Continuous Scanning with Mobile Reader in RFID Systems: an Experimental Study

Lei Xie†, Qun Li‡, Xi Chen†, Sanglu Lu†, and Daoxu Chen†
†State Key Laboratory of Novel Software Technology, Nanjing University, China
‡College of William and Mary, Williamsburg, VA, USA
lxie@nju.edu.cn, liqun@cs.wm.edu, hawkxc@163.com, {sanglu,cdx}@nju.edu.cn

ABSTRACT

In this paper, we show the first comprehensive experimental study on mobile RFID reading performance based on a relatively large number of tags. By making a number of observations regarding the tag reading performance, we build a model to depict how various parameters affect the reading performance. Through our model, we have designed very efficient algorithms to maximize the time-efficiency and energy-efficiency by adjusting the reader's power and moving speed. Our experiments show that our algorithms can reduce the total scanning time by 50% and the total energy consumption by 83% compared to the prior solutions.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Wireless Communication

Keywords

RFID; Realistic Settings; Algorithm Design; Experimental Study; Model

1. INTRODUCTION

Mobile RFID reading performance is critical to a number of applications that rely on mobile readers. Scanning books in a library or a bookstore, tracking merchandises in a store, all require a mobile reader to be used for continuous scanning over the tags attached to the physical goods and assets. The mobile reader moves continuously to scan a large number of tags effectively compensating for its limited reading range. In those types of mobile reader systems, two performance metrics are highly pertinent: time efficiency to reduce the total scanning time, and energy efficiency to reduce the total power consumption. Unfortunately, there is no realistic model to characterize the performance for mobile RFID reading for a large scale setting. The factors that affect the mobile reading performance are very complicated. For example, the actual scanning time for a number of tags in a realistic scenario is much longer than the time computed for free space, as shown in our experiments. In addition, RFID readers have a wide range of power selections, e.g., the Alien-9900 reader has a maximum power 30.7dBm, which is 30 times larger than the minimum power 15.7dBm. There is no guideline, however, in selecting a suitable power. Therefore, we aim to design an efficient solution to continuous scanning problem for a mobile RFID reader based on experimental study.

Although there have been some experimental studies on reading performance in a stationary RFID system [4, 1, 7], the previous studies have the following limitations. First, previous experiments were usually conducted in a small scale (fewer than 20 tags), which does not capture the complication for a large number of tags. Second, previous work has been focused on reading performance in a close to free space scenario. In reality, path loss, multi-path effect and mutual interference are common and have a big impact to RFID reading process. Third, previous work mainly examined how factors such as distance, coding scheme and frequency, affect reading performance. Very important factors, i.e., the reader's power and tag density, were neglected. Therefore, the previous work does not give a model for RFID reading process in a realistic and large scale setting; in particular, it does not include the power and tag density. Indeed, before we started our work, there was no realistic model which can guide us in designing an efficient tag identification solution in our setting.

We have, thus, conducted comprehensive measurements over a large number of tags in realistic settings by varying various parameters. Surprisingly, we have a few important new findings from the experiments. For example, we have found that the probabilistic backscattering is a ubiquitous phenomenon of the RFID system in realistic settings, i.e., during every query cycle each tag randomly responds with a certain probability, which has an important effect on the reading performance. This observation is contrary to the previous belief that tags respond to a reader with either probability 1 or 0. We have also found it is not wise to blindly increase the reader's power for tag identification, which can degrade the overall performance including the effective throughput and energy consumption. These findings are essential to improving reading performance for a mobile RFID system. Most importantly, we can (1) model the patterns of reading a large number of tags by giving a probabilistic model to capture the major and minor detection region, and (2) model how the reading power and tag density affect the reading performance by proving an empirical mapping.

MobiHoc'13, July 29–August 1, 2013, Bangalore, India.

Based on the effective models, we can then design efficient algorithms, which can dramatically improve the performance, as shown in our real experiments.

We make the following contributions in this paper. (1) We are the first to conduct an extensive experimental study over a relatively large number of tags (up to 240 tags) and a rather high tag density (up to 90 tags per square meter) in realistic settings. To the best of our knowledge, this is the first work to propose a model for investigating how the important parameters including reader's power, moving speed and tag density jointly affect the reading performance. (2) This is also the first work to give a framework of optimizing reading performance based on experimental study. We apply our model to solve the problem of continuous scanning with mobile reader. By carefully adjusting the power and moving speed, we design efficient algorithms to optimize time-efficiency and energy-efficiency. (3) We have a number of novel techniques in making our algorithms practical. For example, our tag density estimation method is extremely simple and accurate. Our algorithm extension to nonuniform tag density is also effective. (4) Being compatible with RFID standard (with no changes to the C1G2 protocols or low-level parameters for commercial RFID readers), our solutions can deliver significant performance gain. Experiment results indicate that, while achieving the same coverage ratio, our practical solutions respectively reduce scanning time by 50% and energy consumption by 83% compared to the prior solutions.

2. RELATED WORK

In RFID systems, a reader needs to receive data from multiple tags. These tags are unable to self-regulate their radio transmissions to avoid collisions. In light of this, a series of slotted ALOHA-based anti-collision protocols [19, 14, 9, 22], as well as tree-based anti-collision protocols [12, 13, 2], are designed to resolve collisions in RFID systems. In order to deal with the collision problems in multi-reader RFID systems, scheduling protocols for reader activation are explored in [18], [21]. Recently, a number of polling-based protocols [10, 5, 23, 3] are proposed, aiming to collect information from RFID tags in a time/energy efficient approach. In order to estimate the number of tags without collecting tag IDs, a number of protocols are proposed [8, 11, 15, 6, 17] to leverage the information gathered in slotted ALOHA protocol for fast estimation of the number of tags.

In order to verify the impact of the physical layer's unreliability, Buettner et al. [4] examine the performance of the C1G2 RFID system in a realistic setting. Aroor et al. [1] identify the state of the technical capability of passive UHF RFID systems using a simple, empirical, experimental approach. Jeffery et al. [7] conduct experiments in realistic settings and find that within each reader's detection range, a large difference exists in reading performance. In order to efficiently identify RFID tags in mobile settings, Xie et al. propose a probabilistic model to set optimized parameters for mobile tag identification [20]. Sheng et al. develop efficient schemes for continuous scanning operations [16], aiming to utilize the information gathered in the previous scanning operations to reduce the scanning time of the succeeding ones. Being different from the previous work, this paper conducts an extensive experimental study over a large scale RFID deployment, and proposes an effective model to depict the regularities of reading performance in realistic settings.

3. PROBLEM FORMULATION

We consider a typical scenario of continuous scanning in realistic settings, i.e., using a mobile reader to identify a large volume of tags deployed over a wide area. We respectively consider a situation where the tags are continuously placed with a uniform/nonuniform density, we seek to execute continuous scanning over the tags along a certain direction. The performance metrics in our consideration are as follows:

- *Time-efficiency*: considering it is time-consuming to identify a large volume of tags in realistic settings, the overall scanning time should be as small as possible.
- *Energy-efficiency*: considering the mobile reader is conventionally battery powered, e.g., a typical battery for the mobile reader has a capacity of 3200mAh with output voltage 3.7v, if we scan the tags with a maximum radiation power 36 dBm, the mobile reader can execute continuous scanning for only 3 hours, therefore, the overall energy used should be as small as possible.
- *Coverage ratio*: due to various issues like path loss in realistic settings, it is difficult to identify all tags with a high probability for one single scanning cycle, therefore, the coverage ratio, i.e., the ratio of the number of identified tags to the total number of tags, should be guaranteed, while each tag should have a uniform probability to be identified.

In regard to the continuous scanning, we define the scanning time as T, the overall energy used as E, and the coverage ratio as C. Assuming the tag density is ρ and the length of the scanning area is l, then the total number of tags is $n = l \cdot \rho$, we denote the overall tag set as S. We assume that each tag $t_j \in S$ is successfully identified with probability of p_j after the continuous scanning. The reader's antenna is deployed towards the tags with a distance of d. We can adjust the parameters including the reader's power p_w and the moving speed v to improve the reading performance. Therefore, during the continuous scanning, the problem is how to efficiently set the parameters p_w and v such that the following objectives can be achieved:

Time-efficiency:

$$\text{minimize } T \tag{1}$$

subject to

$$E \leq \alpha \qquad \text{energy constraint} \tag{2}$$

$$Pr[C \geq \theta] \geq \beta \qquad \text{coverage constraint} \tag{3}$$

$$\forall t_j \in S \quad p_j = p \qquad \text{coverage constraint} \tag{4}$$

Energy-efficiency:

$$\text{minimize } E \tag{5}$$

subject to

$$T \leq \gamma \qquad \text{time constraint} \tag{6}$$

$$Pr[C \geq \theta] \geq \beta \qquad \text{coverage constraint} \tag{7}$$

$$\forall t_j \in S \quad p_j = p \qquad \text{coverage constraint} \tag{8}$$

According to the above formulation, in regard to the time-efficiency, the objective is to minimize the overall scanning time T while the energy constraint and the coverage constraint should be satisfied. The energy constraint requires the energy used should be no greater than a certain threshold α. In regard to the coverage constraint, due to the random factors in the anti-collision scheme and the communication environment, the coverage ratio C cannot guarantee to be deterministically equal or greater than a threshold θ, hence we use the probabilistic approach to denote the requirement. The probability for the coverage ratio C to be equal or greater than θ should be no less than β. Moreover, there could exist multiple feasible solutions to guarantee the coverage constraint, in some of the solutions the tags are detected with nonuniform probabilities. In fairness, we require that each tag t_j in the set S should be detected with a uniform probability p, i.e., the detection probability p_j should be equal to p. Similarly, in regard to the energy-efficiency, the objective is to minimize the overall energy E, while the time constraint and the coverage constraint should be satisfied. The time constraint requires that the scanning time should be no greater than a certain threshold, γ.

4. DERIVING A MODEL FROM REALISTIC EXPERIMENTS

In order to understand how the reader's power and tag density affect the reading performance, while dealing with issues like the path loss, energy absorption, and mutual interference, we illustrate several original findings from our realistic experiments. In our experiments, we use the Alien-9900 reader and Alien-9611 linear antenna with a directional gain of 6dB. The 3dB beamwidth is 40 degrees. The RFID tags used are Alien 9640 general-purpose tags which support the EPC C1G2 standards. We attach the RFID tags onto the books which are placed in a large bookshelf. Each tag is attached onto a distinct book with a unique ID. The bookshelf is composed of 12 grids with 4 columns and 3 rows, the height and width of each grid are respectively 60cm and 75cm. The RFID reader is statically deployed by facing its antenna towards the book shelf. Note that in order to set an appropriate value for the distance between the reader and the bookshelf, it is difficult to directly derive the optimal distance from geometry according to the beamwidth, due to issues like the multipath effect. Therefore, we vary the distance from 0.5m to 3m and measure the number of effectively identified tags while scanning 160 tags uniformly distributed on the shelf. We find that the reader achieves the maximum coverage when the distance is 1.5m. Thus, we set the distance to 1.5m to guarantee the reading performance. This setting is close to a typical noisy condition, which is distinct from the free space condition, since the issues in the realistic applications like the path loss, multi-path effect and energy absorption all exist. Considering that we deploy a relatively large number of tags (up to 240 tags) and a rather high tag density (up to 90 tags per square meter) in realistic settings, the experimental findings from the high tag density deployment can be highly scalable and generalized to rather large scale settings.

On the whole, it took us over 300 hours to conduct an extensive experimental study of up to 240 tags in realistic settings. In order to sufficiently understand how the parameters separately/jointly affect the actual reading performance, we conduct up to 100 various experiments, carrying out lots of experimental comparisons and analysis on the obtained results. In order to keep the statistical characteristics, all results are the averaged results of 500 independent trials. We finally summarize these original findings in 12 figures. In the following experiments, we vary the tag density, ρ, from 10 to 40 tags/grid, while adjusting the reader's power from 20.7dBm to 30.7dBm for performance evaluation. Unless otherwise specified, by default we fix the reader towards the center of the bookshelf, set the reader's power to 30.7 dBm, and repetitively scan the tags for 50 query cycles.

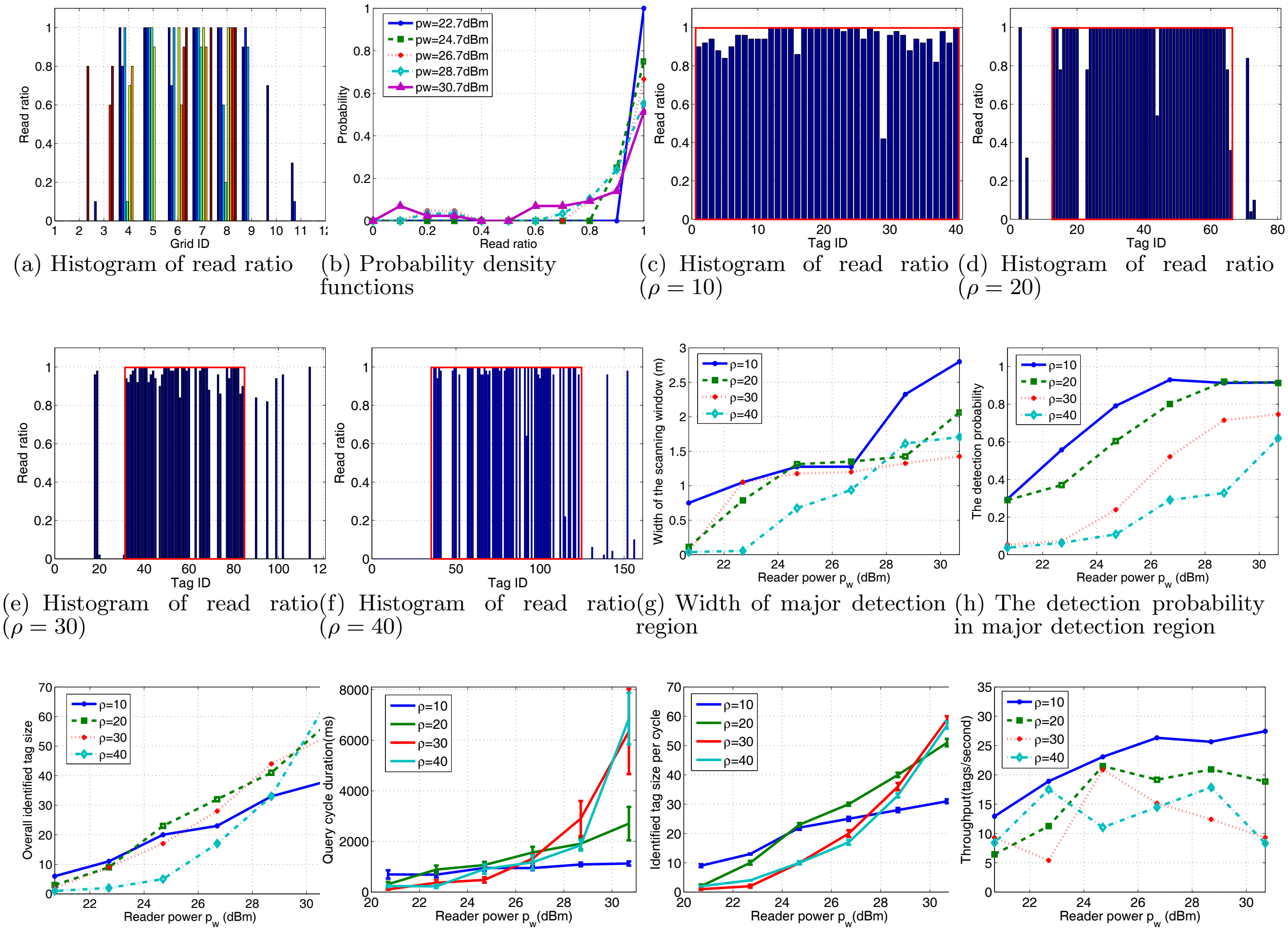
4.1 Experimental findings

4.1.1 Probabilistic backscattering

During the query cycles, each tag responds to the reader with a certain probability between 0 and 1. We uniformly deploy 96 tags in the bookshelf with 8 tags in each grid. The grids on the left/middle/right side are respectively numbered (1,2,3)/(4,5,6,7,8,9)/(10,11,12). In Fig.1(a), we respectively compute the read ratios of each tag in the 12 grids, i.e., the ratio of successful number of responses to the expected number of responses for each tag, and illustrate them in histogram grouped by grid ID. We note that the tags respond to the reader with various probabilities between 0 and 1, although basically no parameters are changed during the repetitive scanning. This observation is contrary to the popular idea that each tag either responds thoroughly or does not respond at all. We think this is probably due to the randomness in the backscattering factors, like the power scattering, multi-path propagation. Furthermore, we vary the reader's power, p_w, from 22.7 dBm to 30.7dBm and obtain the probability density functions for the read ratio. According to Fig.1(b), we note that as the reader's power varies, the distribution of the read ratio also varies. The above observation further implies that, due to the probabilistic backscattering, multiple query cycles are essential to successfully identify a typical tag in the tag set, which may cause massive duplicated readings over other tags in the scanning area.

4.1.2 Major vs minor detection region

Within each reader's detection range, there are two distinct regions: the major detection region where the tags can be identified with high probability, and the minor detection region where the tags can be identified with low probability. We uniformly deploy the tags in a row with 4 grids in the bookshelf, where the tag IDs are sequentially numbered from left to right. The reader's power is set to 30.7dBm. Fig.1(c)-Fig.1(f) show the histogram of each tag's read ratio in the order of tag ID, while varying the tag density, i.e., the number of tags per grid. In order to see the two distinct regions, we use red window to depict the boundary of the major detection region. We observe that within each reader's detection range, the major detection region is the area directly in front of the reader, giving high detection probability, and the minor detection region extends from the end of the major detection region to the edge of the detection range, where the read ratio drops off to zero at the end of the detection range. As the tag density increases, the major detection region gradually shrinks. Note that due to issues like the multipath effect and energy absorption in realistic settings, it is difficult to directly derive the detail parameters of the major/minor detection region from geometrical



(a) Histogram of read ratio (b) Probability density functions (c) Histogram of read ratio ($\rho = 10$) (d) Histogram of read ratio ($\rho = 20$)

(e) Histogram of read ratio ($\rho = 30$) (f) Histogram of read ratio ($\rho = 40$) (g) Width of major detection region (h) The detection probability in major detection region

(i) Overall number of tags identified after 50 query cycles (j) Query cycle duration (k) The number of identified tags per cycle (l) Throughput

Figure 1: Observations from the realistic experiments

principles, since they are also related to other parameters like the tag density.

4.1.3 Marginal decreasing effect

As the reader's power is increasing, the exact read efficiency including the scanning range, the detection probability, as well as the number of identified tags, is not increasing equally with the power. In Fig.1(g)-1(i), we respectively measure the width of the major detection region, the average detection probability (i.e., read ratio) in the major detection region, as well as the overall number of identified tags, while varying the reader's power from 20.7dBm to 30.7dBm. All of the above three variables are increasing while the reader's power increases. However, as the power is increased by 2dBm (i.e, 1.58 times in watt), they mainly increase with a much smaller speed on average. This observation implies that the read efficiency cannot be sufficiently enhanced by purely increasing the reader's power.

4.1.4 Query cycle duration vs the number of identified tags per cycle

As the reader's power increases, the query cycle duration does not increase linearly with the number of identified tags per cycle, causing the variation of the throughput. According to the theoretical analysis in the ideal situation, if the frame size is optimally selected, the expected number of slots as well as the query cycle duration should be linearly increasing with the number of identified tags per cycle. However, in realistic settings that doesn't follow at all. Fig.1(j) and Fig.1(k) respectively show the value of cycle duration τ_c and the number of identified tags per cycle n_c while varying the reader's power. Note that the standard deviation of τ_c is much larger than n_c, which is mainly due to the randomness in the anti-collision scheme. As the reader's power increases, the values of τ_c and n_c are both increasing, however, at different rates. Therefore, the ratio of n_c to τ_c, i.e., the throughput, is also varying.

Fig.1(l) shows the throughput variation with 4 different tag densities. We find that in all cases the throughput achieves the peak value when the reader's power is set to an appropriate value between the minimum and maximum power. The reason is as follows: When the reader's power is set to a small value, the number of activated tags is small, then due to the fairly large inter-query cycle overhead, the throughput is fairly small. As the reader's power increases, more tags are involved in the query cycle, the inter-query cycle overhead is sufficiently amortized, thus the throughput is gradually increased. When the reader's power increases to a fairly large value, the number of collisions in the query cycle is greatly increased, resulting in a large value for the cycle duration, thus the throughput is further decreased. This observation implies that it is neither time-efficient nor energy-efficient to blindly increase the reader's power, an optimal value for the reader's power should be determined.

The above experiment results and observations are obtained from the static situation where the reader is statically deployed. In the mobile situation where the reader is continuously moving, since the moving speed cannot be too large due to the large number of tags to be identified, all the above properties should be preserved. In order to verify this statement, we conduct experiments in mobile situations while varying the moving speed from 0.3m/s to 3m/s. We find that all the obtained results, including the width of major detection region, the detection probability, and the query cycle duration are very close to the static situation. Besides, these experiment results are currently obtained from the settings constructed by the Alien reader and antennas, since the Alien reader and antennas are designed and manufactured according to industrial standard, these results can be applicable to other kinds of commercial readers conforming to the standards. Therefore, it is feasible to apply these parameters to the continuous scanning algorithm.

4.2 Model

Based on the above findings, it is essential to build a model to effectively depict the regularities in reading performance. We first propose a model for probabilistic backscattering, and then a model of the effective scanning window to evaluate the reading performance over multiple tags.

4.2.1 *Probabilistic backscattering*

Suppose an arbitrary tag is separated from the reader at a distance of d. In order for the tag to successfully backscatter the ID message, the reader needs to send a continuous wave to activate the tag. As the tag has a sensitivity threshold t, which is the minimum power required to activate the tag, the incident power to the tag's antenna should be larger than t. It is known that the power budget of conventional RFID systems is *forward-link limited*, which implies that well-designed passive RFID systems are always limited by the tag's sensitivity. Therefore, as long as the reader's power p_w is large enough to activate the tag, the reader is able to resolve the backscattered signal from the tag. We have conducted experiments to evaluate the threshold t. We find that the value of t basically remains unchanged among a certain type of tags.

In the reader's read zone, i.e, the region in which the incident power exceeds the threshold t, it is found that the range is longest along the center and falls off towards the edges. In regard to a plane at a fixed distance from the reader, the incident power varies from the center towards the edges. Besides, the values of incident power has variances since the continuous wave issued from the reader has fluctuations in terms of power. Therefore, assume the reader's power is p_w, in regard to a two dimensional plane at a distance of d from the reader, we respectively use $f_{p_w,d}(x,y)$ and $g_{p_w,d}(x,y)$ to denote the average value and the variance of the incident power in the coordinate (x,y). In the settings where the tags are deployed in a row, we respectively simplify them to $f_{p_w,d}(x)$ and $g_{p_w,d}(x)$. Fig.2 shows a schematic diagram of the average value and variance of the incident power $p'_w(x)$ in the one-dimensional space. Note that conventionally the incident power achieves the maximum value in the center of the read zone, and gradually decreases towards the edges. Meanwhile, in regard to multiple tags deployed in the plane, the incident power is also affected by tag density and multipath effect.

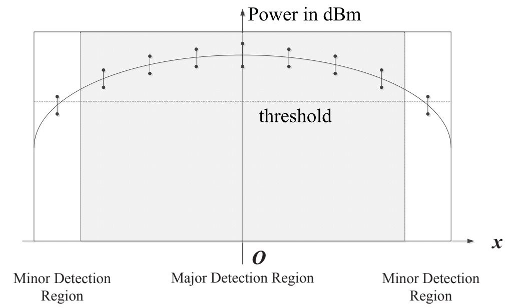


Figure 2: The average value and variance of the incident power in the one-dimensional space: a schematic diagram

In regard to an arbitrary tag in the row, the tag can be successfully identified if and only if the incident power is above the tag's sensitivity threshold t. Due to the fluctuation of the incident power, the tag is successfully identified with some probability, i.e., $Pr[p'_w(x) \geq t]$. In regard to a position x in the effective scanning region, note that once the average value is relatively larger than the threshold t, as the variance is usually relatively small, the detection probability $Pr[p'_w(x) \geq t]$ will be close to 1; similarly, once the average value is relatively smaller than the threshold t, the detection probability $Pr[p'_w(x) \geq t]$ will be close to 0. This property divides the scanning region into two distinct regions, i.e., the major detection region and the minor detection region.

4.2.2 *Effective scanning window over multiple tags*

As we have observed, the reader's effective scanning region can be divided into a major detection region as well as a minor detection region. In the major detection region, most tags can be detected with a probability close to 100%. As the tag density increases, the diffused power cannot guarantee to activate all tags in the major detection region, each tag has a probability to be detected in a random approach. Therefore, we can use the average detection probability to depict the reading performance in this region. The minor detection region is extending from the end of the major detection region to the edge of the effective range, with the detection probability quickly drops off to 0. Based on the above analysis, we use a trapezoidal curve to denote the expected detection probability of tags in the scanning region, as illustrated in Fig.3. In fact, due to the narrow width of the minor detection region, the average probability for a tag to be detected in this region can be negligible. Therefore, in consideration of the actual reading performance, we only need to focus on the major detection region. In the rest of this paper, we use the term *effective scanning window* to denote the major detection region. We use w and p' to denote the width and the average detection probability of the effective scanning window, respectively.

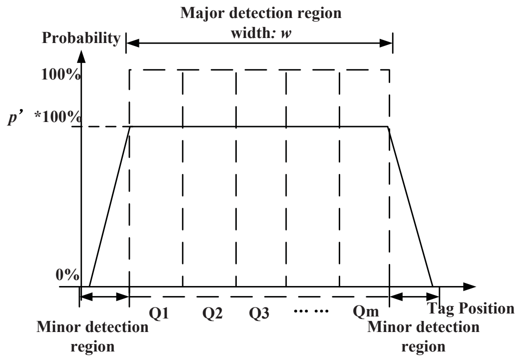


Figure 3: The model of effective scanning window

During continuous scanning, the effective scanning window is continuously moving forward with the mobile reader. Note that there exist overlapping areas between the contiguous scanning windows. During the continuous scanning, each tag gradually enters a minor detection region, then an effective scanning window, finally exits from a minor detection region. While within the effective scanning window, each tag has a probability to be detected for each query cycle. Therefore, in order to guarantee the coverage constraint, multiple query cycles should be issued over each tag while it is within the effective scanning window. Assume that the tags are uniformly deployed along the scanning area, then the number of tags within the effective scanning window is always constant. This infers that the number of tags involved in a query cycle mostly remains unchanged. If the mobile reader is set to a constant power and a constant moving speed, then, after multiple query cycles, each tag has a uniform probability to be detected. This conforms to the requirement in the coverage constraint.

Suppose an arbitrary tag is expected to be queried for m cycles while it is within the effective scanning window, we denote the detection probability in the m query cycles as $p_i(i = 1...m)$. Then, the probability for an arbitrary tag to be identified at least once is as follows:

$$p = 1 - \prod_{i=1}^{m}(1 - p_i) \tag{9}$$

As the reader's power is set to a constant value, due to the uniform tag density, the probability $p_i(i = 1...m)$ in the m query cycles should be uniform. If we use p' to denote the uniform detection probability, then Eq.(9) is further simplified as follows:

$$p = 1 - (1 - p')^m \tag{10}$$

In particular, the value of m is equal to $\frac{\tau_w}{\tau_c}$, here τ_w is the duration in the effective scanning window, and τ_c is the average duration of a query cycle. Moreover, τ_w is equal to the ratio of the window width w to the moving speed v, i.e., $\frac{w}{v}$, hence $m = \frac{w}{v \cdot \tau_c}$. Therefore, in order to increase the detection probability p for an arbitrary tag, it is essential to (1) increase the number of query cycles m as much as possible; (2) increase the detection probability p' as much as possible.

In Fig.1(g), Fig.1(h) and Fig.1(j), we illustrate the value of w, p' and τ_c with various power, p_w. In regard to a fixed tag density, we note that as the value of p_w increases, the value of p', w and τ_c are all monotonically increasing. Moreover, since the value of w increases much more slowly than τ_c, the value of $m = \frac{w}{v \cdot \tau_c}$ is monotonically decreasing with the value of p_w. Therefore, we reach the following conclusion: as the moving speed v decreases, the value of m is monotonically increasing, while the value of p' remains unchanged. As the reader's power p_w increases, the value of p' is monotonically increasing, while the value of m is monotonically decreasing. Thus the value of p_w should be appropriately selected to optimize the performance.

In regard to the coverage constraints in Eq.(3) and Eq.(4), we use the parameter p to denote the probability that a tag is successfully identified after the continuous scanning. Then, according to the binomial distribution, after the overall scanning procedure, the probability for the reader to identify at least $\theta \cdot 100\%$ percent of the overall tags (i.e., $Pr[C \geq \theta]$), is computed as follows:

$$Pr[C \geq \theta] = \sum_{i=\lceil \theta \cdot n \rceil}^{n} C_n^i \cdot p^i \cdot (1 - p)^{n-i} \tag{11}$$

Then it is essential to compute the solution of p to guarantee $Pr[C \geq \theta] \geq \beta$. As $\sum_{i=\lceil \theta \cdot n \rceil}^{n} C_n^i \cdot p^i \cdot (1 - p)^{n-i} = \beta$ is an equation of higher degree, it is rather difficult to directly solve the variable p from the above equation.

In fact, the constraint $Pr[C \geq \theta] \geq \beta$ is equivalent to $Pr[C \leq \theta] \leq 1 - \beta$. In particular, according to *Hoeffding's inequality*, for $\theta \leq p$, it yields the bound

$$Pr[C \leq \theta] \leq \exp(-2\frac{(n \cdot p - n \cdot \theta)^2}{n}).$$

In order for $Pr[C \leq \theta] \leq 1 - \beta$, it is essential to guarantee

$$\exp(-2\frac{(n \cdot p - n \cdot \theta)^2}{n}) \leq 1 - \beta.$$

The solution of p can be directly solved from the above inequality, that is $p \geq \theta^*$, here $\theta^* = \theta + \sqrt{\frac{\ln(1-\beta)}{-2n}}$. This shows that, as long as the detection probability p is no less than θ^* for any tag, the coverage constraint is guaranteed.

5. CONTINUOUS SCANNING WITH MOBILE READER

5.1 Baseline solution

For both the uniform and nonuniform tag distribution, in order to effectively identify all the tags with the mobile reader, conventionally the reader's power is set to maximum and the moving speed is set to a constant value which is small enough. This baseline solution is very straightforward, which however, is neither time-efficient nor energy-efficient since excessive power is used up and the moving speed is slowed down. Besides, a number of tags are interrogated multiple times during continuous scanning, which is unnecessary as each tag only needs to be identified once.

5.2 Solution for uniform tag density

5.2.1 Solution

Without loss of generality, we first propose an optimized solution for the situation with uniform tag density. Considering the objective as well as the energy/time constraint, we need to figure out the optimized value of p_w and v such that the objective is achieved while the coverage constraints are satisfied.

In regard to the coverage constraint, since we need to guarantee $p = 1-(1-p')^m \geq \theta^*$, i.e., $1-(1-p')^{\frac{w}{v\cdot\tau_c}} \geq \theta^*$, it is equivalent to ensure $v \leq \frac{1}{|\ln(1-\theta^*)|}\cdot\frac{w\cdot|\ln(1-p')|}{\tau_c}$. As the value of w, p' and τ_c all depends on the value of p_w, let $w(p_w)$, $p'(p_w)$ and $\tau_c(p_w)$ respectively denote the mapping function from p_w to w, p' and τ_c, then

$$v^* = \frac{1}{|\ln(1-\theta^*)|}\cdot\frac{w(p_w)\cdot|\ln(1-p'(p_w))|}{\tau_c(p_w)}, \tag{12}$$

then, v^* is the maximum allowable moving speed to satisfy the coverage constraint.

Since the length of the scanning area is l, the overall scanning time $T = \frac{l}{v}$, and the overall used energy $E = T\cdot p_w = \frac{p_w\cdot l}{v}$. Therefore, considering the time-efficiency, in order to minimize T, it is equivalent to maximize v. Then, according to Eq.(12), it is essential to maximize $\frac{w\cdot|\ln(1-p')|}{\tau_c}$. It is known that as the value of p_w increases, the value of $w\cdot|\ln(1-p')|$ and τ_c are both monotonically increasing, thus an optimized value of p_w should be selected to minimize T. Considering the energy constraint $E \leq \alpha$, the optimal value p_w^* can be computed according to the following formulation:

$$\text{maximize } y_T = \frac{|\ln(1-p'(p_w))|\cdot w(p_w)}{\tau_c(p_w)} \tag{13}$$

subject to

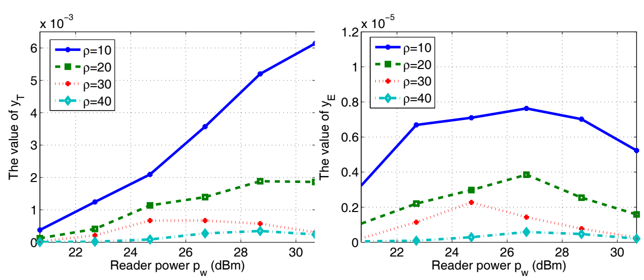
$$\frac{|\ln(1-p'(p_w))|\cdot w(p_w)}{p_w\cdot\tau_c(p_w)} \geq \frac{l\cdot|\ln(1-\theta^*)|}{\alpha} \tag{14}$$

Considering the energy-efficiency, in order to minimize E, it is equivalent to minimize $\frac{p_w}{v}$, then according to Eq.(12), it is essential to maximize $\frac{|\ln(1-p')|\cdot w}{p_w\cdot\tau_c}$. Therefore, considering the time constraint $T \leq \gamma$, the optimal value p_w^* can be computed according to the following formulation:

$$\text{maximize } y_E = \frac{|\ln(1-p'(p_w))|\cdot w(p_w)}{p_w\cdot\tau_c(p_w)} \tag{15}$$

subject to

$$\frac{|\ln(1-p'(p_w))|\cdot w(p_w)}{\tau_c(p_w)} \geq \frac{l\cdot|\ln(1-\theta^*)|}{\gamma} \tag{16}$$



(a) The value of y_T (b) The value of y_E

Figure 4: Compute the value of y_T and y_E with various values of p_w

In regard to a certain tag density ρ, by enumerating the candidate values of the power p_w, we can compute the value of y_T and y_E. Fig.4(a) and Fig.4(b) respectively illustrate the value of y_T and y_E while varying the reader's power p_w. We note that there exist a maximum value of y_T and y_E for each tag density. In regard to a specified tag density ρ, while satisfying the time/energy constraint, we can use the power p_w^* for the maximum value of y_T or y_E as the optimal parameter and compute the corresponding moving speed v^* according to Eq.(12). In this way, the optimal solution (p_w^*, v^*) for time/energy efficiency can be generated. Therefore, in regard to various tag densities ρ, we can collect the performance parameters like w, p' and τ_c in advance, pre-compute the optimal pairs of (p_w^*, v^*), and store them in a table. When dealing with an arbitrary tag density, we can directly use the optimal pair of (p_w^*, v^*) to achieve the time/energy efficiency.

5.2.2 *Estimate the tag density*

According to the measured data in realistic settings, it is known that the tag density ρ has an important effect on the performance metrics. In situations where the tag density cannot be pre-fetched or the tag density varies along the forwarding direction, it is essential to accurately estimate the current tag density, such that the optimized parameters (p_w^*, v^*) can be effectively computed. Due to the probabilistic backscattering property, it is difficult to directly estimate the tag density according to the observed number of empty/singleton/collision slots [8, 11]. Furthermore, current commercial RFID readers do not expose these low-level data to upper-layer applications. Therefore, it is essential to estimate the tag density in a more practical way.

According to Fig.1(k), we note that if the reader's power p_w is set to a certain value, the number of identified tags per cycle n_c is varying as the tag density ρ varies, with a very small standard deviation. Table 1 shows further details for the average values of n_c. These are obtained through 50 repetitive experiments with various values of ρ and p_w. Due to the small variance of n_c, there is a very stable pattern between n_c and ρ that varies with p_w. Therefore, given a reference tag density ρ_i, we can depict the values of n_c with various powers as a vector $V_i = \{n_{i,1}, n_{i,2}, ..., n_{i,s}\}$, here s is the number of power levels. Then, in regard to an unknown tag density ρ, assume the corresponding vector is $V = \{n_1, n_2, ..., n_s\}$, we can estimate the value of ρ by comparing V with the vectors of reference tag densities. Therefore, we propose an algorithm to estimate the tag density, by leveraging the k-nearest neighbor method, as shown in Algorithm 1.

$p_w=$	20.7	22.7	24.7	26.7	28.7	30.7
$\rho=10$	9	13	22	25	28	31
$\rho=20$	2	10	23	30	40	51
$\rho=30$	1	2	10	20	36	59
$\rho=40$	2	4	10	17	33	57

Table 1: The number of identified tags per cycle

In Algorithm 1, the similarity $sim(V, V_i)$ is actually calculated by using the cosine value of the angle between the two vectors, hence the value of similarity is between 0 and 1. We use the k-nearest neighbor method to estimate the tag density based on k-nearest reference tag densities. The estimated tag density ρ is computed using an inverse distance weighted average with the k-nearest multivariate neighbors, here the distance is defined as $1-sim(V, V_i)$. Since the value of n_c has a rather small variance, the accuracy of the estimated tag density can be guaranteed if the number of samplings m is fairly large. In the algorithm, the mobile

Algorithm 1 Tag density estimation algorithm

INPUT: $V_i = \{n_{i,1}, n_{i,2}, ..., n_{i,s}\}$: the vectors for reference tag densities $\rho_i (i = 1...h)$.

PROCEDURE

Set the reader's power to various levels $p_{w,1}, ..., p_{w,s}$. In regard to each power level $p_{w,j} (j \in [1, s])$, issue m query cycles to get the average value of the number of identified tags per cycle as n_j. Assemble them as a vector $V = \{n_1, n_2, ..., n_s\}$.

for $i \in [1, h]$ **do**

Compute the similarity between V and V_i as follows:

$$sim(V, V_i) = \frac{V \cdot V_i}{|V| \cdot |V_i|} = \frac{\sum_{j=1}^{s} n_{i,j} \cdot n_j}{\sqrt{\sum_{j=1}^{s} n_j^2} \cdot \sqrt{\sum_{j=1}^{s} n_{i,j}^2}}.$$

end for

Sort the value of $sim(V, V_i)$ in decreasing order. Find the first k items of ρ_i according to $sim(V, V_i)$, say $\rho_1', ..., \rho_k'$.

Compute $\rho = \sum_{i=1}^{k} \rho_i' \cdot w_i$, here

$$w_i = \frac{1/(1 - sim(V, V_i) + \epsilon)}{\sum_{i=1}^{k} 1/(1 - sim(V, V_i) + \epsilon)} (\epsilon > 0).$$

OUTPUT: The estimated tag density ρ.

reader is required to obtain the value of n_c in multiple power levels, which increases overhead in both time and power usage. In regard to the uniform tag density, since the tag density is only necessary to estimate once, this overhead can be effectively amortized by the following multiple query cycles. In regard to the nonuniform tag density, the tag density needs to be continuously estimated. The algorithms for fast tag size estimation [8, 11, 15] can be used to reduce the overhead. In regard to the selection of k, k should be set to neither a too small value nor a too large value for accurate estimation, the optimal value depends on the exact deployment, conventionally k should be set to 2 or 3 for performance consideration.

This algorithm is very practical and fully compatible with the EPC C1G2 standard, and does not require to obtain any low-level parameters for commercial RFID readers.

5.3 Extensions for nonuniform tag density

In the above solution, we assume that the tag density is always uniform. In some applications, the tags are not uniformly deployed. While the mobile reader is continuously scanning the tags, the tag density may always change along the forward direction. In this situation, the constant moving speed and power for the mobile reader is no longer suitable to improve performance. Note that in conventional situations, the tag density changes slowly along the forward direction. Therefore, in regard to each query cycle, we can assume the tag density within the effective scanning window is close to uniform, since the cycle duration is usually small. Therefore, we can reduce the situation with nonuniform tag density into multiple snapshots with fairly uniform tag density. In each query cycle, the mobile reader can be reset with the optimal values of p_w and v according to the nearby tag density. In this way, the reading performance can be effectively improved by dynamically adjusting the reader's power p_w and the moving speed v.

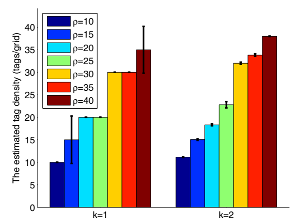
6. PERFORMANCE EVALUATION

We evaluate the performance in realistic settings. The experiment settings are the same as the realistic settings in Section IV, except that in this experiment we use the Alien-9900 reader as the mobile reader to move forward for continuous scanning.

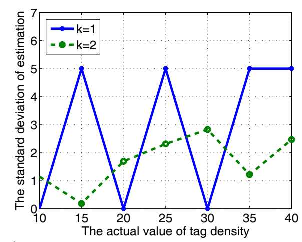
6.1 Evaluate the performance in unform tag density

In order to evaluate the performance in unform tag density, we deploy the tags in a row with 4 grids in the shelf, while varying the tag densities from 10 tags/grid to 40 tags/grid, the length of the scanning area is 3m.

6.1.1 The accuracy of tag density estimation



(a) The average value and standard deviation for estimated tag density



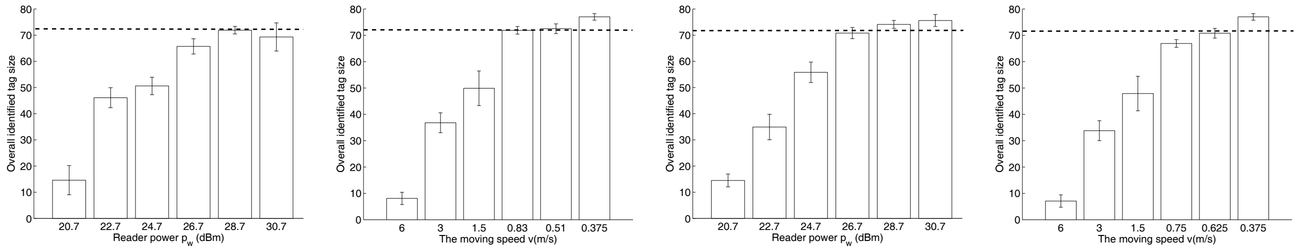
(b) The estimation error of tag density estimation

Figure 5: Evaluate the accuracy of tag density estimation

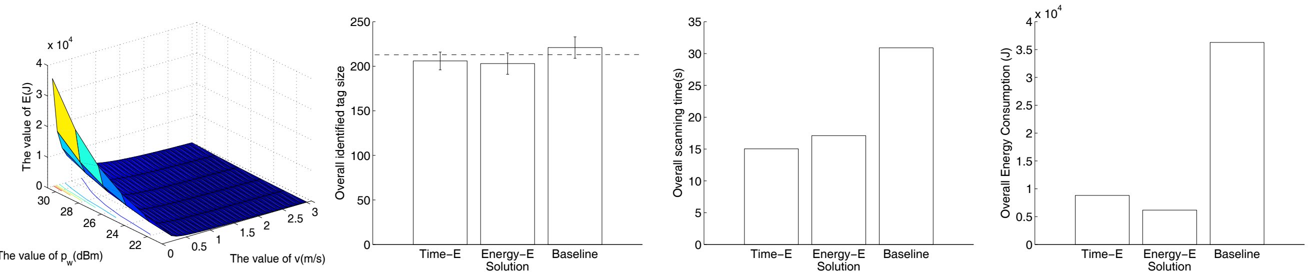
In order to perform tag density estimation, we utilize Table 1 as the reference set. Then, while varying the tag density from 10 to 40 tags/grid in increments of 5 tags/grid, we estimate the tag densities based on the number of identified tags per cycle, i.e., n_c. For each power level, we collect 10 samples of n_c for estimation. As the k-nearest neighbor method is used in the estimation, we respectively set k to 1 and 2 for the estimation. Fig.5(a) illustrates the estimated values as well as the standard deviations for various tag densities. We find that both the 1-nearest neighbor method (1NN) and the 2-nearest neighbor method (2NN) achieve fairly good performance in terms of estimation accuracy. As 1NN can only select one tag density with the nearest property, the estimate accuracy declines when the exact tag density is between two reference tag densities in Table 1. In comparison, 2NN achieves a much higher estimation accuracy since it can effectively select two closest reference tag densities to estimate the tag densities. Fig.5(b) further compares the estimation error of the two methods. We use the standard deviation as the metric for the estimation error. We note that the standard deviation for 1NN is fluctuating between 0 and 5 tags/grid while the standard deviation for 2NN is relatively stable below 3 tags/grid. This infers that in conventional cases 2NN achieves a much better performance than 1NN in terms of estimation accuracy.

6.1.2 The coverage ratio

According to the analysis in Section IV, in regard to the coverage constraint $Pr[C \geq \theta] \geq \beta$, we need to guarantee $p \geq \theta^*$, here $\theta^* = \theta + \sqrt{\frac{\ln(1-\beta)}{-2n}}$. Without loss of generality, we set θ^* to 90% in our experiments. This means, on average 90% of the tags should be identified after continuous scanning. According to Fig.4, the optimal values of the



(a) The coverage of various powers (time-efficiency)
(b) The coverage of various moving speed (time-efficiency)
(c) The coverage of various powers (energy-efficiency)
(d) Coverage of various moving speed (energy-efficiency)



(e) The energy consumption with various power and moving speed
(f) Coverage ratio for time-efficient, energy-efficient and baseline solution
(g) Scanning time for time-efficient, energy-efficient and baseline solution
(h) Energy consumption for time-efficient, energy-efficient and baseline

Figure 6: Experiment results in realistic settings

power p_w^* and moving speed v^* for the mobile reader are computed in Table 2. The corresponding scanning time and energy are also illustrated.

TE	$\rho = 10$	$\rho = 20$	$\rho = 30$	$\rho = 40$
p_w^*	30.7dBm	28.7dBm	26.7dBm	28.7dBm
v^*	2.65m/s	0.83m/s	0.29m/s	0.15m/s
T^*	1.13s	3.6s	10.3s	20s
E	1326.62J	2667.6J	4820.4J	14820J

EE	$\rho = 10$	$\rho = 20$	$\rho = 30$	$\rho = 40$
p_w^*	26.7dBm	26.7dBm	24.7dBm	26.7dBm
v^*	1.57m/s	0.625m/s	0.29m/s	0.12m/s
T	1.91s	4.8s	10.3s	25s
E^*	893.88J	2243.75J	3038.5J	11700J

Table 2: Optimal parameters for time-efficiency (TE) and energy-efficiency (EE)

We vary the tag densities ρ from 10 tags/grid to 40 tags/grid and verify the coverage ratio with various configurations for the parameters p_w and v. Due to the limitation of space, we only illustrate the results for $\rho = 20$ tags/grid, the other results are very similar to them. We run each experiment 20 times to obtain the average value and the standard deviation of the number of identified tags. In regard to the time-efficiency, Fig.6(a) shows the coverage of various power levels, while fixing the moving speed to the optimal value $v^* = 0.83$ m/s, the dashed line denotes the threshold for 90% coverage. We note that as the power increases, the number of identified tags gradually increases to the threshold and then further decreases. The threshold is only achieved while the power is set to the optimal value 28.7dBm. Fig.6(b) shows the coverage of various moving speed, while fixing the power to optimal value $p_w^* = 28.7$ dBm. We note that as the moving speed decreases, the number of identified tags gradually increases to cross the threshold for 90% coverage, the coverage is right at the threshold when the speed is set to the optimal value 0.83 m/s. The above results show that, while guaranteeing the coverage ratio, the optimal settings (p_w^*, v^*) can achieve much better time-efficiency than other settings.

Similarly, in regard to the energy-efficiency, Fig.6(c) and Fig.6(d) respectively show 1) the coverage of various power levels while fixing the moving speed to the optimal value $v^* = 0.625$ m/s; 2)the coverage of various moving speed, while fixing the power to optimal value $p_w^* = 26.7$ dBm. It is found that, among various (p_w, v) parameter pairs, the optimal solution (p_w^*, v^*) achieves the coverage right at the required threshold. These results infer that, while guaranteeing the coverage ratio, the optimal settings (p_w^*, v^*) can achieve much better energy-efficiency than other settings.

6.1.3 The time/energy-efficiency

In regard to the time-efficiency, since the length of the scanning area is 3m, the scanning time $T = \frac{3\text{m}}{v}$. Therefore, while guaranteeing the coverage ratio, a high speed v is preferred. Note that as the value of v increases from 0 to 1m/s, the scanning time rapidly decreases from $+\infty$ to 3s; as the value of v further increases, the scanning time decreases rather slowly from 3s to 0, while the coverage ratio can be decreased rapidly. Considering the marginal decreasing effect, the moving speed should be appropriately selected for cost-effective consideration. In regard to the energy efficiency, it is known that the overall energy consumption E is proportional to the power p_w and inverse to the moving speed v. Fig.6(e) illustrates the value of E while varying the value of p_w and v. Moreover, in order to guarantee the coverage ratio, the values of p_w and v are mutually restricted. Recall that Fig.4(a) actually illustrates the maximum allowable moving speed $v = \frac{y_T}{|\ln(1-\theta^*)|}$ for reader's power p_w

with various tag densities. Integrating with both figures, we can effectively derive the minimum energy to satisfy the coverage constraint.

6.2 Evaluate the performance in non-unform tag density

In order to evaluate the performance in non-unform tag density, we nonuniformly deploy 240 tags in a row with 12 grids, while varying the tag density from $\rho = 10$ to $\rho = 30$ tags/grid, the length of the scanning area is 9m. We set θ^* to 90% in our experiments, which infers that, on average 90% of the tags should be identified after the continuous scanning. We compare our time-efficient solution (*Time-E*) and energy-efficient solution (*Energy-E*) with the baseline solution (*Baseline*). In regard to the baseline solution, we set the reader's power to its maximum value, i.e., 30.7dBm, which is also the standard configuration for conventional commodity readers. The moving speed is set according to the optimal value $v^* = 0.29$m/s when $\rho = 30$ to tackle the worst case. In Fig.6(f), we evaluate the coverage ratio for the three solutions. Both *Time-E* and *Energy-E* achieve the coverage which is very close to the 90% coverage ratio. *Baseline*'s coverage is slightly more than this threshold, since *Baseline* uses the maximum power and lowest speed. In Fig.6(g) and Fig.6(h), we respectively evaluate the overall scanning time and energy consumption for continuous scanning. Both *Time-E* and *Energy-E* achieves much better performance than *Baseline* in regard to the two metrics. Here, *Time-E* saves more than 50% of the scanning time compared with *Baseline*, and *Energy-E* saves more than 83% of the energy consumption compared with *Baseline*.

7. CONCLUSION

This paper considers how to efficiently identify RFID tags with a mobile reader, from the experimental point of view. We conduct measurements over a large volume of tags in realistic settings, and propose efficient algorithms for continuous scanning. Our experiments show that our algorithms can reduce the total scanning time by 50% and the total energy consumption by 83% compared to the prior solutions. We believe this work gives much insight and inspiration for devising optimized algorithms for reading a large number of tags in realistic settings.

Acknowledgments

This work is supported in part by National Basic Research Program of China (973) under Grant No. 2009CB320705; National Natural Science Foundation of China under Grant No. 61100196, 61073028, 61021062; JiangSu Natural Science Foundation under Grant No. BK2011559. Qun Li is supported in part by US National Science Foundation grants CNS-1117412 and CAREER Award CNS-0747108.

8. REFERENCES

[1] S. R. Aroor and D. D. Deavours. Evaluation of the state of passive uhf rfid: An experimental approach. *IEEE Systems Journal*, 1(2):168–176, 2007.

[2] D. Benedetti, G. Maselli, and C. Petrioli. Fast identification of mobile rfid tags. In *Proc. of IEEE MASS*, 2012.

[3] K. Bu, B. Xiao, Q. Xiao, and S. Chen. Efficient misplaced-tag pinpointing in large rfid systems. *IEEE Transactions on Parallel and Distributed Systems*, 23(11):2094–2106, 2012.

[4] M. Buettner and D. Wetherall. An empirical study of uhf rfid performance. In *Proc. of MobiCom*, 2008.

[5] S. Chen, M. Zhang, and B. Xiao. Efficient information collection protocols for sensor-augmented rfid networks. In *Proc. of INFOCOM*, 2011.

[6] H. Han, B. Sheng, C. C. Tan, Q. Li, W. Mao, and S. Lu. Counting rfid tags efficiently and anonymously. In *Proc. of INFOCOM*, 2010.

[7] S. R. Jeffery, M. Garofalakis, and M. J. Franklin. Adaptive cleaning for rfid data streams. In *Proc. of VLDB*, 2006.

[8] M. Kodialam and T. Nandagopal. Fast and reliable estimation schemes in rfid systems. In *Proc. of ACM MobiCom*, 2006.

[9] S. Lee, S. Joo, and C. Lee. An enhanced dynamic framed slotted aloha algorithm for rfid tag identification. In *Proc. of MobiQuitous*, 2005.

[10] T. Li, S. Chen, and Y. Ling. Identifying the missing tags in a large rfid system. In *Proc. of ACM Mobihoc*, 2010.

[11] T. Li, S. Wu, S. Chen, and M. Yang. Energy efficient algorithms for the rfid estimation problem. In *Proc. of INFOCOM*, 2010.

[12] J. Myung and W. Lee. Adaptive splitting protocols for rfid tag collision arbitration. In *Proc. of ACM MobiHoc*, 2006.

[13] T. L. Porta, G. Maselli, and C. Petrioli. Anti-collision protocols for single-reader rfid systems: temporal analysis and optimization. *IEEE Transactions on Mobile Computing*, 10(2):267–279, 2011.

[14] C. Qian, Y. Liu, H.-L. Ngan, and L. M. Ni. Asap: Scalable identification and counting for contactless rfid systems. In *Proc. of ICDCS*, 2010.

[15] C. Qian, H.-L. Ngan, and Y. Liu. Cardinality estimation for large-scale rfid systems. In *Proc. of PerCom*, 2008.

[16] B. Sheng, Q. Li, and W. Mao. Efficient continuous scanning in rfid systems. In *Proc. of INFOCOM*, 2010.

[17] B. Sheng, C. C. Tan, Q. Li, and W. Mao. Finding popular categoried for RFID tags. In *Proc. of ACM Mobihoc*, 2008.

[18] S. Tang, J. Yuan, X. Y. Li, G. Chen, Y. Liu, and J. Zhao. Raspberry: A stable reader activation scheduling protocol in multi-reader rfid systems. In *Proc. of ICNP*, 2009.

[19] H. Vogt. Efficient object identification with passive rfid tags. In *Proc. of Pervasive*, 2002.

[20] L. Xie, B. Sheng, C. C. Tan, H. Han, Q. Li, and D. Chen. Efficient tag identification in mobile rfid systems. In *Proc. of INFOCOM*, 2010.

[21] L. Yang, J. Han, Y. Qi, C. Wang, T. Gu, and Y. Liu. Season: Shelving interference and joint identification in large-scale rfid systems. In *Proc. of INFOCOM*, 2011.

[22] Y. Yin, L. Xie, S. Lu, and D. Chen. Efficient protocols for rule checking in rfid systems. In *Proc. of ICCCN*, 2013.

[23] Y. Zheng and M. Li. Fast tag searching protocol for large-scale rfid systems. In *Proc. of ICNP*, 2011.