Frogeye: Perception of the Slightest Tag Motion

Lei Yang*†, Yong Qi†, Jianbing Fang*, Xuan Ding*, Tianci Liu*, Mo Li‡
† CST, Xi'an Jiaotong University, China
* School of Software and TNLIST, Tsinghua University, China
‡ School of Computer Engineering, Nanyang Technological University
young@tagsys.org, qiy@mail.xjtu.edu.cn, {bing, xuan, tian}@tagsys.org, limo@ntu.edu.sg

***Abstract*—Existing methods in RFID systems often employ presence or absence fashion to detect the tags' motions, so they cannot meet motion detection requirement in many applications. Our recent observations suggest that the signal strength backscattered from the tag is hypersensitive to its position, inspiring us to perceive the tag motion through its radio signal strength changes. Motion perception is not trivial and challenged by weak stability of strength in that any other interference or noise may incur significant changes as well, resulting in high false positives. To tackle this issue, we propose to model the strength via the Mixture of Gaussian Model (MoG). The problem is thus converted to foreground segment in computer vision with the help of Strength Image, where the technique of MoG based background subtraction is employed. We then implement a prototype using commercial off-the-shelf products. The evaluation results show that the slightest tag motion ($\sim 10cm$) can be precisely perceived, and the accuracy is up to 92.34% while the false positive is suppressed under 0.5%.**

***Keywords*—RFID, Motion detection, Background substraction, Mixture of Gaussian Model**

I. Introduction

An RFID system typically consists of a large number of RFID readers and tags. Tags are attached to products for labeling, and usually have no battery supply so they have to be activated within the interrogation area of a reader. The reader interrogates the tags and collects their IDs via continuous RF waves, without the need of in sight or touch. In contrast to conventional identification technologies, like barcodes, RFID systems have many advantages, such as non-optical proximity, long transmission range, and quick identification.

The initial motivation of RFID is to automatically identify objects fast and conveniently, but its potential applications have been widely studied in recent years for areas. Two classical scenarios are presented as follows. *Securing valuable objects.* Museums and art galleries use RFID to track valuable artifacts, which is treated as a cost-efficient solution when the price of nowadays RFID tags can be as low as 1 USD per capita. A key requirement is that any movement of the tag (and thus the associated item) can be accurately captured, which might be as slight as centimeters. *Mining customer's behaviors.* To stay competitive, a large number of data mining techniques have been introduced to help supermarket managers better understand consumers' behaviors. Unfortunately, the techniques are generally confined to the consumers' purchase data. The motion of the RFID tags on the products again implies rich behavioral information of the customer, e.g., taking the product from the shelf and returning it back. Again, precise and accurate detection of tag motion is of essential importance in this application. At first glance, there is no any connection between the above two scenarios. Actually, both focus on the surveillance of tag motions: the first needs an alert when valuable objects are moved; the second requires behavior records when the products are taken off the shelf.

978-1-4799-3360-0/14/$31.00 © 2014 IEEE

To the best of our knowledge, few works have been studied on the surveillance of tag motion. Most previous works detect tag motion based on its absence or presence in ranges of different RFID readers, which gives very coarse granularity but cannot precisely detect slight tag motion at centimeter level. The recent RFID tag localization approaches [1]–[3] can work with better precision, but still at meter level which is not high enough. Besides, tag localization often incurs high deployment cost, including the deployment of numerous RFID readers and anchors.

In this study, we propose and design Frogeye, a system that exploits the signal strength changes of the backscattered from moving tags to provide a hypersensitive approach for tag motion perception. In theory, the power received by the reader which is backscattered from the tag, goes as the inverse fourth power of the distance in between. Such a characteristic makes the signal strength hypersensitive to tag positions. Our empirical studies validate the hypotheses, and reveal expected and encouraging results. Strength change based motion perception approach provides three obvious benefits: First, no additional devices are required because the radio signal strength is the most common parameter provided by the commercial off-the-shelf (COTS) readers. Second, there is no need for pre-deployed anchors or infrastructures. Third, open and welcoming surveillance can be offered without line of sight, which allows persons and objects to move between the reader and tags.

Motion perception at high accuracy is not trivial because the signal strength has *weak stability* in that arbitrary other interference or noise may incur significant changes. In detail, the thermal noise enables the signal strength to vibrate in a small range. Multi-path enhances or weakened the strength. To address this issue, we utilize the Mixture of Gaussian (MoG) to model such weak stability. Our approach includes three steps. In the first step, a preprocessing is applied to divide the original reading stream into read frames. In the second step, the frame is projected into a strength image. An improved MoG based background subtraction (MoG-BS) method from computer vision is applied to detect the foreground pixels in the third step. We also design the mechanism to discover the moved tags through the detected foreground pixels.

Compared with the existing methods, Frogeye advances

the perception in terms of sensitivity and stability. Our major contributions are summarized as follows.

• We conduct extensive statistical analysis of strength collected in a real-life office, showing that the strength are indeed hypersensitive to tags' positions, but suffers from weak stability where the strength values are highly clustered in a small range due to thermal noise, and enhanced or weakened due to multi-path effect. We then present a MoG to accurately characterize the weak stability.

• We propose Frogeye, to perceive the slight of tag motion. This approach takes a snapshot of tags' positions through their backscattered strength every several read cycles, producing a sequence of strength frames. The MoG-BS is leveraged to detect any foreground pixels, namely moved tags in our scenario.

• We implement the Frogeye using pure COTS RFID devices, ImpinJ R420 reader, and evaluate it at varying parameter choices. The evaluation results show that the slightest tag motion can be precisely perceived ($\sim$ 10 cm on average), and the accuracy is up to 92.34% while the false positive is suppressed under 0.5%.

The remainder of the paper is structured as follows. We introduce and model the weak stability of signal strength in Section II. The main design of Frogeye is presented in Section III. The implementation and evaluation are given in Section IV. The related works about RFID system are overviewed in Section V. Finally, Section VI concludes this paper.

II. Modeling The Weak Stability

In this section, we introduce the theoretical background of the RFID technique and perform several empirical studies to model the weak stability of strength.

A. Intuition

Passive tags not equipped with batteries do not use a radio transmitter. Instead, they use modulation of the reflected power from the tags. Referring to [4], we can construct a mathematical statement of the power relationships using the Friis equation. Defining the gains of the reader's and tag's antenna G_{reader}, G_{tag}, a backscatter transmission loss T_b, we can describe the power backscattered by the tag as:

$$P_{\text{RX,reader}} = P_{\text{TX,reader}} G_{\text{reader}}^2 G_{\text{tag}}^2 \left(\frac{\lambda}{4\pi r}\right)^4 T_b^2$$

where λ is the wavelength and r is the distance between reader and tag. The above equation indicates that the power received by the reader and backscattered from the tag, goes as the inverse fourth power of the distance. *In other words, the power backscattered by the tag is hypersensitive to the tag's positions.* This characteristic inspires and motivates us to perceive the tag motion by leveraging the changes in its signal strength. But the intuition becomes reality only when the backscattered strength behaves relatively stable. Namely, the strength backscattered from a stationary tag should remain unchanged.

We place a tag in front of the antenna in a quiet lab without any interference. The strength collected from different positions are shown in Figure 1, where the d indicates the distance (cm) between the reader and tag. We observe that (1) the strength is indeed hypersensitive to the distance. The strength difference is very noticeable even if the two positions are very close ($\sim 10cm$). About the strength sensitivity, we further conduct evaluation in Section IV. (2) Unfortunately, the result is not as stable as expected, because the value occupies several units even when the tag remains in a same distance. We call this phenomenon *weak stability*.

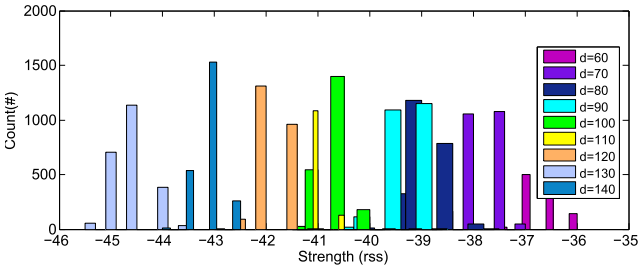


Fig. 1. Strength distribution

In fact, it is still full of challenges to perceive the slightest tag motion despite the strength backscattered by the tag being hypersensitive to the motion, because the electronic component's thermal vibration also brings changes. Especially when the strength is interfered, its changes are as significant as when the tag is moved. It is easy to mistakenly consider a stationary object moved. Aiming to distinguish these changes caused by real motion from the thermal vibration or interference, we next offer empirical insights into the weak stability and statistical methods to model them.

B. Modeling the Thermal Noise

As illustrated in Figure 1, the strength of stationary tag does not always remain constant in the manner we expected. There is a range of vibration whatever effort we make to enable environment interference free. Actually, we think such vibration mainly comes from the thermal noise of the electronic components rather than the environment. Therefore, we model the strength using Gaussian distribution $N(\mu, \sigma^2)$. We believe this model is reasonable because a lot of natured phenomena follows the Gaussian distribution, especially thermal noise from internal electronic components, which mainly contribute the vibration. Then the probability that target tag x has strength value x_t at time t is estimated as:

$$\eta(x_t) = \frac{1}{\sqrt{2\pi|\sigma^2|}} \exp\left(-\frac{(x_t-\mu)^T(x_t-\mu)}{2|\sigma^2|}\right)$$

The Gaussian model depends on the tag position and features of the electronic components. The preview determines the center of vibration and the latter constrains its range. To verify the Gaussian model, we first use the Quantile-quantile plot (QQ plot) to visually display whether the X is from a Gaussian distribution (Gaussian model). If the distribution of X is normal, the plot will be linear compared with the ideal theoretical distribution. We use the interference-free trace from the first measurement. Figure 2 plots collections from two different antennas. In the figure, the x-value of the point is the theoretical quantile and the y-value is the quantile of the data trace. The dashed line is the straight line connected by

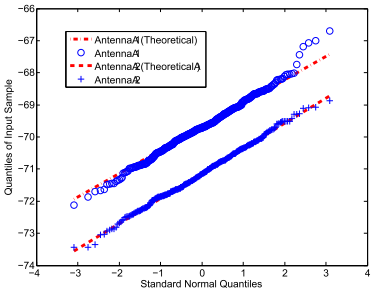


Fig. 2. QQ plot

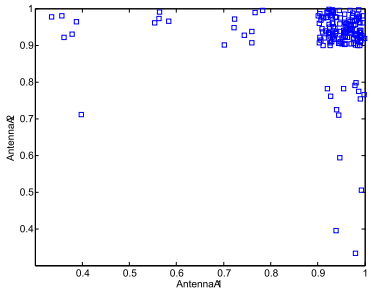


Fig. 3. JB test

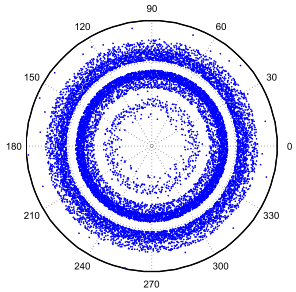


Fig. 4. The strength cloud

the first and third quartile calculated theoretically. If the traces follow the theoretical Gaussian distribution, their quantile is distributed around the dashed line. As we expect, the points in the figure are linear and very close to the theoretical line. Second, we employ Jarque-Bera Test (JB test) of 0.95 significance level to quantitatively evaluate the goodness-of-fit of the Gaussian model. The JB test is a widely adopted tool to test whether the sample data matches a Gaussian distribution. In detail, we can fit the Gaussian model by maximum likelihood estimation given a collected samples, $\{x_1, \cdots, , x_t\}$, as follows.

$$\mu \approx \bar{x} = \frac{1}{t}\sum_{i=1}^{t} x_i \text{ and } \sigma^2 \approx \frac{1}{t}\sum_{i=1}^{t}(x_i - \bar{x})^2$$

Since the trace is recorded according to the time flow, we divide the whole time recorded into equal size windows. The JB test is then applied for each window to test the goodness-of-fit for the estimated Gaussian model. The pass rate is calculated every 100 windows. The test is performed on 100 tag traces. Figure 3 illustrates the pass rate using the JB test. If points are clustered at the top and right-hand corner, then the fitness is good. We observe that the modelability of the strength using Gaussian model is rather convincing.

C. Modeling the Interference

In the second experiment, we bring an inference object when collecting the strength from the target stationary tag. We let one interference object (a person) work slowly through the intermediate region between the reader and target tag. In the experiment, a total of 13,797 strength values are recorded. Figure 4 plots the tag's 'strength cloud' in which the point is scattered using one strength value in a random direction. The point in the polar origin has the strongest strength. It is evident that: (1) being different from the case without the inference, the strength vibration does not concentrate on a sole level any more. Instead, there are three clear levels which correspond to three cloud layers in the figure. Clearly, the strength is enhanced or decayed by the person's movement. (2) the thickness of each layer is equally likely. Essentially, the vibration ranges are very similar. This indicates that the thermal noise is still the leading factor contributing to the vibration whether the tag is interfered with or not.

These phenomenons can be explained by the multipath effect. There exist several paths for the backscattered signal propagating from tag to reader. The signal strength propagating through different paths varies a lot due to the path length. The reader prefers to resolve the strongest signal from all paths. Note that the strongest does not always come from the shortest path because two signal propagation may either cancel out or reinforce each other. When the interference object gets close to the tag, it may block some propagation paths and lead to the propagation jumping among the multiple paths, resulting in the strength migrates from one level to another.

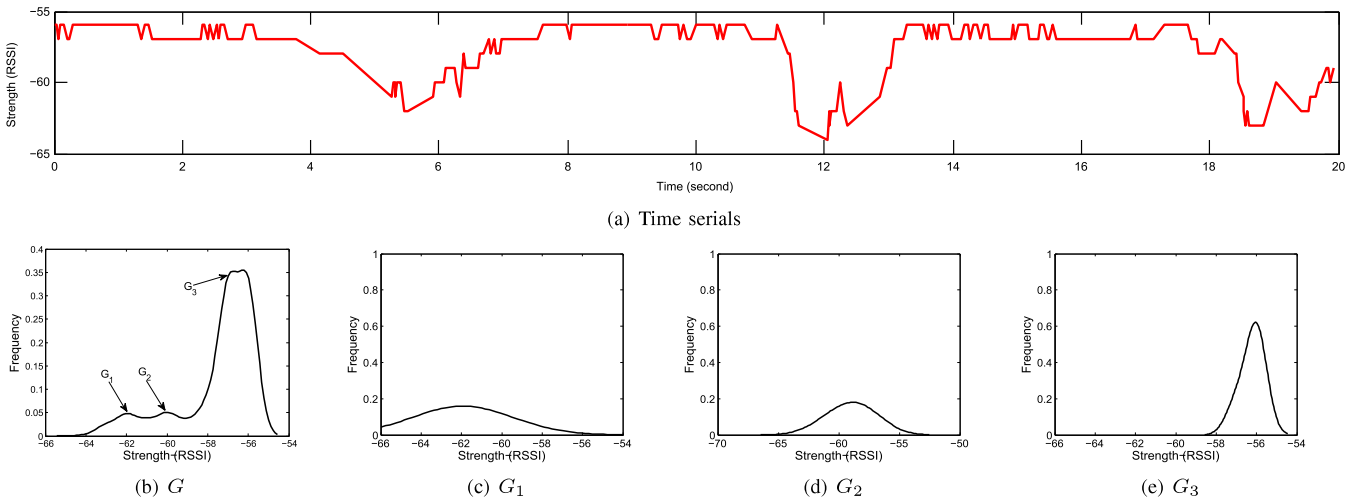
From a long-term perspective, the strength exhibits multi-modal characteristics where the distribution is likely composed of multiple Gaussian models. For example, Figure 5(a) shows the strength changes in the above measurement over a short period of time ($20s$). Some of the time the strength regularly vibrates without interference while some of the time it may be weakened. Figure 5(b) shows the fitted Gaussian curve for this tag' strength. It is clear that the strength distribution is multi-modal containing more than one peak, so a single Gaussian model cannot exactly depict the truth. Actually, the multi-modal model is comprised of three single Gaussian models in the example, as shown in Figure 5(c), 5(d), and 5(e). The final distribution can be considered as a weighted mixture of three, $G = w_1G_1 + w_2G_2 + w_3G_3$. To cope with such a complicated scenario, the Mixture of Gaussian Model (MoG) based approach is employed here. We present a generalization to this approach. The strength is modeled by a mixture of K Gaussian models. The probability that the tag x has strength x_t at time t is estimated as:

$$\begin{aligned}\eta(x_t) &= \sum_{i=1}^{K} w_i \eta_i(x_t, \mu_i, \Sigma_i) \\ &= \sum_{i=1}^{K} \frac{w_i}{(2\pi)^{-\frac{d}{2}}|\Sigma_i|^{\frac{1}{2}}} \exp\left(\frac{1}{2}(x_t - \mu_i)^T \Sigma_i (x_t - \mu_i)\right)\end{aligned}$$

where w_i is the weight, μ_i is the mean and d is the reader's antenna number. Σ_i is the covariance for the i^{th} Gaussian model. Actually, the parameter x_t and μ_t are vectors with d dimension, explained later in Section III. Since the strengths are independently captured by d antennas, we set $\Sigma_i = \sigma_i^2\mathbf{I}$ where $\mathbf{I}$ is an identify matrix with d dimension. Then probability is reduced to

$$\eta(x_t) = \sum_{k=1}^{K}\prod_{j=1}^{d} \frac{w_k}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_t - \mu_i)^2}{2\sigma_j^2}\right) \quad (1)$$

The K is determined by the potential number of propagation paths. Usually, from 3 to 10 are used according to our experience. We can see the weak stability of strength



(a) Time serials

(b) G (c) G_1 (d) G_2 (e) G_3

Fig. 5. The tag was captured for 20 seconds when a person randomly walked around the tag. (a) Strength values over 20 seconds. (b) Gaussian model fitted with the values over 20 seconds. (c) The model G_1 fitted over $[10s, 14s]$. (d) The model G_2 fitted over $[4s, 8s]$. (e) The model G_3 fitted over $[14s, 18s]$.

is actually mainly determined by the environment, *e.g.* the propagation paths, instead of the interference object. In theory, *the peaks in MoG are as many as the potential propagation paths*. Each Gaussian model in MoG corresponds a potential signal propagation and its weight is related to the time interval that the propagation takes effect in this path. Note that unlike the active communication, the backscattered signals from tag are too weak to penetrate the interference objects to generate a new vibration level. In addition, despite the interference object may create a new path to backscattered the signals, their effects are very limited on the MoG, as will be shown in Section IV. This fact motivates us to build the model through the history of strength and leverage the model to perceive tag motions.

III. FROGEYE

In this section, we introduce the technique of MoG based background subtraction to perceive the tag motion via strength image. We call our approach *Frogeye*, because the natural construction of a frog retina enables its eyes to be hypersensitive to the moved objects but blind to static ones.

A. Overview

Our basic idea is to detect the 'significant' changes of the backscattered signal for perception of tag motion. There is a high probability that the tag moved when its strength changed significantly. The naive method is to compare the latest strength with the last one. If their difference exceeds a threshold, the motion is reported. However, the weak stability may trigger the significant changes as well. For example, the strength may migrate from one vibration level to another level far away when the original propagation path is blocked by the interference object. In this situation, the strength changes exhibit significant as the tag moves, giving rise to mistakenly determination, the false positive.

We find our problem is very similar to the foreground segmentation in computer vision, which is to segment the foreground pixels that "significantly differ" between the latest image of sequence and the previous images. To associate these two issues, we project the target tags into a strength image in which the strengthes of each tag are mapped to a row of pixels. The fact is that the values of these related pixels significantly change and these pixels become to foreground when the related tag moves. Therefore, as long as the foreground is detected, the moved tags are discovered. A lot of methods are proposed for foreground segmentation. We adopt the most popular one, MoG based background subtraction (MoG-BS), due to two reasons. First, MoG-BS well tackles illumination vibration and background motion. These two problems resemble our challenges incurred by the weak stability. Second, both strength and pixel value can be modeled using the MoG. In detail, the history of each a pixel is applied to build a MoG composing of K Gaussian models. When a new pixel value is captured, the value is compared with the built K models. If the value matches any one of the K models, the pixel is consider as background. Otherwise the foreground. After the foreground are segmented, we can easily reverse them to moved tags since the projection remains unchanged during the process.

B. System Architecture

Our goal is to determine whether the tag is really moved when its strength changes. We describe the high level flow of information through its architecture, and present the internal details later. Our approach contains four components. The reader repeatedly interrogates tags using multiple antennas and produces a stream of readings in its detection field. The first component is to split the original reading stream into many read cycles by time. Then the preprocessing aggregates m read cycles into a read frame. The read frame is the processing unit fed into the following components. In the second component, a strength image is constructed using the readings in one read frame. In the third component, we model each pixel using MoG to determine whether the pixel belongs to the foreground. This component outputs a serials of labels that indicates each pixel is either a foreground or background. In the last component, the system identifies the moved tags through the foreground and outputs the motion event with a specific probability to upper applications.

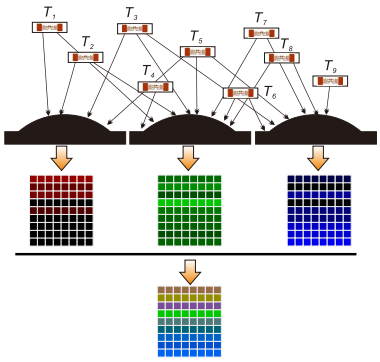


Fig. 6. The strength image

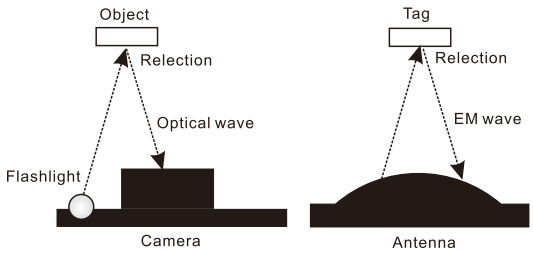


Fig. 7. Rationale behind

C. Preprocessing

The reader circularly scans the target space. Therefore, the input of our system is a stream of readings. A commercial reader usually contains d antennas successively scheduled by a time division algorithm. Therefore each read cycle can be further divided into small parts denoted as *antenna cycles*. The reader uses each antenna to collect one set of readings from the target tags during an antenna cycle. If the tag is located within the overlapping region by multiple antennas, it will be collected many times. We define a higher kernel term, denoted as *read frame*. The read frame is a set of readings collected during m consecutive read cycles where m is a user specific parameter. The main task of preprocessing is to split the strength flow into a sequence of read frames.

D. Constructing Strength Image

After separating the frames from the strength flow, we project the readings in one read frame to a strength image. Suppose the reader monitors n tags, $\{T_1, T_2, \cdots, T_n\}$, in the target space and all of them are known in advance. Then each read frame should contains $n \times m \times d$ readings. Note some tags may be beyond the range of some antenna. In this situation, the system creates a default reading for the tag, whose strength is set to the default value, *e.g.*-90. We project a read frame of readings into a strength image I as follows:

$$I = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{pmatrix}$$

In the image, each row is uniquely mapped to a same tag. The mapping fashion between the tags and rows is arbitrary as long as their mapping remains constant during the processing. Each column represents a read cycle. The whole image contains a total of m columns. Formally, given a strength image, the element $x_{i,j}$ represents a reading strength from the i^{th} tag (T_i) collected in the j^{th} read cycle of the frame. Since the tag is read multiple times by d antennas in one read cycle, the $x_{i,j}$ is indeed a vector. We consider them as the color values coming from different color channels (related to different antennas), such as red, green, blue, and alpha channels. Namely, $x_{i,j}$ is a vector with d dimension. Actually, our approach is independent on the dimension of pixel. Even if there are more than four strengths captured, they will be fully considered. An example is shown in Figure 6. There are 9 tags in the target space and three antennas are deployed. After a read frame, each antenna creates an image with 9×8 resolution ($m = 8$). These three images are synthesized to a final image, which is going to be fed into the next step (foreground detection). Through image construction, the original reading stream eventually becomes a sequence of strength images, denoted as $\mathcal{I} = \{I_1, I_2, \cdots, I_t\}$.

Why would we project the strength into an image? At the first glance, there is no connection between an optical image in computer vision and tag's signal strength. Actually, we think the physical rationale behind them are similar. As shown in Figure 7, *(i)* both of them focus on capturing the projections of objects. In optics, the light wave coming from the flashlight is reflected by the object and then captured by the camera. In our scenario, the wave becomes an electromagnetic wave. The wave is reflected by the tag and captured by the antenna. *(ii)* the optical image is an optical projection of the object while the backscattered strength can be considered as an electromagnetic projection of the tag. *(iii)* hence, the reader antenna can be considered a special 'camera' producing the tags' imaging. The optical camera outputs optical images while the RFID antenna provides a strength image. *(iv)* when the objects are moved in the real target space, the pixels reflected from the objects significantly change in the images. The similar thing occurs in our scenario. The pixels reflected from the tags change when the tag moves. Therefore, we think the strength image represents a snapshot of current tags' positions. Part of the images change when the tags move. *(v)*, importantly, as we model the weak stability of tag's strength, the pixel' intensity can be modeled by the MoG as well.

E. Foreground Detection

Foreground detection from a sequence of images is a fundamental problem in computer vision, which has been well studied for many years [5]–[7]. Its goal is to segment the foreground pixels corresponding to moving objects, such as vehicles and humans, from the rest of an image, *i.e. background*. The *foreground* is the set of pixels that "significantly differ" between the latest image of the sequence and the previous images, given a set of images of the same scene taken at several different times.

Background subtraction is the most popular method for foreground segmentation, especially under those circumstances with a relatively dynamic background. It detects the foreground in an image by taking the difference between the current image and the reference background. The main idea behind it is to

automatically generate and maintain a representation of the background that is then used to classify the pixels as background or foreground. The MoG-BS has been one of the most popular background subtraction techniques in the computer vision because of its robustness to subtle illumination changes and background motions. Here we employ and improve the MoG-BS from [8]. This algorithm is able to fast model the MoG using the training set and adaptively update the models during applying.

Modeling background with history : To make the problem more precise, we generally denote a sequence of strength images as $\mathcal{I} = \{I_1, I_2, \cdots, I_t\}$. For any pixel $I(x)$ in an image I, $x \in \mathbb{R}^l$ and $I(x) \in \mathbb{R}^d$. l is the dimension of the image. Typically, when $l = 1$, the image is one dimension and can be represented as a vector. When $l = 2$, the image is a matrix. Typically, $d = 1$ for a gray-scale image or $d = 3$ for RGB color images, but other values are allowed. At any time t, the history of a particular pixel x is

$$\{x_1, \cdots, x_t\} = \{I_i(x,), 1 \leq i \leq t\}$$

The history of such a pixel is applied to model the pixel values using the mixture of K Gaussian models. Initially, the K Gaussian models are given by an initial large variance ω_k and low prior weight σ_k. These K Gaussian models are ordered by the criterion of $r_k = \omega_k/|\sigma_k|$. This order supposes that a background pixel corresponds to a high weight with a small variance because the background is more present than foreground. On the next frame, we have a new pixel value x_{t+1}. We say the Gaussian model is 'matched' if x_{t+1} is within $\lambda\times$ standard deviation of this Gaussian model. Here, we use the Mahalanobis distance to compare two vectors. Suppose the matched model $N(\mu_{k,t}, \sigma^2_{k,t})$, the match rule is formulated as follows:

$$(x_{t+1} - \mu_{k,t})^T \cdot (x_{t+1} - \mu_{k,t}) < \lambda^2 \cdot |\sigma^2_{k,t}|$$

Considering each pixel, there are two results:

Case 1: A match is found with one of the K Gaussian models, termed as G_k. In this case, the pixel is classified as background.

Case 2: No match is found with any of the K Gaussian models. In this case, the pixel is classified as foreground.

According to the above results, the parameters of each model must be updated to make next foreground detection. Using the match result, two cases can occur such as in the foreground detection:

Case 1: A match is found with one of the K Gaussian models. For the matched model G_k, we increase its weight, adjust the mean closer to x_t, and decrease the variance as follows.

$$\omega_{k,t+1} = (1 - \alpha) \times \omega_{k,t} + \alpha$$

where α is a learning rate.

$$\begin{aligned} \mu_{k,t+1} &= (1 - \gamma) \times \mu_{k,t} + \gamma \times x_{t+1} \\ \sigma^2_{k,t+1} &= (1 - \gamma) \times \sigma^2_{k,t} + \gamma \times (x_t - \mu_{t+1})^2 \end{aligned}$$

where $\gamma = \alpha \times \eta\left(x_{t+1}|(\mu_k, \sigma_k)\right)$ and $\eta(\cdot)$ refers to Equation 1. For the unmatched models, we decrease their weights:

$$\omega_{i,t+1} = (1 - \alpha) \times \omega_{i,t}$$

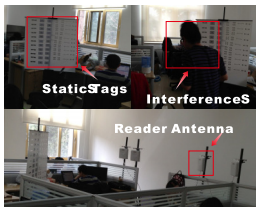


Fig. 8. Stationary scene

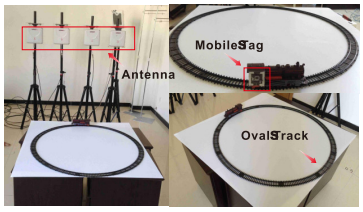


Fig. 9. Mobile scene

Case 2: No match is found with any of K Gaussian models. In this case, the model with least order is replaced by a new model with a mean equals to the new pixel value x_{t+1}, and initial large variance and low prior weight.

Foreground detection : Above algorithm is an adaptive process. In the training phase, we cloud set a bigger learning rate. The training phase ends when the K models become stable. In applying phase, we stop learning by setting the learning rate to zero. With regard to the dynamical situation in which the environment slightly changes, the smaller learning rate, *e.g.* 0.001, is allowed to let the system self-adaptively update. Especially it is able to accommodate the influence of the new few propagation paths produced by the interference objects. The foreground detection outputs a set of pixels labeling as foreground.

F. Motion Refinement

Since we have a row of pixels mapping to one tag, the results of foreground detection are a mixture of background and foreground for each row. Suppose there are f pixels belonging to the foreground out of m values in a particular row. We define the probability of tag motion as $p = f/m$. Therefore, the outputs of foreground detection for each read frame are reversely mapped to a vector:

$$\{< T_1, p_1 >, < T_2, p_2 >, \cdots, < T_n, p_n >\}$$

where T_i is a reversely mapped tag and p_i denotes the probability that this tag moves. The simple method is to use majority voting to fuse the mixture in which the tag is determined moved only when $p \geq 0.5$.

IV. Implementation and Evaluation

In this section, we present the implementation and also conduct performance evaluation on the prototype.

A. Implementation

Hardware: In our implementation, we employ R420 reader from ImpinJ. Total 100 tags with 4 kind of models are used. Four reader antennas polarized in horizontal direction are deployed in a line. The transmission power is adjusted at $30mW$.

Software: We adopt LLRP protocol to communicate with the reader. This protocol was ratified by EPCglobal in April 2007. We adjust the configuration of reader to immediately report reading whenever tag is detected. The client code is implemented by Java language with help of Apache Math library.

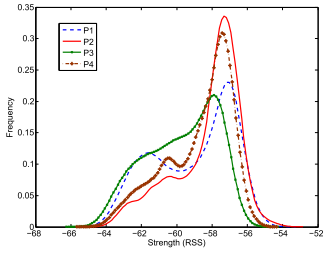


Fig. 10. Interference similarity

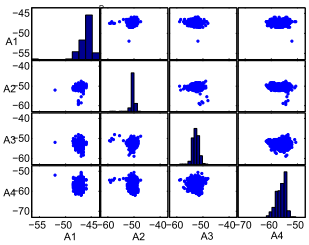


Fig. 11. The strength scatter matrix

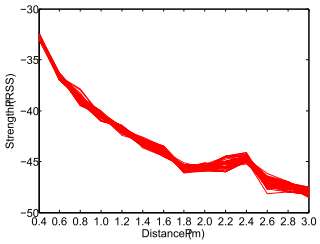


Fig. 12. Strength vs. distance

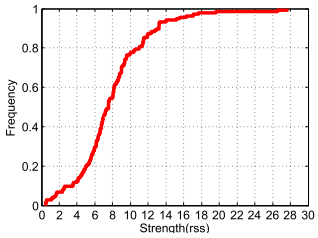


Fig. 13. CDF of sensitivity

Parameter choices: Frogeye comes with a number of different parameters. Four key parameters are employed as follows. Frame size is set to $m = 10$. The learning rate α is set to 0.1 for training phase but 0.001 when detecting. The number of Gaussian K is a user estimated parameter, which really depends on the environment. In experiment, we set $K = 10$ to adapt potential propagation paths. Note that it does not matter to use large K because the redundant models will be naturally ignored according to our algorithm. Since we use the ROC curve to measure the accuracy, the match threshold λ is considered as the operating parameter varying from 0.01 to 2.0.

B. Evaluation Methodology

In terms of the false positives, we deploy a white board on which 100 tags are attached, in the office, as shown in Figure 8. There are 20 persons staying in the room. These persons, whose weights and heights vary in $60kg \sim 90kg$ and $1.6m \sim 1.8m$, may interfere the signal strength when getting close to the board. We use four antennas continuously reading these tags for 24 hours. A total of $1,130,997$ readings are collected. The traces collected in the first 6 hours are employed as the training set to model background models for each tag, and the left 18-hour traces are fed for testing the false positives. On the other hand, to measure the true positives, we attach tags on a toy train which moves along an oval track in a constant speed $(0.7m/s)$. The track illustrated in Figure 9 has a decimeter of $1m$. We collect traces for 5 minutes as training set before the train takes movements. The number of true positive rate equals the number of perceived motions divided by the total read times when the tag moves.

Metrics: We express the perception accuracy by using the *Receiver Operating Characteristic* (ROC) curve. The vertical axis of the ROC curve is the true positive rate (TPR) which is the total number of perceived true positives divided by the number of real positives in ground truth. The horizontal axis of the ROC curve is the rate of false alarms, *e.g.* false positive rate (FPR), divided by the number of negative events in ground truth.

We define the *sensitivity* to measure the sensibility of backscattered signals. This metric is defined as the absolute the ratio of RSS difference to the distance of two positions, formulated as follows:

$$\text{Sensitivity} = \frac{|RSS_A - RSS_B|}{\sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2}}$$

where RSS_A and RSS_B are the strength values respectively collected from position $A(X_A, Y_A)$ and position $B(X_B, Y_B)$. The unit is rss/m, namely, the strength changes (rss) in unit distance (m).

The *minimum perception displacement* (MPD) is used as the second metric to measure the system sensibility. MPD is the distance from the position where tag is trained to the one that the tag is firstly perceived in motion,

Baseline: For comparison, we also employ other three methods to perform the evaluation.

(1) *Localization*. There are many localization methods proposed in the literature. We use the most popular one in RFID system, LANDMARC [2]. Its basic idea is to calculate the weighted average of the four nearest tags' locations. The weights are determined by the rss values. (2) *The minimum variance (MV)*. In this method, the latest 50 strengths are buffered to dynamically compute the current mean value. If the new strength deviates from the mean over than a threshold, the tag is determined in motion. (3) *Frame differencing based subtraction (FD-BS)*. The simplest method used for foreground segmentation in computer vision. The foreground is determined if the continuous two frame image difference is greater than a threshold.

C. Perception Similarity

Intuitively, the individual differentiation may incur different interference impacts. Can we use the interference impacts from one object to model the others? To answer this question, we let four persons independently impose interference on the strength. These persons' weights and heights are $P_1(74.5kg, 173cm)$, $P_2(58.5kg, 164cm)$, $P_3(64kg, 170cm)$, and $P_4(69kg, 171cm)$. Figure 10 plots the fitted probability densities of strength interfered by these four persons. Visual inspection shows that the four curves nearly have the same shape. All of them can be modeled by two Gaussian models. The only difference is that the individual Gaussian model has different weights. As discussed in Section II, the weight of every individual Gaussian model in MoG is highly related to the time interval that the model takes effect. In our algorithm, the weight only affects the matching order rather than the result. Our experiment fully shows the interference object takes very limited impacts on the models.

Figure 11 plots seven-hour data containing $57,603 \times 4$ readings for a stationary tag collected by four antennas. In the figure, four histograms represent the strength statistics respectively collected by the four antenna $A1 \sim A4$. The scatter matrix illustrates the strength results collected by any two antennas. From the figure, we observe that either the strength readings from one antenna or the combination from

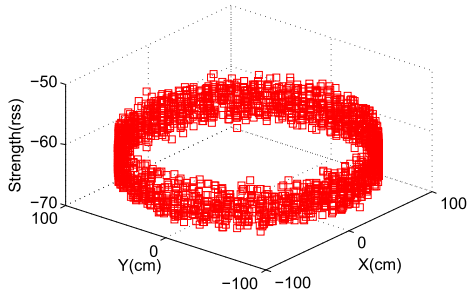


Fig. 14. Tracked strength

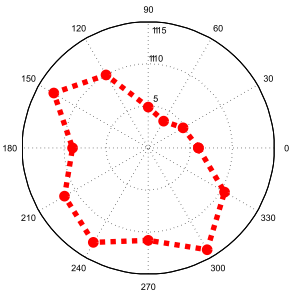


Fig. 15. MPD

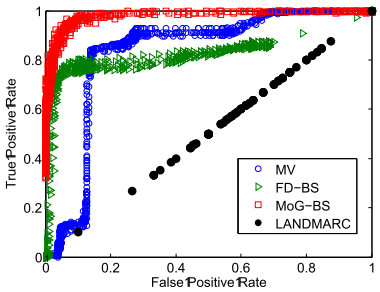


Fig. 16. Accuracy

any two antennas are highly clustered. There are little points beyond the clusters. Although the four antennas are deployed in a line and very close (interval of $\sim 30cm$ for each other), the shape of cluster exhibits a little difference. This shows the strength distribution depends on the space relationship between reader antennas and tags. It further validates our hypothesis that the Gaussian model can depict the tags' positions is reasonable.

D. Perception Sensitivity

To verify the sensibility, we change the distance between the antenna and tags from $30cm$ to $300cm$. The strength results are plotted in Figure 12. In total, the curve decreases as the distance increases and the slope of the curve equals 0.25 approximately, which well follows the theory (Note that the RSS is the log-scale presentation of power). On the other hand, we also calculate the sensitivity of any two points out of 30 points in the curve. The CDF of the sensitivity results are shown in Figure 13. We observe that 50% sensitivity achieves about $7rss$ changes and about 10% cases have $13rss$ above changes in the unit distance. Correspondingly, the received powers have $5\times \sim 20\times$ difference because of the following equation.

$$\frac{P_A}{P_B} = 10^{\frac{1}{10}(10\log P_A - 10\log P_B)} = 10^{\frac{1}{10}(RSS_A - RSS_B)}$$

Such differences are so noticeable that any commercial device is able to detect them.

Considering the toy train, we can calculate its read positions according to the read time and the train's speed. Figure 14 plots the tracked strength along the track. We can see that the strength changes in a rather wide range ($\sim 15rss$) even when the neighboring positions are very close. We think the tag's orientation also contributes a lot to these changes during the movement, because it is a well-known fact that the orientation seriously affects the strength. The tag's orientation is defined by the angle between reader and tags. Note that the tag orientation does not influence our approach because the stationary tag does not change its orientation. By contrary, it may help motion perception when the tag moves with the orientation changes.

To measure the sensitivity without orientation' influence, we keep the tag orientation unchanged and move the tag deviating from the trained position. Figure 15 shows the MPD in various directions where the antennas locate in the north. We can see that (1) the MPDs in different direction are not the same. (2) The directions between $(-60^\circ \sim 60^\circ)$ are the most sensitive in which the average MPD equals $6cm$. This is because the tag in our experimental room is closer to the west wall. The more complex environment introduces more potential propagation paths of signal, requiring more Gaussian models to describe. Obviously, the sensitivity decreases when more models are used, because the real motion may be mistakenly determined as being interfered. (3) The average MPD is $10cm$, which means the motion can be perceived by the system as long as the tag moves away by $10cm$. The result is rather hypersensitive.

E. Perception Accuracy

In this section, we apply the four methods to data traces and measure their perception accuracy. The ROC curves by these methods are plotted in Figure 16. We can see that any of strength changes enabled methods has a better performance compared with localization. In detail, given a FPR of 0.5%, MoG-BS achieves 92.34% TPR while FD-BS and MV respectively have 75.6% and 3% FPR. The accuracy of MoG-BS is far better than two others. The MV uses several of the latest strength history to predict the tag motion. However, when the tag moves in a regular mode like the train, its changes is easy to be hidden by the latest history. So there are two obvious steps in the figure. In the first step, the threshold is too large to detect the motion, but the motion is suddenly detected when the threshold drops to a special value. Being different with MV, FD-BS only uses the last one history strength, hence its accuracy is higher than MV from the beginning because it has ability to resist the regular motion. However, FD-BS has a fata drawback that there is a chance to find a movement way enabling the same strength changes in any two adjacent read positions. In this situation, FD-BS fails to detect the motion. MoG-BS conquers the above weakness because it learns from the environment instead of the tag's current status. As long as the environment is unchanged, its perception accuracy always remains in a high level. LANDMARC is the worst method behaving as a random coin because its localization accuracy is around $2m$ but our movement is confined in $1m$.

V. Related Work

Localization: Early work on localization relied on RSS and proximity [2], [3], [9]–[11]. Zhu *et al.* propose a fault-tolerant localization approach for RFID reader [3]. Shangguan *et al.* [9] study on the problem of tag order under mobile environment. Sen *et al.* [10] offers a new localization algorithm using the PHY layer information. Liu *et al.* [11] propose a novel

localization approach that utilizes the interference to position the target. **Detection of missing tags:** The existing work [12]–[16] utilize the presence-or-absence fashion to monitor the tag's situation. Tan [12] *et al.* is the first to address this problem . In practice, it is necessary to report the missing ones without collecting all tags, because the later must stay in the range by default. Li [13] *et al.* further study the issue but propose a deterministic algorithm which reduces the time by 88.9% more. Luo [14] *et al.* consider the energy perspective for battery-powered active tags when detecting the missing tags. Zheng [15] *et al.* presents a physical layer missing tag identification scheme which effectively makes use of the lower layer information and dramatically improves operational efficiency. Gong [16] *et al.* propose a accurate approximation scheme for large-scale RFID cardinality estimation. As stated before, our work has a futher fine-grained surveillence on the target. **Fast identification:** Anti-collision is another important topic in RFID area which aims to conduct the fast identification [17]–[24]. Chen [17] *et al.* address the problem of collecting information from sensor-augmented tags. They use multi-hash to spread the transmission slots. Xie *et al.* consider how to efficiently identify tags on the moving conveyor [18]. Jeffery [19] *et al.* design a statistical method to clean the uncertain tag readings. Yang [20] *et al.* and Tang [21] *et al.* discuss how to improve the throughput by avoiding the reader collision. Yang [22] *et al.* introduce the fast identification in anti-counterfeiting for a batch of rfid tags. Liu [23] *et al.* presents a new method to conduct adaptive continuous ccanning in large-scale RFID systems. **Polling query:** Assuming all tag IDs in the range are known in advance, these papers [25]–[27] design protocols of polling queries to identify concerned tags with no need to collect them all. Shen *et al.* is one of the earliest works abandoning collecting all tags but designing polling query to the target tags whose categories are popular [25]. Qiao *et al.* propose tag-ordering polling protocol that can reduce per-tag energy consumption by more than an order of magnitude [26]. Gong *et al.* [27] present batch authentication by informative counting. **Foreground segmentation:** Detecting regions of change in multiple images of the same scene taken at different times is of widespread interest in computer vision [5], [8]. The basic adaptive model is used in [5]. Chris and Grimson [8] discuss modeling each pixel as a mixture of Gaussians and using an online approximation to update the model.

VI. Conclusion

Real-time object surveillance is an important task in RFID system. Existing methods in RFID systems often employ presence or absence fashion to detect the tags motions, so they cannot meet motion detection requirement in many applications. In this paper, we offer an innovate approach to explore the tag motion through its backscattered signal changes. We believe this is a step forward in the area of real-time monitoring using RFID technique.

Acknowledgment

This work is supported in part by SF China Major Program 61190110, NSF China under Grant No.(61272460,60933003) and Ph.D. Programs Foundation of Ministry of Education of China under Grant No. 20120201110010. We acknowledge the support from the codes of USRP2reader from the Open RFID Lab (ORL) project for protocol analysis [28].

References

[1] M. Bouet and A. L. dos Santos, "Rfid tags: Positioning principles and localization techniques," in *IFIP Wireless Days*, 2008.

[2] L. Ni, Y. Liu, Y. Lau, and A. Patil, "Landmarc: indoor location sensing using active rfid," *Wireless networks*, vol. 10, no. 6, pp. 701–710, 2004.

[3] W. Zhu, J. Cao, Y. Xu, L. Yang, and J. Kong, "Fault-tolerant rfid reader localization based on passive rfid tags," in *Proc. of IEEE INFOCOM*, 2012.

[4] D. Dobkin, *The RF in RFID: passive UHF RFID in practice*, 2008.

[5] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Real-time tracking of the human body ," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997.

[6] N. Friedman and S. Russell, "Image segmentation in video sequences: A probabilistic approach," in *Uncertainty in Artificial Intelligence*, 1997.

[7] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Transactions on Systems, Man, and Cybernetics*, 2004.

[8] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. of IEEE CVPR*, 1999.

[9] L. Shangguan, Z. Li, Z. Yang, M. Li, and Y. Liu, "Otrack: Order tracking for luggage in mobile rfid systems," in *Proc. of IEEE INFOCOM*, 2013.

[10] S. Sen, B. Radunovic, R. R. Choudhury, and T. Minka, "Spot localization using phy layer information," in *Proc. of MobiSys*, 2012.

[11] Y. Liu, Y. Zhao, L. Chen, J. Pei, and J. Han, "Mining frequent trajectory patterns for activity monitoring using radio frequency tag arrays," *IEEE TPDS*, vol. 23, no. 11, pp. 2138–2149, 2012.

[12] C. C. Tan, B. Sheng, and Q. Li, "How to monitor for missing rfid tags," in *Proc. of IEEE ICDCS*, 2008.

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

[14] W. Luo, S. Chen, T. Li, and S. Chen, "Efficient missing tag detection in rfid systems," in *Proc. of IEEE INFOCOM*, 2011.

[15] Y. Zheng and M. Li, "P-mti: Physical-layer missing tag identification via compressive sensing," in *Proc. of IEEE INFOCOM*, 2013.

[16] W. Gong, K. Liu, X. Miao, and H. Liu, "Arbitrarily accurate approximation scheme for large-scale rfid cardinality estimation," in *Proc. of IEEE INFOCOM*, 2014.

[17] S. Chen, M. Zhang, and B. Xiao, "Efficient Information Collection Protocols for Sensor-augmented RFID Networks," in *Proc. of IEEE INFOCOM*, 2011.

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

[19] S. R. Jeffery, M. Garofalakis, and M. J. Franklin, "Adaptive cleaning for RFID data streams," in *VLDB*, 2006.

[20] L. Yang, J. Han, Y. Qi, W. Cheng, T. Gu, and Y. Liu, "Season: Shelving interference and joint identification in large-scale rfid systems," in *Proc. of IEEE INFOCOM*, 2011.

[21] 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 IEEE ICNP*, 2009.

[22] L. Yang, J. Han, Y. Qi, and Y. Liu, "Identification-free batch authentication for rfid tags," in *Proc. of ICNP*, 2010.

[23] H. Liu, W. Gong, X. Miao, K. Liu, and W. He, "Towards adaptive continuous scanning in large-scale rfid systems," in *in Proc. of IEEE INFOCOM*, 2014.

[24] S. Qi, Y. Zheng, M. Li, L. Lu, and Y. Liu, "Collector: A secure rfid-enabled batch recall protocol," in *Proc. of IEEE INFOCOM*, 2014.

[25] B. Sheng, C. Tan, Q. Li, and W. Mao, "Finding popular categories for RFID tags," in *Porc. of ACM MobiHoc*, 2008.

[26] Y. Qiao, S. Chen, T. Li, and S. Chen, "Energy-efficient polling protocols in RFID systems," in *Proc. of ACM MobiHoc*, 2011.

[27] W. Gong, K. Liu, X. Miao, Q. Ma, Z. Yang, and Y. Liu, "Informative counting: fine-grained batch authentication for large-scale rfid systems," in *Proc. of ACM MobiHoc*, 2013.

[28] "Open rfid lab," http://pdcc.ntu.edu.sg/wands/ORL.