAS scheme. The scheme comprises of three components: Estimation Interference Device (EID), interference model, and interference algorithm. We first present the design of EID. We then construct an interference model and utilize the model to inspire the design of our interference algorithm.

3.1 Estimation Interference Device

We deploy an EID with each batch of tags along the entire logistic flow. EID behaves like a normal tag, yet is more powerful. According to the interference algorithm, EID emits interfering signals in the entire estimation frame. Note the malicious reader cannot know the pattern of interference. The reader receives the interfering signals emitted from EID concurrently with the raw signals from the tags. Meanwhile, EID delivers a secret interference factor, which determines the interference pattern, to the valid reader through a secure channel. The channel can be a wired connection or an encrypted wireless connection. As a result, only valid readers that are aware of the interference factor can accurately estimate the cardinality of tags. Note that to confuse malicious reader the EID randomly produce ID every time.

3.2 Interference Model

In Section 2, we introduce two estimators, ZE and CE, which utilize the numbers of idle slots and collision slots to estimate the cardinality of tags. Intuitively, if we disturb them via interference, the estimation result will dramatically change. In this subsection, we analyze how the interference ratio (IR) influences the estimation error (EE). Let $n_0$ be the number of idle slots under interference, while the real number of idle slots is $n_0$; $n_c$ be the number of collision slots under
interference, while the real number of collision slots is \( n_c \).

**Definition 1 Interference Ratio (IR) and Estimation Error (EE):** Let \( \hat{t}_0^\prime \) denote the estimation result of tags using ZE, and \( \hat{t}_c^\prime \) denote the estimation result of tags using CE. We define \( r_0 = (n_0 - n_0)/n_0 \) as the interference ratio of idle slots and \( e_0 = (t_0^\prime - t)/t \) as the error of ZE. We define \( r_c = (n_c - n_c)/n_c \) as the interference ratio of collision slots and \( e_c = (t_c^\prime - t)/t \) as the error of CE.

Based on Definition 1, we find if \( e_0 \) (or \( e_c \)) equals 0, which means there is no interference, the \( r_0 \) (or \( r_c \)) equals 0. If \( e_0 \) (or \( e_c \)) < 0, the estimation under interference is smaller than the actual estimation, or vice versa. The objective of MEAS here is to make the absolute value of the estimation error as large as possible.

**Theorem 1:** Let the load factor \( \rho = t/f \), the relationship between the interference ratio and estimation error of ZE and CE is as following:

(a) For ZE:
\[
e_0 = -\ln(1 + r_0)/\rho
\]
where \(-1 \leq r_0 \leq (1 - \epsilon^\rho)/\epsilon^\rho\)

(b) For CE:
\[
(1 + r_c)(1 + \rho) - (1 + \rho + \rho e_c)\epsilon^{-\rho e_c} = r_c\epsilon^\rho
\]
where \(-1 \leq r_c \leq (1 + \rho)/(\epsilon^\rho - \rho - 1)\)

Proof:

a) According to ZE estimator formula in Table 1, we have
\[
f \epsilon^{-\frac{t_0}{\epsilon}} = n_0^\prime
\]
Since \( t_0^\prime = (1 + e_0)\frac{t}{\epsilon} \), \( n_0^\prime = (1 + r_0)n_0 \), and \( n_0 = f \epsilon^{-\frac{t}{\epsilon}} \), we can get \( e_0 = -\ln(1 + r_0)/\rho \)

Since \(-1 \leq n_0^\prime \leq f\), \( r_0 \) should be \(-1 \leq r_0 \leq (1 - \epsilon^\rho)/\epsilon^\rho\)

b) According to CE estimator formula in Table 1, we have
\[
f - t_c^\prime \epsilon^{-\frac{t_c}{\epsilon}} - f \epsilon^{-\frac{t_c}{\epsilon}} = n_c^\prime
\]
Since \( t_c^\prime = (1 + e_c)\frac{t}{\epsilon} \), \( n_c^\prime = (1 + r_c)n_c \) and \( n_c = f - t_c^\prime \epsilon^{-\frac{t_c}{\epsilon}} - f \epsilon^{-\frac{t_c}{\epsilon}} \), we have
\[
(1 + r_c)(1 + \rho) - (1 + \rho + \rho e_c)\epsilon^{-\rho e_c} = r_c\epsilon^\rho
\]
Since \(-1 \leq n_c^\prime \leq f\), \( r_c \) should be
\[-1 \leq r_c \leq (1 + \rho)/(\epsilon^\rho - \rho - 1)\]

This completes the proof. \( \square \)

According to Theorem 1, we depict above relationships in Figure 1. Without interference, IR equals 0. If we increase the number of idle or collision slots, IR > 0. We define these cases as positive interferences. Correspondingly, if we decrease the number of idle or collision slots, IR < 0. We define these cases as negative interferences. From the results, we have several useful observations, which can be leveraged to design effective interference schemes.

1) Increasing the absolute value of IR achieves better interference effects. This is because both the positive interference and negative interference can lead to estimation errors.

2) For ZE, the interference effects of the positive and negative interference are different. Using negative interference, a same increment of IR will incur a larger estimation error compared to using positive interference. This observation indicates that the interference effect of negative interference is better than that of positive interference towards ZE.

3) For CE, the observation is opposite to that of 2). The interference effect of positive interference is better than that of negative interference towards CE.

4) Based on 2) and 3), both the negative interference on ZE and positive interference on CE argument the error of estimation results. In addition, if we concurrently launch these two interference mechanisms, the interference effects of them will not counteract each other.

The above observations facilitate us to design the optimal interference scheme. Based on 2) and 3), we attempt to strength the negative interference on ZE and positive interference on CE for better interference effect. To this end, we prefer to utilize the ranges of IR as [-1, 0] for ZE and [0, 1] for CE. Correspondingly, interference will work on the second quadrant on ZE and first quadrant on CE in Figure 1, respectively. In practice, we can decrease the number of idle slots or increase the number of collision slots accordingly. The fourth observation motivates us to utilize the negative interference on ZE and positive interference on CE simultaneously. In this case, the overlapped interference would reinforce the interference effect.

**Definition 2 Guard Range:** Assume the minimum IR must be attained is \( \delta \) where \( \delta \in [-1, 1] \). The guard range of interference is defined as \( [-\delta, \delta] \).

Within the guard range, the interference scheme cannot guarantee to make estimation error sufficiently large. In real RFID systems, we can make use of \( \delta \) to tune the interference effect of interference scheme.

![Figure 1: IR vs. EE on CE and ZE](image-url)
3.3 Interference Algorithm of MEAS

We leverage the observations presented in previous subsection to design interference algorithm of MEAS. Intuitively, we can minimize the number of idle slots or maximize the number of collision slots for increasing IR. In the estimation frame, the numbers of slots, however, follow $n_0 + n_1 + n_c = f$. Note that EID does not know the distribution of $n_0$, $n_1$, and $n_c$ in the tags’ responses (indeed, there is no need for EID to do so). We can only use the interference signals generated by EID to change the numbers of slots of two tags’ responses.

A naive solution is to convert all $n_0$ idle slots to collision slots. Unfortunately, this is impossible because EID is not aware which slot is idle in the tags’ responses. Another simple solution for EID is to interfere every slot in the estimation phase. This converts all $n_0$ in tags’ responses to $n_1$, and all $n_c$ in tags’ responses to $n_c$. Such an interference mode is not secure. In this case, all idle slots convert to single slots and all single slots convert to collision slots under the interference. By observing the number of single slots, the adversary is able to know the number of idle slots in tags’ responses, and then utilize ZE to do estimations.

To amend this flaw, we introduce an interference factor, denoted as $\gamma$, to MEAS. The $\gamma$ is a probability. In each slot of the estimation frame, EID sends the interference with the probability $\gamma$. The $\gamma$ is used to effectively disturb the distribution of slots in a frame, while providing the ability of interference blanking to valid readers.

Note our interference scheme is not like jamming the wireless communications [12] between the tag and reader. EID is fully compatible with ALOHA based protocols. Therefore, the adversary cannot distinguish the raw signals of tags from the interference signals from EID. Since the $\gamma$ is only shared between EID and valid readers, the interfering process is similar to an encryption procedure. The ‘key’ of MEAS is the $\gamma$.

Based on above analysis, the crucial task of the interference algorithm is to select appreciate $\gamma$ and enable the interference radio IR beyond the guard range $\delta$, as we discussed in previous subsection. The $\gamma$ is highly related to the guard range $\delta$, frame size $f$, and the number of tags $t$. We will discuss their relationship in next section. In practice, it is easy to manually configure $\delta$ and retrieve $f$ from the reader’s estimate command. However, there are still several challenges when determining $\gamma$ and designing the interference algorithm:

1) EID does not know the $t$ in advance. The value of $t$ is unknown at the beginning of estimation phase, and hence poses difficulty to set proper $\gamma$. Another common scenario is that the number of tags may frequently change. For example, an update of stock in a retailing warehouse often changes $t$. Our strategy is to allow EID randomly to select a $\gamma$ at initialization. After each estimation, EID adaptively tunes the $\gamma$ according to the last estimation result.

2) How to evaluate the interference effect? The adversary may randomly guess the number of slots no matter how EID interferes. As long as the guessed number of idle or collision slots are close to the actual value, the adversary will also accurately estimate the number of tags. We adopt entropy to evaluate the interference uncertainty. A maximum entropy model will facilitate the determination of $\gamma$. We will detail the entropy based evaluation and the determination $\gamma$ in next section.

3) The adversary may have some prior knowledge about the $t$, for example, a supplier wants to track the quantity changes of a batch of items in the wholesaler and retailer. In this case, the supplier knows IR after the first estimation. Then he can obtain the $\delta$. If the $\delta$ is fixed, he may deduce the $t$ in subsequential estimations even if the $t$ changes. Therefore, we should let $\delta$ dynamically shake after each estimation. Based on above analysis, we propose our Maximum Entropy based Adaptively Shaking scheme (MEAS), as shown in Algorithm 1.

4. Interference Optimization

To optimize MEAS, we adopt entropy to measure the uncertainty of interference. In this section, we first define the rule of interference, and then define the entropy of adversary’s estimators for measuring the effectiveness and security of MEAS. We also illustrate the relationship between interference factor $\gamma$ and the entropy, which guides us to determine an optimal $\gamma$.

---

**Algorithm 1: MEAS interference algorithm**

**INPUT:** Guard range $\delta$, Frame size $f$, Tag number $t$

**INTERFERENCE PROCEDURE**

1: IF receive estimation command
2: ELSE IF last estimation $t == 0$
3: Randomly select $\gamma$
4: ELSE
5: Obtain random decimal $r$
6: value $= \delta + (1 - \delta) \cdot r$ // make $\delta$ shaking
7: Calculate $\gamma$ according to maximum entropy.\
8: END
9: ELSE IF slot $i$ arrives
10: Produce random decimal $k$ that value between 0 and 1
11: IF $k < \gamma$ // the probability of $k < \gamma$ is $\gamma$
12: Emit interference
13: ELSE
14: Keep silent
15: END
16: ELSE
17: Sending the interference factor $\gamma$
18: END
4.1 Interfering Rule

The objective of selecting a proper interference factor is twofold. First, we expect to introduce sufficiently large errors to adversary estimation via EID’s interference. Second, we must guarantee the interference pattern is unable to be guessed by malicious attackers.

For simplifying the description, we use ‘0’ to denote the idle slot, ‘1’ denotes the single slot, and ‘X’ denote the collision slot. In this way, the slots of a frame can be described as a vector with f elements and every element has a value as 0, 1, or X. We define the vector of tags’ signals as raw vector, the vector of EID’s signals as interference vector, and the vector of signals finally received by a reader as mimic vector. The relationship of them is mimic vector = raw vector \( \Pi \) interference vector, where \( \Pi \) is a special operating rule defined in Table 2. In the table, the dark row represents the slot types of raw vector and the dark column represents the slot types of interference vector. Note that an interference vector only contains ‘0’ and ‘1’ that means whether the EID does not interfere or interfere.

Figure 2 illustrates an example of interference vector. For easy illustration, we deliberately group the elements according to the values of ‘0’, ‘1’, and ‘X’ in the raw vector. In this example, the raw vector is [0 0 0 0 1 1 1 X X X] and the interference vector is [0 0 1 1 0 1 1 0 1 0]. Thus the mimic vector received by the reader is [0 0 1 1 1 X X X X X] according to the \( \Pi \) rule.

The goal of adversary is to guess the numbers of ‘0’s and ‘X’s in the raw vector. Obviously, based on our interference scheme the adversary knows that the number of ‘0’s in the mimic vector is smaller than the number of ‘0’s in the raw vector, and the number of ‘X’ in the mimic vector is larger than the number of ‘X’ in the raw vector.

![Figure 2: Interference example](image)

<table>
<thead>
<tr>
<th>raw vector</th>
<th>interference vector</th>
<th>mimic vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0 0 0 1 1 1 X X X</td>
<td>0 0 1 1 1 0 1 1 0 1 0</td>
<td>0 0 1 1 1 X X X X X</td>
</tr>
</tbody>
</table>

\( X_0 = l \), \( X_1 = k \), \( X_2 = m \)

4.2 Measuring the Uncertainty of Adversary’s ZE Estimator

An interference scheme should keep sufficient uncertainty of the mimic vector towards adversary’s estimators. In this paper, we adopt entropy to describe the uncertainty. We use the indicator random variable \( X_0 \) to denote the number of ‘0’s, \( X_1 \) to denote the number of ‘1’s, and \( X_2 \) to denote the number of ‘X’s in the mimic vector, respectively. Thus, we have

\[
0 \leq X_0 \leq n_0, \\
n_0 - X_0 \leq X_1 \leq n_0 + n_1, \\
n_0 - X_0 - X_1 \leq X_2 \leq n_0 + n_1 + n_2, \\
\]

and \( X_0 + X_1 + X_2 = f \), where \( n_0, n_1, \) and \( n_2 \) are the number of ‘0’s, ‘1’s, and ‘X’ in the raw vector. We analyze the entropy of adversary’s ZE, and the ZE estimator.

If \( X_0 = l \) and \( X_1 = k \), there are \( l \) non-interfered ‘0’s, \( (n_0 - l) \) interfered ‘0’s; \( k \cdot (n_0 - l) \) non-interfered ‘1’s, and \( n_0 - (k \cdot (n_0 - l)) \) interfered ‘1’s in the raw vector. As the example in Figure 2, \( l = 2 \), \( k = 3 \), \( n_0 = 4 \), \( n_1 = 3 \), and \( n_2 = 3 \), thus, there are 2 non-interfered ‘0’s, 2 interfered ‘0’s, 1 non-interfered ‘1’s, and 2 interfered ‘1’s in the raw vector. The probability of \( X_0 = l \) and \( X_1 = k \) is given by:

\[
Pr(X_0 = l, X_1 = k) = \binom{n_0}{l} \cdot (1 - \gamma)^{n_0 - l} \cdot \gamma^{l} \\
= \binom{n_0}{l} \cdot (1 - \gamma)^{n_0 - l} \cdot \gamma^{l - (n_0 - l)} \\
= \binom{n_0}{l} \cdot \frac{\gamma^{l - k}}{(1 - \gamma)^{2l + k - n_0}} \\

\]

When the adversary uses ZE, the adversary attempts to guess the number of ‘0’s in the raw vector via the mimic vector. Obviously, the adversary knows the ‘0’ number of the raw vector is no less than \( X_0 \).

Based on the \( \Pi \) rule, the ‘0’s in the raw vector will be converted to ‘1’s under the interference. When adversary receives a ‘0’ in the mimic vector, he can immediately know that the values of those slots in both the raw vector and interference vector are ‘0’. The adversary will be confused when receiving a ‘1’ in the mimic vector, which may directly come from the raw vector without interference or be converted from a ‘0’ in the raw vector due to the overlying of a ‘1’ of interference vector. Hence, according to maximum entropy principle, if the adversary has no any knowledge about the interference vector, his best choice is to make a random guess on whether the slot ‘1’ is interfered or not. In this case, the probability of success is 0.5 [13].

In fact, the adversary does not need exactly guess the numbers of ‘0’s or ‘1’s in the raw vector via the mimic vector, because it is also acceptable to the adversary if he restricts the error of numbers of ‘0’s or ‘1’s within a reasonable range. Here, we use the guard range as such a range. Under the adversary’s guess, the total number of ‘0’s in the raw vector equals the number of ‘0’s in the mimic vector plus the number of ‘1’s that are converted from ‘0’s in the raw vector.
Table 3: MEAS IBEs

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZE+</td>
<td>( f(1 - \gamma)e^{\frac{-t}{\tau}} = n_a )</td>
</tr>
<tr>
<td>CE+</td>
<td>( f - (t + f - \gamma t)e^{\frac{-t}{\tau}} = n_c )</td>
</tr>
</tbody>
</table>

Table 4: Simulation setup

<table>
<thead>
<tr>
<th>Case</th>
<th>Number of Tags</th>
<th>Frame Size</th>
<th>Guard Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>50</td>
<td>100</td>
<td>0.1</td>
</tr>
<tr>
<td>II</td>
<td>500</td>
<td>300</td>
<td>0.2</td>
</tr>
<tr>
<td>III</td>
<td>5000</td>
<td>3000</td>
<td>0.3</td>
</tr>
<tr>
<td>IV</td>
<td>50000</td>
<td>30000</td>
<td>0.4</td>
</tr>
</tbody>
</table>

**Theorem 2:** Given a guard range with \( \delta \), the probability that the adversary guesses the correct number of ‘0’ s via the mimic vector is the guard range is denoted by \( P_{ZE}(l, k, \delta) \). Then

\[
P_{ZE}(l, k, \delta) = \sum_{i=[(1-\delta)n_0]-1}^{\lfloor (1+\delta)n_0 \rfloor} C_k^i 0.5^k,
\]

where \( l \) is the number of ‘0’ s and \( k \) is the number of ‘1’ s in the mimic vector.

Proof: Let \( X_6 \) denote the number of ‘0’ s that the adversary guesses via the mimic vector. Then

\[
P_{ZE}(l, k, \delta) = \Pr \left( [(1-\delta)n_0 - 1]) \leq l \leq (1+\delta)n_0 \right)
\]

\[
= \sum_{i=[(1-\delta)n_0]-1}^{\lfloor (1+\delta)n_0 \rfloor} C_k^i 0.5^k
\]

This completes the proof. \( \square \)

We use the average entropy to measure the uncertainty of adversary’s ZE. We then compute the average entropy of adversary’s ZE, denoted as \( \omega_{ZE} \).

**Definition 3 Uncertainty of Adversary's ZE:** The entropy of adversary’s ZE, denoted by \( H_{ZE} \), can be computed as follows:

\[
H_{ZE}(l, k, \delta) = -E_{ZE}(l, k, \delta) \log_2 E_{ZE}(l, k, \delta) + (1 - E_{ZE}(l, k, \delta)) \log_2 (1 - E_{ZE}(l, k, \delta));
\]

then the uncertainty of adversary’s ZE \( \omega_{ZE} \) is

\[
\omega_{ZE}(\delta, n_0, n_1) = \sum_{l=0}^{n_0} \sum_{h=n_0-l}^{n_0+n_1-1} \Pr(X_0 = l, X_1 = k) H_{ZE}(l, k, \delta)
\]

The analysis based on the entropy gives us a guideline to set appropriate interference factor \( \gamma \). Obviously, a higher entropy leads to a larger uncertainty to adversary’s estimators. The optimal \( \gamma \) for ZE is the interference factor that is able to maximize the \( \omega_{ZE} \).

5. **Interference Blanking Estimators**

Due to the interference, the previous ZE and CE estimators cannot be used in MEAS. To enable the valid reader perform accurate estimation, MEAS provides two interference blanking estimators, ZE+ and CE+, for canceling interference in cardinality estimation. Note that ZE+ and CE+ only work after retrieving interference factor from EID. We design new expect formulas in Theorem 5 that involves the interference factor for designing ZE+ and CE+.

**Theorem 5:** If MEAS performs an interference, the expects of \( N_0, N_1, \) and \( N_C \) are as follows.

\[
E[N_0] \approx f(1 - \gamma)e^{\frac{-t}{\tau}} + (1 - \gamma)te^{\frac{-t}{\tau}}
\]

\[
E[N_1] \approx f - (t + f - \gamma t)e^{\frac{-t}{\tau}}
\]

Proof: Slot \( i \) will be idle if none of tags reply and EID also keeps silent in slot \( i \). Therefore,

\[
E[N_0] = \sum_{t=1}^{L} \Pr(X_i = 1) = f \left( \frac{1}{1 - \gamma} \right)^{t} \approx f(1 - \gamma)e^{\frac{-t}{\tau}} + (1 - \gamma)te^{\frac{-t}{\tau}}
\]

\[
E[N_1] = \sum_{t=1}^{L} \Pr(Y_i = 1) = f \left( \frac{1}{1 - \gamma} \right)^{t-1} \approx f - (t + f - \gamma t)e^{\frac{-t}{\tau}}
\]

This completes the proof. \( \square \)

Based on Theorem 5, we can achieve two new estimators: ZE+ and CE+. Indeed, ZE+ and CE+ have the capability of anti-interference from EID, and provide the accurate estimation to valid readers. The formula is shown in Table 3.

6. **Performance Evaluation**

In this section, we evaluate the performance of MEAS through comprehensive simulations.

6.1 **Simulation Methodology and Metrics**

In our simulation, the reader can exactly detect all the slots in a frame. Each tag has 96 bit ID, which is a standard setting in EPC Gen [14]. In the four test cases, the number of tags ranges from 50 and 50000, which represents different application scenarios in real RFID systems. Each simulation is executed 1000 rounds with the similar parameters. The simulation setup is shown in Table 4. To examine the effect of MEAS, we employ the following two metrics: normalized error (NE) and normalized variance (NV). Let \( t_i' \) is the estimated number of tags by the estimator in \( i \)th experiment and \( t \) is the actual number of tags, we define the normalized error (NE) and normalized variance (NV) as

\[
NE = \frac{\left| \frac{1}{m} \sum_{i=1}^{m} t_i' - t \right|}{t}
\]

\[
NV = \frac{1}{m} \sum_{i=1}^{m} \frac{|t_i' - t|}{t}
\]
only shows the estimated results using CE, CE+, and ACE for case I in Figure 8. Figure 9 illustrates the comparisons between AZE and ACE and shows the interference of EID is always effective on AZE than ACE.

7. Related Work

Utilizing collision to launch cardinality estimation on RFID tags has been extensively studied. Z. Zhou et al. analyze three types of collisions in multi-reader environments: tag-tag collision, reader-tag collision, and reader-reader collision [15, 16]. All cardinality estimators mainly employ the first collision [17, 18]. Kodialam and Nandagopal [11] propose two estimation algorithms, Unified Simple Estimator (USE) and Unified Probabilistic Estimator (UPE). As a replicate-sensitive estimation protocol, LoF proposed by Chen. Q [19] addresses the multiple-reading problem in large-scale RFID systems. On the other hand, privacy issue is also critical in RFIDs. Many approaches have been proposed to achieve private authentication in RFID system [8-10, 16, 20, 21]. All of these works focus on authenticating individual tag, while we mainly consider the privacy for a batch of tags. To the best of our knowledge, the most similar work to MEAS is the block tag [7], which is capable of protecting tags within a specific zone. The block tag, however, is based on tree based anti-collision protocols and belonging to deterministic algorithms. Thus, the block tag is unsuitable for cardinality estimation. The works in [22] introduce the ‘privacy-masking’ mechanism to protect RF communication. Those works merely works on the physical layer and cannot support the privacy protection for cardinality estimation.

8. Conclusion

In this paper, we focus on the privacy issue of tags’ cardinality estimation in RFID systems. We propose a novel scheme, MEAS, to utilize RF interference for achieving private cardinality estimation. MEAS employs an interfering device EID to disturb the estimation of invalid readers, and provides accurate estimation for valid readers via two interference blanking estimators, ZE+ and CE+.

Table 5: Normalized Error of ZE, ZE+ and AZE

<table>
<thead>
<tr>
<th>Estimators</th>
<th>Case I</th>
<th>Case II</th>
<th>Case III</th>
<th>Case IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZE</td>
<td>0.0026</td>
<td>0.0034</td>
<td>0.000256</td>
<td>0.000129</td>
</tr>
<tr>
<td>ZE+</td>
<td>0.028</td>
<td>0.017</td>
<td>0.0000553</td>
<td>0.0023</td>
</tr>
<tr>
<td>AZE</td>
<td>1.9768</td>
<td>0.65</td>
<td>0.7998</td>
<td>0.9081</td>
</tr>
</tbody>
</table>

Table 6: Normalized Variance of ZE, ZE+ and AZE

<table>
<thead>
<tr>
<th>Estimators</th>
<th>Case I</th>
<th>Case II</th>
<th>Case III</th>
<th>Case IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZE</td>
<td>0.0026</td>
<td>0.0444</td>
<td>0.0144</td>
<td>0.000129</td>
</tr>
<tr>
<td>ZE+</td>
<td>0.0281</td>
<td>0.1384</td>
<td>0.0434</td>
<td>0.0044</td>
</tr>
<tr>
<td>AZE</td>
<td>1.9768</td>
<td>0.7</td>
<td>0.8038</td>
<td>0.8038</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENT

This work is supported in part by National Natural Science Foundation of China (NSFC) (No.60933003, No.60736016, No.60873262, and No.60903155), National High Technology Research and Development Program of China (863 Program) under Grants No.2009AA01Z116, National Basic Research Program of China (973 Program) under Grants No. 2011CB302705 and No. 2010CB328004, China Postdoctoral Science Foundation funded project (No.20090461298), Hong Kong Innovation and Technology Fund GHP/044/07/LP and ITP/037/09/LP, the Science and Technology Research and Development Program of Shaanxi Province under Grant No.2008KW-02, and IBM Joint Project.

REFERENCES

[20] Q. Yao, Y. Qi, J. Han, J. Zhao, X. Li, and Y. Liu, "Randomizing RFID Private Authentication," in Proceedings of IEEE PerCom, 2009