RFID Cardinality Estimation with Blocker Tags

Xiulong Liu, Bin Xiao, Keqiu Li, Jie Wu, Alex X. Liu, Heng Qi and Xin Xie

Presenter: Dr. Bin Xiao
The Hong Kong Polytechnic University, Hong Kong
csbxiao@comp.polyu.edu.hkt
Background & Motivation

- **Radio Frequency Identification.**
- An identification system that consists of chip-based tags, readers, and a back-end.
- Each tag has a unique 96-bit ID to identify the tagged object.
RFID Background

- Two types of RFID tags:
  - Passive tags and Active tags
RFID Background

RFID vs. Bar-code

• Advantages of RFID over bar-code:
  □ remote access
  □ non-line-of-sight reading
  □ multiple simultaneous accesses
  □ large rewritable memory
Background & Motivation

Supply Chain Management

Object Tracking

Environment Monitoring

Pets Management

Anti-Counterfeit
The widely-used RFID tags impose serious privacy concerns.

Reason: When C1G2 tags are interrogated by an RFID reader, no matter whether the reader is authorized or not, they blindly respond with their IDs and other stored information (such as manufacturer, product type, and price) in a broadcast fashion.
What woman wants her dress size to be publicly readable by any nearby scanner?

Who wants the medications and other contents of a purse to be scannable?

Who wants his or her location to be tracked and recorded based on the unique ID number in their shoes or other clothing?

An effective solution to this privacy issue is to use commercially available blocker tags.
What are blocker tags?

A blocker tag is an RFID device that is preconfigured with a set of known RFID tag IDs, which we call blocking IDs. The blocker tag behaves as if all tags with its blocking IDs are present.
How blocker tags protect the privacy?

A blocker tag protects the privacy of the set of genuine tags whose IDs are among the blocking IDs of the blocker tag because any response from a genuine tag is coupled with the simultaneous response from the blocker tag; thus, the two responses always collide and attackers cannot obtain private information.

The genuine tag always collides with the blocking tag having the same ID
We are concerned with the problem of RFID (population size) estimation with the presence of blocker tags.  

**Problem Definition:** given (1) a set of unknown genuine tags $G$ of unknown size $g$, (2) a blocker tag with a set of known blocking IDs $B$, (3) a required confidence interval $\alpha \in (0,1]$, and a required reliability $\beta \in [0,1)$, we want to use one or more readers to estimate the number of genuine tags in $G$, denoted as $\hat{g}$, so that $P\{|\hat{g} - g| \leq g\alpha\} \geq \beta$
Problem Formulation

- To the best of our knowledge, this paper is the first to investigate RFID estimation with the presence of a blocker tag.
- None of the existing estimation schemes considers the presence of a blocker tag. Furthermore, none of them can be easily adapted to solve this problem.
Problem Formulation

- How about turning off the blocker tag and then using prior RFID estimation schemes to estimate the number of genuine tags?

- Turning off the blocker tag will give attackers a time window to breach privacy, especially for the scenarios in which RFID estimation schemes are being continuously performed for monitoring purposes.
REB Protocol

- RFID Estimation scheme with Blocker tags
- The communication protocol used by REB is the standard *framed slotted Aloha protocol*. 
**REB Protocol**

- **Detailed Steps:**
  - **Step 1:** the reader broadcasts a value \( f \) and a random number \( R \) to query all tags (including blocker tags), where \( f \) is the number of slots in the forthcoming frame. Then, each tag computes a hash \( H(ID, R) \% f \) to select a slot to respond.
REB Protocol

• Detailed Steps:
  • **Step 1:** the reader broadcasts a value $f$ and a random number $R$ to query all tags (including blocker tags), where $f$ is the number of slots in the forthcoming frame. Then, each tag computes a hash $H(ID, R) \% f$ to select a slot to respond.

- 0 represents no tag responds
- 1 represents only one tag responds
- 2+ represents two or more tags simultaneously respond and create a collision
REB Protocol

- **Step2**: As we know the blocking IDs, we can virtually execute the framed slotted Aloha protocol using the same frame size $f$ and random number $R$ for the blocking IDs; thus, we get another vector.
• **Step 2**: As we know the blocking IDs, we can virtually execute the framed slotted Aloha protocol using the same frame size \( f \) and random number \( R \) for the blocking IDs; thus, we get another vector.

- 0 represents no tag chooses this slot.
- 1 represents only one tag chooses this slot.
- 2+ represents two or more tags choose this common slot.
**REB Protocol**

- **Step 3**: we count two numbers: $N_{00}$, which is the number of slot $i$ such that both $V_1[i] = 0$ and $V_2[i] = 0$, and $N_{11}$, which is the number of slots $i$ such that both $V_1[i] = 1$ and $V_2[i] = 1$.

The Key Insight:
- The smaller $N_{00}$ is, the larger $|B \cup G|$ is.
- The larger $N_{11}$ is the larger $|B - G|$ is.
We theoretically proved that $N_{00}$ monotonously decreases with the increase of $|B \cup G|$; and $N_{11}$ monotonously increases with the increase of $|B - G|$.

Therefore, from the observed values of $N_{00}$ and $N_{11}$, we can estimate $|B \cup G|$ and $|B - G|$, respectively. Then, we can calculate the number of genuine tags, i.e., $|G| = |B \cup G| - |B - G|$.
REB Protocol

- **Practical Issue**: The frame size should be set as no more than 512. To scale to a large tag population, the reader uses a persistence probability $p \in (0, 1]$ to virtually extend the frame size $f$ to $f/p$, but actually terminates the frame after the first $f$ slots.

- Fundamentally, each tag participates in the actual frame of $f$ slots with a probability $p$. 
Theoretical Analysis

- **Functional Estimator:**

  \[ \hat{g} = -\frac{f}{p} \ln \left( \frac{N_{00}}{f} \right) - \frac{fN_{11}}{pN_{00}} \], where \( f \) is the observed frame size, \( p \) is the persistence probability, \( N_{00} \) is the number of persistent empty slots, \( N_{11} \) is the number of persistent singleton slots.
Theoretical Analysis

- Variance of the Estimator:

\[ \text{Var}(\hat{g}) = \frac{1}{fp^2} e^\frac{u_p}{f} (b'^2p^2 + f^2 - b'fp) - \frac{f}{p^2}, \]

where \( f \) is the observed frame size, \( p \) is the persistence probability, \( u = |B \cup G| \), and \( b' = |B - G| \).
Theoretical Analysis

• Refined Estimation with \( k \) Frames:
• We repeat \( k \) independent frames with different seeds, and use the average estimation result \( \widehat{g}_k' = \frac{1}{k} \sum_{j \in [1,k]} \widehat{g}_j \) to refine the estimation of REB, where \( \widehat{g}_j \) is the estimate derived from the \( j \)-th frame.
Theoretical Analysis

• **Termination Condition:**
• If the frame number $k$ satisfies: $k \geq Z_{\beta} \frac{\alpha u_{p_j}}{f_j} \sqrt{\sum_{j \in [1,k]} \left[ \frac{1}{f_j p_j^2} e^{f_j} \left( b'^2 p_j^2 + f_j^2 - b' f_j p_j \right) - \frac{f_j}{p_j^2} \right]}$,

where $f_j$ and $p_j$ are the frame size and persistence probability used in the $j$-th frame.
Avoiding Premature Termination:

$k \geq \frac{Z_\beta}{g_\alpha} \sqrt{\sum_{j \in [1,k]} \left[ \frac{u_p_j}{f_j} e^{f_j} \left( b'^2 p_j^2 + f_j^2 - b' f_j p_j \right) - \frac{f_j}{p_j^2} \right]}$.

If we directly use the estimated values $\hat{b}', \hat{u}, \hat{g}$ to calculate the R.H.S. of this inequality, $k$ may have a chance to be larger than it, which is not true and REB will have a premature termination.
Theoretical Analysis

- \( \delta \)-sigma method to avoid premature termination.
- When calculating the R.H.S. of the termination inequality, we use the upper/lower bounds on \( b', u, g \).
  - Upper bounds: \( \hat{x} \uparrow = \hat{x} + \delta \sqrt{Var(\hat{x})} \);
  - Lower bounds: \( \hat{x} \downarrow = \hat{x} - \delta \sqrt{Var(\hat{x})} \);
- Here, \( x \) could be \( b', u, \) or \( g \).
- Three-sigma rule indicates \( \delta = 3 \) is large enough.
• **Optimization**: frame size $f$ and persistence probability $p$.

• **For the first frame**, we simply set $f = 512$ and $p = \frac{512}{\hat{u}}$, where $\hat{u}$ is the number of total tags that can be fast estimated by the existing estimation protocols, e.g., ART [Mobicon 12].

• **For the other frames**, we can leverage the information obtained from previous frames to optimize $f$ and $p$. 

---

**Theoretical Analysis**
Theoretical Analysis

- **Optimization:** the Persistence Probability \( p \)
- **For a fixed frame size** \( f \), the goal of optimizing \( p \) is to minimize the estimation variance \( \text{Var}(\hat{g}) \).

\[ \text{Var}(\hat{g}) \text{ is a convex function of } p \]

**The first-order derivation**

\[ \frac{\partial \text{Var}(\hat{g})}{\partial p} = 0 \]

**Algorithm 1:** Optimizing \( p_{x+1} \) for the \((x+1)^{th}\) frame.

- **Input:** \( \hat{g}, \hat{b}, \hat{q}, \hat{y}, \text{ and } f \)
- **Output:** The optimized \( p_{x+1} \) for the \((x+1)^{th}\) frame.

1. \( \delta = 0.0001; \)
2. \( p_{\text{low}} = \frac{1}{\gamma}; \)
3. \( p_{\text{high}} = 1; \)
4. **while** \( p_{\text{high}} - p_{\text{low}} > \delta \) **do**
5. \( p = (p_{\text{low}} + p_{\text{high}})/2; \)
6. **Calculating** \( \frac{\partial \text{Var}(\hat{g})}{\partial p} \) in Eq. (22);
7. **if** \( (\frac{\partial \text{Var}(\hat{g})}{\partial p} > 0) \) **then**
8. \( p_{\text{high}} = p; \)
9. **else**
10. \( p_{\text{low}} = p; \)
11. **end if**
12. **end while**
13. \( p_{x+1} = (p_{\text{low}} + p_{\text{high}})/2; \)
14. return \( p_{x+1} \).
Theoretical Analysis

• Optimization: the frame size \( f \)

• We target finding an optimal \( f \) to minimize the expected remaining execution time.

• Minimize \((f + 1) \times y\)

• s.t. \( x + y \geq \frac{Z_{\beta}}{g\alpha} \sqrt{\sum_{j \in [1,x]} \text{Var}(\hat{g}_j) + y\text{Var}(\hat{g})} \)

• \( f \in \{2,4,8,16,\ldots,512\} \)

• Here, \( x \) is the number of frames that have already been executed. \( y \) is the number of frames that need to be further executed.
Performance Evaluation

1. Verifying the Optimized $f$ and $p$.

The values of $f$ and $p$ approach their overall optimal values after a few frames.
Performance Evaluation

• 2. Estimation Reliability.

Our REB ($\delta = 1$) can meet the required accuracy under different simulation settings.
Performance Evaluation

3. Time Efficiency: Impact of $|U|$

The ratio of three types of IDs is fixed to 1:1:1. The total tag number $|U|$ varies.

When $|U|=50000$, our REB runs 33x faster than the fastest tag identification protocol.
4. Time Efficiency: Impact of Tag Ratio

Our REB persistently runs tens of times faster than the existing protocols.
Conclusion

• We take the first step to address the problem of RFID estimation with Blocker tags.
• The proposed REB protocol is compliant with the commodity EPC C1G2 standard, and does not require any modifications to off the-shelf RFID tags.
• REB can guarantee any degree of estimation accuracy specified by the users.
• Extensive simulation results reveal that REB is tens of times faster than the fastest identification protocol with the same accuracy requirement.
Thanks for your attention!

Q & A