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

Outline

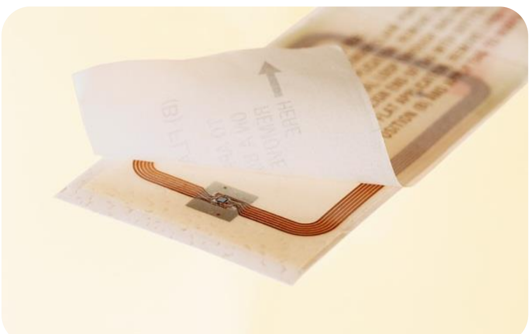
1	Background & Motivation
2	Problem Formulation
3	RFID Estimation scheme with Blocker tags (REB)
4	Theoretical Analysis
5	Performance Evaluation
6	Conclusion

Background & Motivation

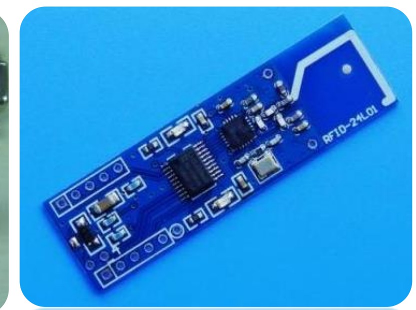
- Radio Frequency Identification.
- An identification system that consists of chip-based tags, readers, and a back-end.
- Each tag has a unique 96-bit ID to identify the tagged object.

RFID Background

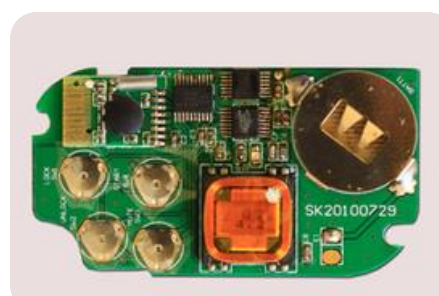
- Two types of RFID tags:
  - ☐Passive tags and Active tags

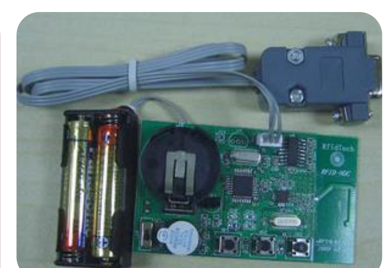


Passive tags

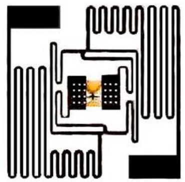


Active tags

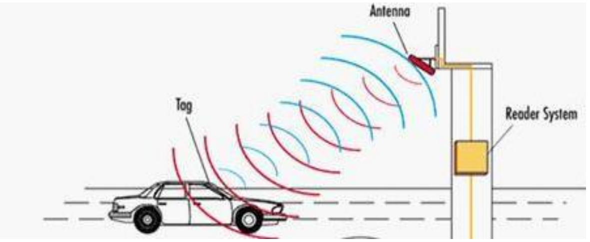


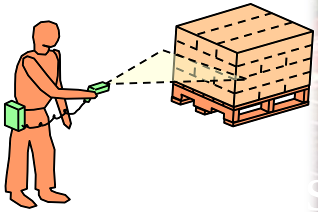


RFID Background

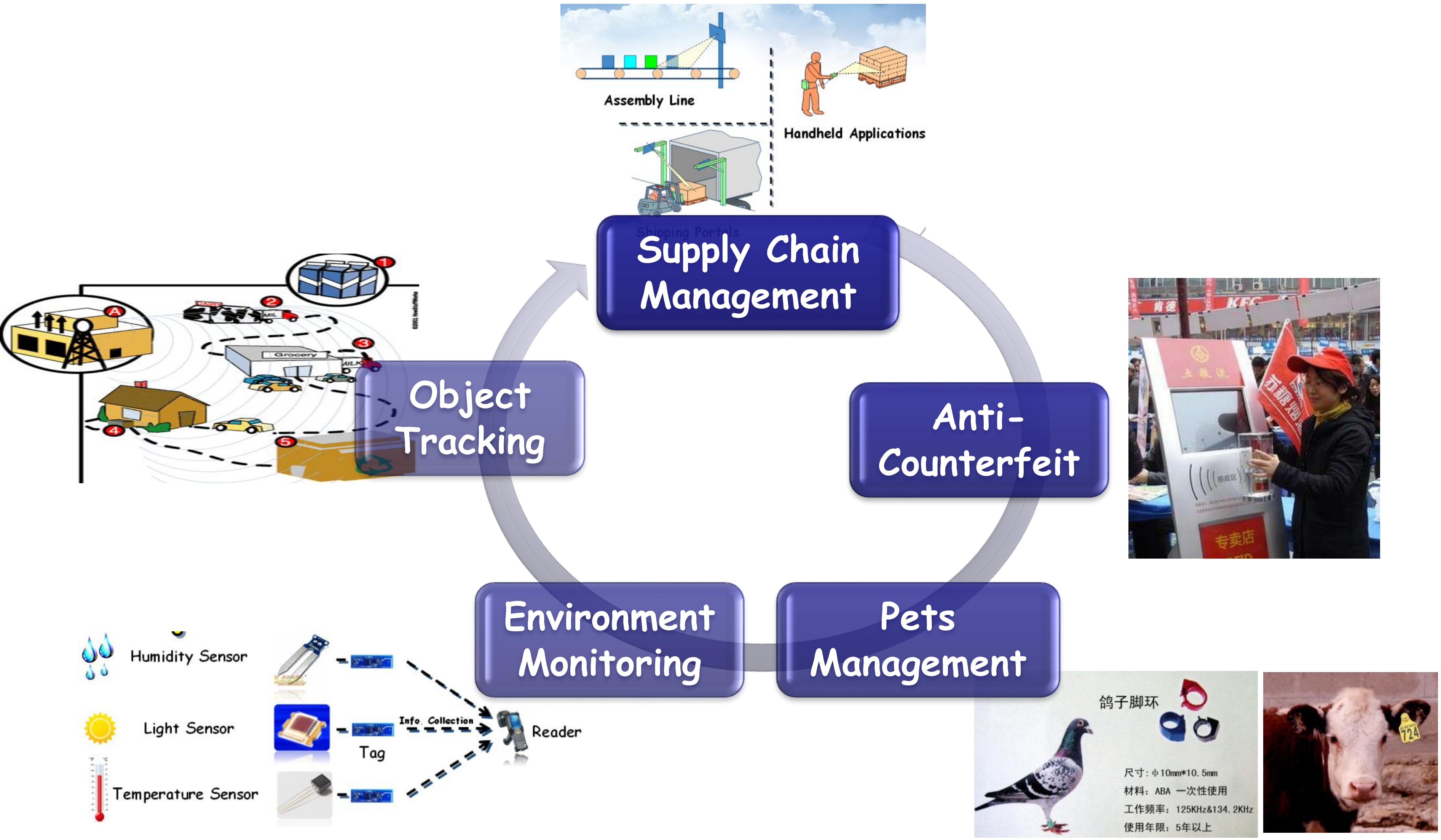
RFID  *vs.* Bar-code

- Advantages of RFID over bar-code:
  - ☐remote access
  - ☐non-line-of-sight reading
  - ☐multiple simultaneous accesses
  - ☐large rewritable memory





Background & Motivation

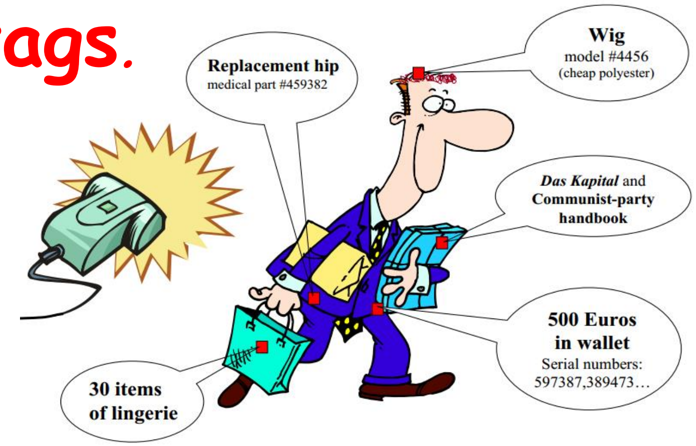


Background & Motivation

- 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.

Background & Motivation

- **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**.



Background & Motivation

- **What are blocker tags?**

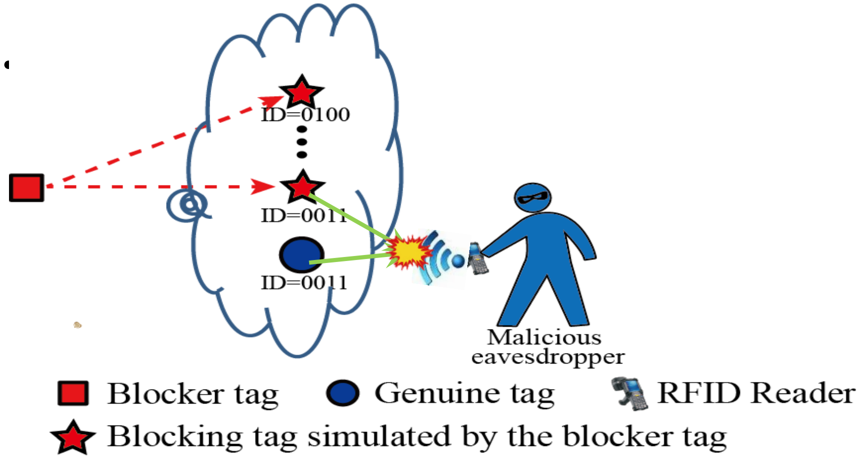
- A blocker tag is an RFID device that is preconfigured with a set of known RFID tag IDs, which we call blocking IDs. The blocker tag behaves as if all tags with its blocking IDs are present.

Background & Motivation

- **How blocker tags protect the privacy?**

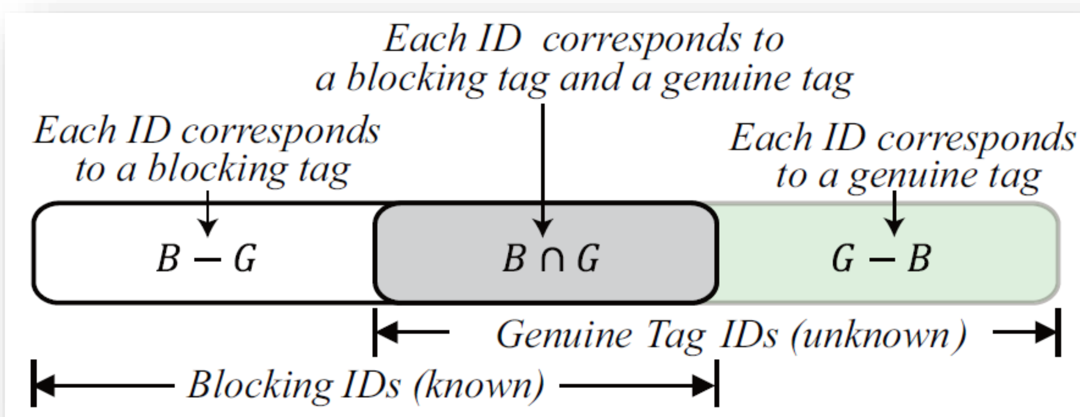
- A blocker tag protects the privacy of the set of genuine tags whose IDs are among the blocking IDs of the blocker tag because any response from a genuine tag is coupled with the simultaneous response from the blocker tag; thus, the two responses always collide and attackers cannot obtain private information.

The genuine tag always collides with the blocking tag having the same ID



Problem Formulation

- 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

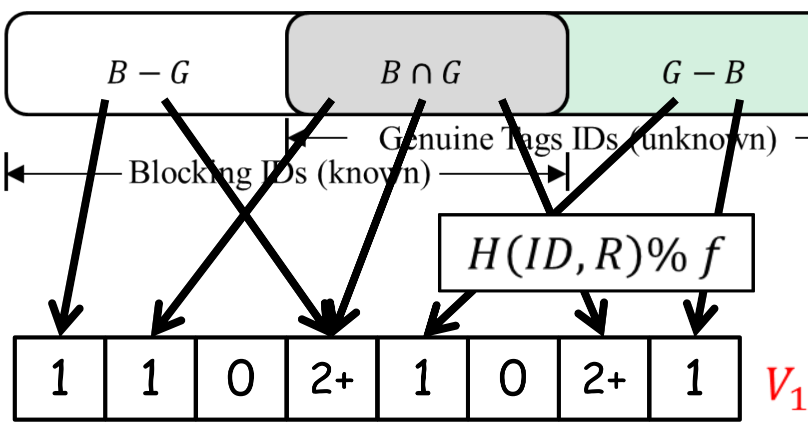
- **R**FID **E**stimation scheme with **B**locker tags
- The communication protocol used by REB is the standard *framed slotted Aloha protocol*.

REB Protocol

- Detailed Steps:
- **Step1:** 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:
- **Step1:** 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.



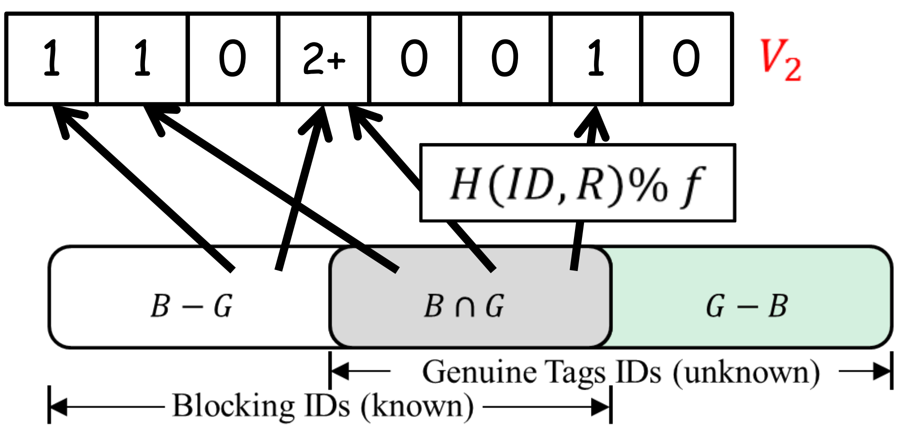
- 0 represents no tag responds
- 1 represents only one tag responds
- 2+ represents two or more tags simultaneously respond and create a collision

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.

REB Protocol

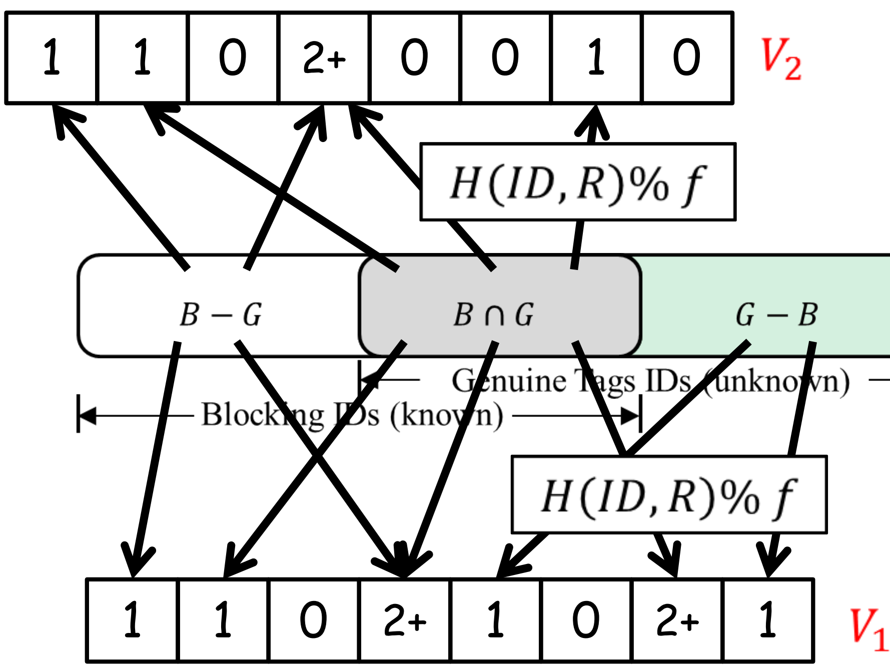
- **Step2**: As we know the blocking IDs, we can virtually execute the framed slotted Aloha protocol using the same frame size f and random number R for the blocking IDs; thus, we get another vector.



- 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

- **Step3**: 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.

REB Protocol

- 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., $|G| = |B \cup \mathrm{G}| - |B - G|$.

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:**
- $Var(\hat{g}) = \frac{1}{fp^2} e^{\frac{up}{f}} (b'^2 p^2 + f^2 - b' f p) - \frac{f}{p^2}$, 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: $k \geq \frac{z_\beta}{g\alpha}\sqrt{\sum_{j\in[1,k]}[\frac{1}{f_j p_j^2}e^{\frac{up_j}{f_j}}\left(b'^2 p_j^2 + f_j^2 - b' f_j p_j\right) - \frac{f_j}{p_j^2}]}$, where f_j and p_j are the frame size and persistence probability used in the j-th frame.

Theoretical Analysis

- **Avoiding Premature Termination:**

$$k \geq \frac{z_\beta}{g\alpha}\sqrt{\sum_{j\in[1,k]}[\frac{1}{f_j p_j^2} e^{\frac{up_j}{f_j}}(b'^2 p_j^2 + f_j^2 - b' f_j p_j) - \frac{f_j}{p_j^2}]},$$

If we directly use the estimated values $\widehat{b'}, \hat{u}, \hat{g}$ 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

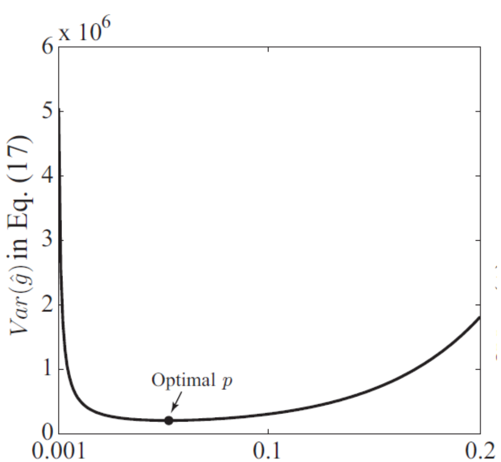
- δ-sigma method to avoid premature termination.
- When calculating the R.H.S. of the termination inequality, we use the upper/lower bounds on b', u, g.
- Upper bounds: $\hat{x} \uparrow = \hat{x} + \delta\sqrt{Var(\hat{x})}$;
- Lower bounds: $\hat{x} \downarrow = \hat{x} - \delta\sqrt{Var(\hat{x})}$,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

Theoretical Analysis

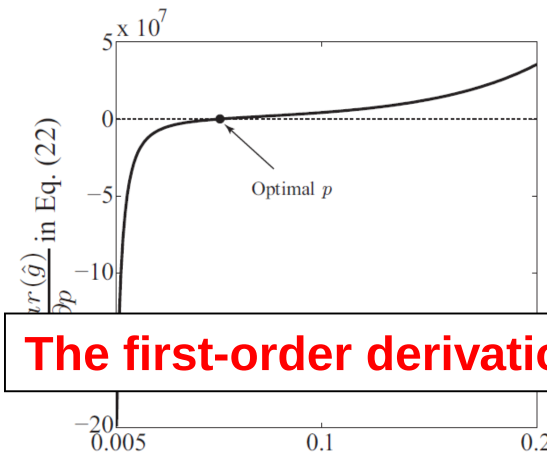
- **Optimization:** frame size f and persistence probability p.
- For the first frame, we simply set $f = 512$ and $p = \frac{512}{\hat{\boldsymbol{u}}}$, where $\hat{\boldsymbol{u}}$ is the number of total tags that can be fast estimated by the existing estimation protocols, e.g., ART [Mobicom 12].
- For the other frames, we can leverage the information obtained from previous frames to optimize f and p.

Theoretical Analysis

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



$Var(\hat{g})$ is a convex function of p



The first-order derivation

The optimal p makes $\frac{\partial Var(\hat{g})}{\partial p} = 0$

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

Input: $\hat{u}_{\overline{x}}$, $\hat{b'}_{\overline{x}}$, $\hat{g}_{\overline{x}}$, and f.
Output: The optimized p_{x+1} for the $(x+1)^{th}$ frame.

$\delta = 0.0001$;
$p_{low} = \frac{1}{\hat{u}_{\overline{x}}}$;
$p_{high} = 1$;
while $p_{high} - p_{low} > \delta$ **do**
  $p = (p_{low} + p_{high})/2$;
  Calculating $\frac{\partial Var(\hat{g})}{\partial p}$ in Eq. (22);
  if $(\frac{\partial Var(\hat{g})}{\partial p} > 0)$ **then**
    $p_{high} = p$;
  else
    $p_{low} = p$;
  end if
end while
$p_{x+1} = (p_{low} + p_{high})/2$;
return p_{x+1};

Binary search algorithm

Theoretical Analysis

- **Optimization: the frame size f**
- We target finding an optimal f to minimize the expected **remaining execution time**.
- Minimize $(f+1) \times y$

The remaining execution time

- s.t. $x + y \geq \frac{z_\beta}{g\alpha}\sqrt{\sum_{j\in[1,x]} Var(\widehat{g_j}) + yVar(\hat{g})}$
- $f \in \{2,4,8,16, \dots, 512\}$
- 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.

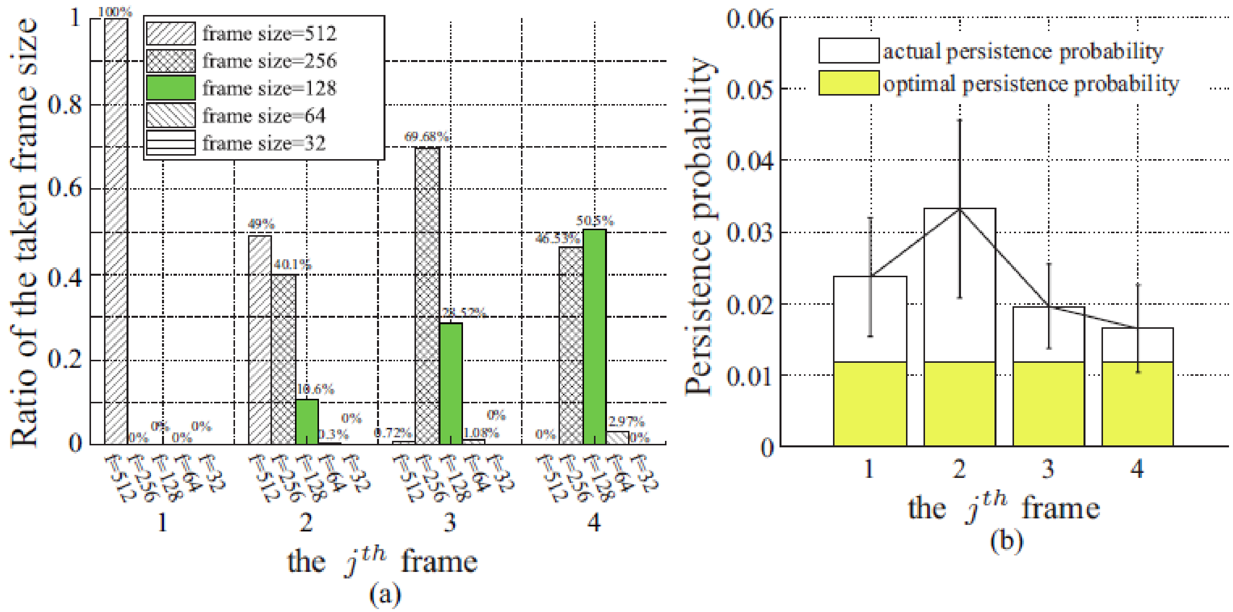


Fig. 3. Verifying the optimized settings of f and p. $|B-G| = 5000$, $|B \cap G| = 5000$, $|G-B| = 5000$. $\alpha = 10\%$, $\beta = 90\%$. (a) Verifying the optimized f. (b) Verifying the optimized p.

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

Performance Evaluation

- 2. Estimation Reliability.

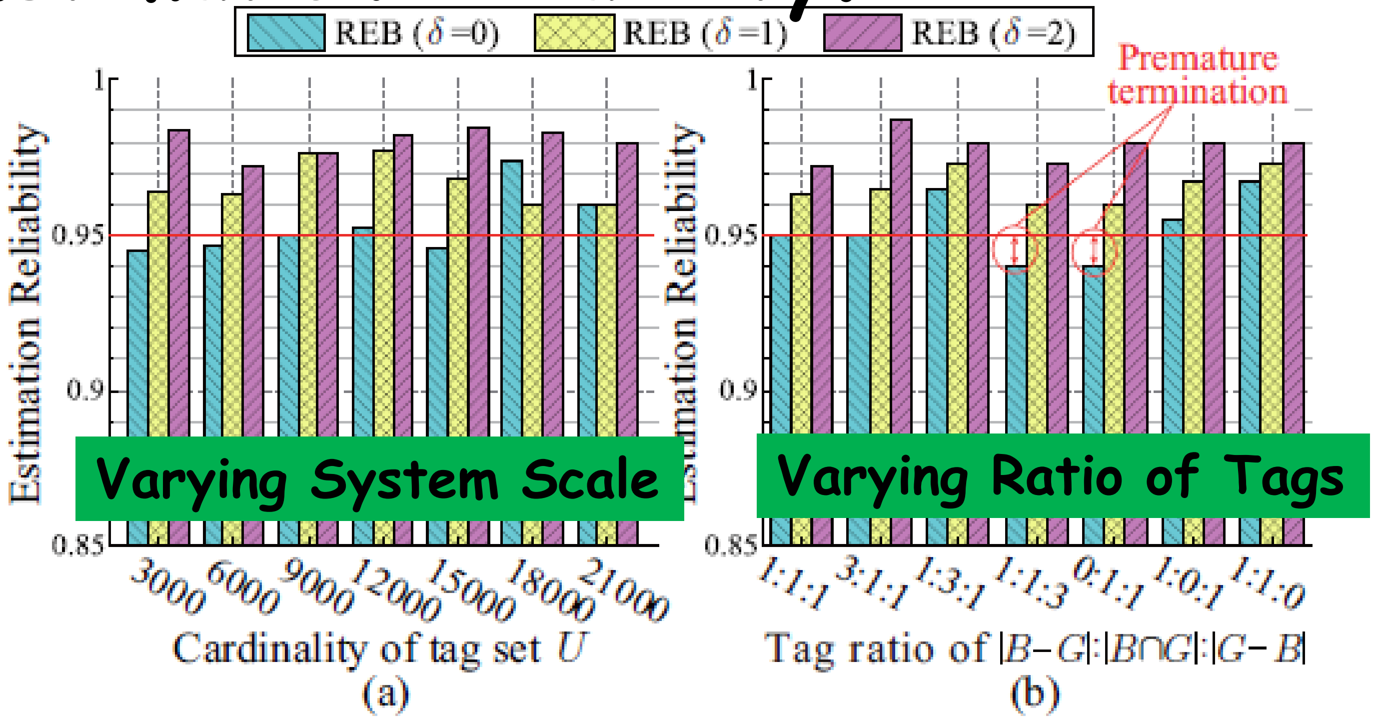


Fig. 4. Evaluating the reliability of REB. $\alpha = 5\%$, $\beta = 95\%$. (a) Tag ratio $|B - G|:|B \cap G|:|G - B|$ is fixed to $1:1:1$, and u varies from 3000 to 21000. (b) u is fixed to 9000, and tag ratio varies.

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

Performance Evaluation

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

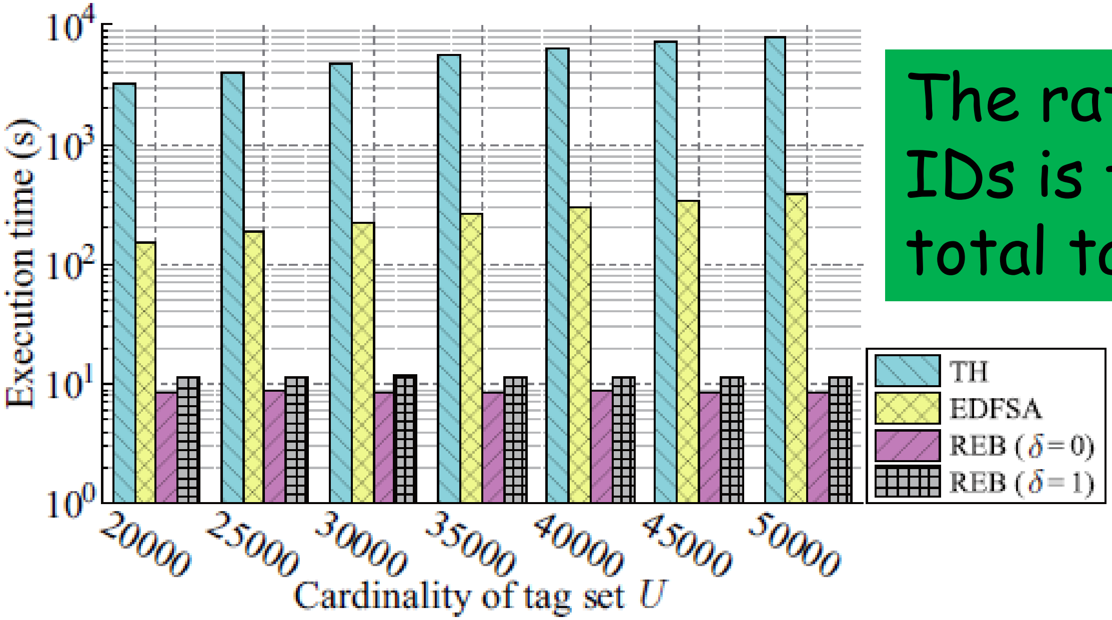


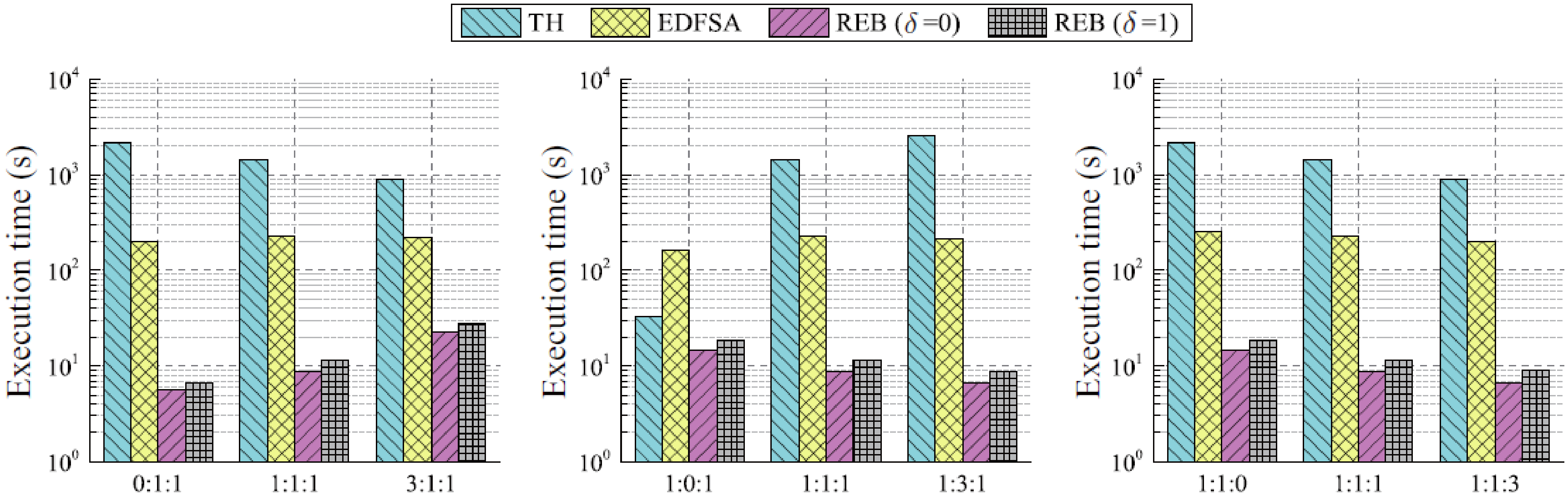
Fig. 5. Evaluating the time-efficiency of protocols with varying u. Tag ratio of $|B-G|$:$|B \cap G|$:$|G-B|$ is fixed to $1:1:1$ and $\alpha = 5\%$, $\beta = 95\%$.

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.

Performance Evaluation

- 4. Time Efficiency: Impact of Tag Ratio



Varying the ratio of $|B - G|$

Varying the ratio of $|B \cap G|$

Varying the ratio of $|G - B|$

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