Efficient Localization based on Imprecise Anchors in RFID System

Xiang Lu, Lei Xie, Yafeng Yin, Wei Wang, Baoliu Ye, Sanglu Lu
State Key Laboratory for Novel Software Technology, Nanjing University, China
Email: {luxiang, yyf}@dislab.nju.edu.cn, {lxie, ww, yebl, sanglu}@nju.edu.cn

Abstract—With the rapid proliferation of RFID-based applications, RFID tags have been deployed into pervasive spaces in increasingly large numbers, e.g., the shelves of super markets are filled with tag-labeled items. Conventional localization schemes usually leverage precise anchor nodes to help compute the position of objects. However, it is usually difficult to find or deploy enough anchor nodes for accurate localization. In this paper, we propose solutions to locate the mobile users based on imprecise anchors in RFID systems. A large number of tags with approximate locations are used as anchor nodes to compute the user’s locations. We thus present a time-efficient localization scheme to continuously tracking the mobile users. Experimental results indicate that our solutions can accurately locate the mobile users in a real-time approach. The improved method’s accuracy is more than 30% better than the base solution.

Index Terms—RFID; Tags filled environment; Localization; Reader-based; location estimation

I. INTRODUCTION

The proliferation of wireless and mobile devices has fostered the demand for context-aware or location-based services in pervasive applications [1–4]. For example, while a user is shopping in the super market, the relevant advertisements can be popped up in time when he is approaching a shelf with specified goods. In regard to these applications, location is viewed as one of the most significant factors. Conventional localization schemes mainly leverage precise anchor nodes to help compute the user’s location [5–8]. These anchor nodes are intentionally deployed for localization. However, it is difficult to find or deploy enough anchor nodes for accurate localization, due to the very expensive deployment cost.

With the rapid proliferation of RFID-based applications [9–13], RFID tags have been deployed into pervasive spaces in increasingly large numbers, e.g., in the libraries or super markets, the shelves are always filled with tag-labeled items. This provides us a new opportunity for localization. Since the same type of items are usually placed in a specified range, the location of these items can be approximately known according to the layout of the merchandise. For example, considering the case where Crest toothpastes are known to be placed in a certain range \((x_a, x_c)\) of a shelf. If the RFID reader has scanned tags with category identification to the Crest toothpastes, then the reader is supposed to be close to this range, as the reader usually has a rather limited scanning range. Therefore, if a mobile user carries an RFID reader while he is walking around, e.g., the reader is deployed in the shopping cart, it is possible to locate the user according to the categories of the scanned tags and the layout of the merchandize. When multiple categories of tags are scanned, the reader’s location can be computed more accurately by sufficiently leveraging the layout information of these categories.

Based on the above understanding, we propose solutions to locate the mobile user based on imprecise anchor nodes in RFID systems in this paper. Such localization solutions need to resolve two key challenges in order to provide precise and timely user location. Firstly, the anchor positions are imprecise. Items of the same category can be placed in a range rather than on a precise point on the shelf. Furthermore, the reader may miss some of the nearby RFID tags due to imperfect radio environments. Our localization scheme must handle such uncertainty to provide reliable user locations. Secondly, most RFID readers take more time to scan nearby RFID tags when the density of RFID is increased. As our system utilizes a large number of existing RFID tags for localization, we need to find a way to return useful information in a timely manner even if the reader is surrounded by thousands of RFID tags.

We thus present a time-efficient localization scheme to continuously tracking the mobile user. Experimental results indicate that our solutions can accurately locate the mobile users by a real-time approach. The main contributions of this paper are summarized as follows:

- We propose an efficient localization scheme based on the large number of tags scanned by the RFID reader, it does not require any additional precise anchor nodes, thus effectively avoids the significant deployment cost. To the best of our knowledge, this is the first localization work based on imprecise anchor nodes in RFID systems.
- In order to continuously track the mobile user, we propose two time-efficient algorithms to accurately locate the moving user, i.e., the Category Cardinality based Protocol and the RSSI based Protocol. By reasonably adjusting the reader’s power, the localization can be executed in a real-time approach within the specified time delay.

II. RELATED WORK

Many research works using RFID for indoor localization have been carried out in recent years [14–18]. Liu et al. propose a new localization algorithm (ARSS) based on RSSI, ARSS figures out the distribution relationship between unknown nodes and anchor nodes [14]. Babic et al. propose a
novel localization method for moving object using integration of passive RFID tags and scene analysis technique [15]. LANDMARC [16] is a tag localization prototype in indoor environment. By utilizing extra reference with fixed reference tags to help location calibration, it can increase location accuracy without deploying large numbers of RFID readers. Zhu et al. propose a fault-tolerant RFID reader localization approach to solve the problem of frequent occurred RFID faults [17]. Xiao et al. propose a novel environmental-adaptive indoor positioning approach using RSSI [18]. The signal propagation model and model parameters are updated in a closed-loop feedback correction manner.

III. PROBLEM FORMULATION

Shopping malls, warehouses and public libraries, may have massive RFID tags deployed in the environment (There are many objects or books attached with tags on the shelves). Our objective is to use these existing tags to localize the mobile user equipped with an RFID reader, e.g., the reader may be attached to a shopping cart. When the user moves with the reader in these environments, the reader can interrogate the tags within the scanning range and get their identities for localization.

As shown in Fig.1, there are many shelves in the localization environment. The star is the target needed to be located. As the shelves’ locations are fixed, we can use one-dimension coordinate system to describe the target’s position. In the figure, the $y_0$ is known. By estimating the position $x$ in the aisle, we finally get the exactly location of the target.

As shown in Fig. 2, the shelf is divided into several blocks and each block contains multiple objects represented by small cubes. The blocks are distinguished by their colors in the figure. The objects in one block belong to the same category, which is represented in tag ID. The width of the blocks can also be different from each other. We describe a block as $[x_{i,s}, x_{i,e}, l_i]$. $x_{i,s}, x_{i,e}$ are the start point and end point of the block, $l_i$ represents the layer of the shelf, and the height of each shelf layer is $h$. The exact position of each object in the block is unknown, while the location of each block is known.

IV. OBSERVATIONS FROM THE REALISTIC EXPERIMENTS

Because of the issues like path loss, energy absorption and mutual interference, the RSSI distribution are always not idealized. In order to get these information in the realistic environment, we provide several observations from 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 65 degrees. The RFID tags used are Alien 9640 general-purpose tags which support the EPC C1G2 standards. We place the tags on a shelf with 5 rows, the distance between two rows is 60cm.

In the following experiments, we vary the tag density from 2 to 10 tags per meter, while adjusting the reader’s power from 15.7dBm to 30.7dBm. Unless otherwise specified, in default situation we fix the reader towards the center of the shelf, scan the tags for 50 query cycles repetitively. And the distance between the antenna and the shelf is 1.5m.

TABLE I

<table>
<thead>
<tr>
<th>Table Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_a$</td>
<td>Detection regions’ radius in different tag density and power.</td>
</tr>
<tr>
<td>$T_b$</td>
<td>Relationships among tags size, Power and Tag density.</td>
</tr>
<tr>
<td>$T_c$</td>
<td>Relationship between detected tags size and scanning time.</td>
</tr>
<tr>
<td>$T_d$</td>
<td>Relationship between RSSI value and distance to the center of the detection region.</td>
</tr>
</tbody>
</table>

1) Tag density effects: In the same scanning range, the density $\rho$ of the tags is the main influencing parameter. A high density $\rho$ causes more tags to be identified in the range. Meanwhile, as the tag density increases, the major detection regions radius gradually decreases. Fig.3 shows the relationship of detection region, reader’s power, and tag density. We can find that when the power is fixed, as the tag density increases, the width of the detection region decrease. The Fig.4 shows the relationship of the number of identified tags, the reader’s power, and the tag density. Based on these two relationships, we make two training data sets $T_a$ and $T_b$ by experiments. Then we can compute the tag density and the...
radius of the major region with the known reader’s power by linear interpolation method.

2) Scanning time: When the number of tags becomes large, the identification time increases, as shown in Fig.5. If we want to localize objects in a very short time by reading the tags, then the number of tags in the detection region should be limited. For example, when the number of tags equals to 120, the reading time approximates to 1.5s, which cannot be accepted in the real-time location system. Therefore, we should reduce the number of identified tags (e.g. 90) in the detection region to reduce the reading time (e.g., 1s), which can be achieved by decreasing the scanning power. We get a training data set 
\[ T_c \]
which shows the relationship between detected tag size and scanning time.

3) RSSI distribution: In regard to the tags in the same detection region, their values of RSSI are different from each other. Fig.6 shows the RSSI distribution for tags on the single row. We can find that the tags in the major detection region have the higher RSSI values, while the tags in minor detection region have the lower RSSI values. In fact, the RSSI value is mainly affected by the distance and the angle between the tag’s surface and the antenna’s radiation direction towards the reader. The shorter distance to the center of the detection region, the higher RSSI value we will get. Here, we get a training data set 
\[ T_d \]
which shows the relationship between RSSI and distance to the center of the detection region.

V. Baseline Solutions

A. Category Cardinality based Protocol (CCP)

In this baseline solution, we set the reader’s power to the default value(30.7dbm). Then the reader scans the tags to get the categories of identified tags. We average the locations of the detected categories to get the estimated position. We use the number of identified tags to calculate the weight. Because the start point \( x_{i,s} \) and the end point \( x_{i,e} \) of the category’s range are known, the estimation method could be expressed as follows.

\[
x = \frac{\sum_{i=1}^{k} w_i \cdot x_{i,s} + x_{i,e}}{\sum_{i=1}^{k} n_i},
\]

\[
w_i = \frac{n_i}{\sum_{i=1}^{k} n_i}
\]

Here, \( k \) is the number of categories for identified tags. \( w_i \) is the weight factor for category \( c_i \), which is ratio of the number of identified tags in category \( c_i \) to the number of all the identified tags in the scanning area. In this method, we use the middle value of the category’s range \( \frac{x_{i,s} + x_{i,e}}{2} \) to represent the category’s position. Finally, \( \hat{x} \) is the computed position of the target.

B. RSSI based Protocol (RP)

In the previous baseline solution, we just use the number of identified tags in each category. However, the RSSI is a very important factor to measure the distribution of tags. As mentioned in Section IV, the higher the RSSI, the closer the tag to the center of the detection region.

With this knowledge, several methods based on RSSI can be used to compute the real position of the reader. The basic one is like the previous method, which just replaces the tag size \( n_i \) with total RSSI \( s_i \) in each identified category, and then calculate the weighted mean of the identified categories’ positions in x coordinate as \( \frac{x_{i,s} + x_{i,e}}{2} \). The estimation method could be expressed as follows.

\[
x = \frac{\sum_{i=1}^{k} w_i \cdot x_{i,s} + x_{i,e}}{\sum_{i=1}^{k} s_i},
\]

\[
w_i = \frac{s_i}{\sum_{i=1}^{k} s_i}
\]

These basic schemes by using default power to read as many tags in different categories as possible. As mentioned before, the larger power causes the number of tags in the scanning range to increase, therefore the reader needs more time to finish the scanning process.

Besides, in the tag size based baseline solution, the weight is not reliable due to the variances in tag distributions. For instance, maybe there is a category containing just a small number of tags, thus, even all the tags in the category are identified, it always has a small weight in the computing process. In the RSSI based baseline solution, when the categories in the major region contain just a small number of tags, but the categories in the minor region contain a lot of tags. The imbalanced numbers of tags will skew the weight and cause large error in the estimation position.

Therefore, these limitations lead to the low efficiency of the baseline solutions.

VI. Localization based on Imprecise Anchors

A. Compute the optimal value of power

Our algorithm tries to reasonably adjust the power of the reader, while keeping it at a reasonable level to make sure the delay satisfy the limits \( t_l \). As shown in Algorithm 1, we use four steps to compute the optimal power. First, we use a pre-scan process to get the identified tag size. By setting the pre-scan power to an empirical value \( p_0 \), we can get an identified tag size \( n_0 \). Second, based on the training data set \( T_h \), we can compute the tag density \( \rho \) by an interpolation method.(For example, when the power is 26.7dBm and the number of identified tags is 50, the intersection is between 5 tags/m and 10 tags/m, and it is closer to 10tags/m, thus we set it 8 tags/m.) Third, suppose that we take time \( t_0 \) in the pre-scan, then we have the time of \( t_l - t_0 \) left. As mentioned in Section IV, there is a relationship between the identified tag size and scanning time as shown in Fig.5. Therefore, we can get an target tag size limit of \( n' \) with respect to the remaining time of \( t_l - t_0 \) by an interpolation method based on the training data set \( T_c \). Finally, from the relationship shown in Fig.4, by using the \( \rho \) and \( n' \), we can compute the optimal power \( \hat{p} \). The algorithm is shown in Algorithm 3.

B. Category Matching based Protocol (CMP)

1) Motivation: During the scanning process, we can get the number of tags in each identified category. We denote them
Algorithm 1 Compute the optimal value of power
1: INPUT: $t_l$: the limit of the delay; $p_0$: the power of the pre-scan;
2: PROCEDURE
3: Set reader’s power to $p_0$, get the identified tag size $n_0$.
4: From the relationship between identified tags size $n_0$ and power $p_0$ in different densities which is showed in training data set $T_c$, compute the tag density $\rho$ by intersection.
5: From the relationship between identified tags size $n_0$ and scanning time ($t_l - t_0$) which is showed in training data set $T_c$, get the acceptable tag size $n'$ by intersection.
6: Use the relationship in the training data set $T_b$, again, we can compute the $\hat{p}$ with the tags size $n'$ and tag density $\rho$ by intersection.
7: OUTPUT: the optimal value of power $\hat{p}$

Algorithm 2 Category Matching based protocol
1: INPUT: $V_0 = \{n_{0,1}, n_{0,2}, ..., n_{0,s}\}$: the vectors for the numbers of tags in each identified category; $C_i = \{x_{i,1}, x_{i,r}, l_i\}, i \in \{1, 2, ..., N\}$: the position of tag category $C_i$;
2: PROCEDURE
3: Call Alg.1 to get the optimal power $p_0$, tag density $\rho$.
4: Based on the $T_2$, get the minor region radius $r$.
5: Get the left point $x_l$ and the right point $x_r$ of the covered range by the identified categories.
6: $x = x_l + r; i \leftarrow 1$;
7: while $x < (x_r - r)$ do
8: Get the $V_i$ at each potential position;
9: Compute the similarity between $V_0$ and $V_i$ as follows:
$$\text{sim}(V_0, V_i) = \frac{V_0 \cdot V_i}{|V_0| \cdot |V_i|} = \frac{\sum_{j=1}^{s} n_{i,j} \cdot n_{0,j}}{\sqrt{\sum_{j=1}^{s} n_{0,j}^2} \cdot \sqrt{\sum_{j=1}^{s} n_{i,j}^2}}$$
10: $x \leftarrow x + \Delta d; i \leftarrow i + 1$;
11: end while
12: Sort the value of $\text{sim}(V_0, V_i)$ in descending order. Find the first $k$ nearest position $x_1, ..., x_k$ according to $\text{sim}(V_0, V_i)$.
13: Compute $\hat{x} = \sum_{i=1}^{k} x_i \cdot w_i$, here
$$w_i = \frac{1/(1 - \text{sim}(V_0, V_i) + \varepsilon)}{\sum_{i=1}^{k} 1/(1 - \text{sim}(V_0, V_i) + \varepsilon)} (\varepsilon > 0)$$
14: OUTPUT: The computed position $\hat{x}$.

Fig. 7. Matching the identified tags’ number as a vector. We treat the positions covered by the identified categories as the potential positions. From the left to the right of these potential positions, by making each potential position as the center of a circle and the major detecting region’s radius as the circle’s radius, we can calculate a series of tag size vectors for the identified categories as shown in Fig.7. By matching each calculated vector with other identified vectors, we can get $k$-nearest vectors which have the highest similarities, as shown in Algorithm 2. Based on the $k$-nearest neighbor algorithm, we can compute the final position.

2) Algorithm: As shown in Fig.7, the circle filled with light color represents the major detecting region. The $r$ is the major detecting region’s radius. By the scanning results, we can get the major detecting region of each categories $C_i$ and the tag size $n_i$ in each category. We denote these numbers of tags in each category $C_i$ as a vector $V_0(n_{0,1}, n_{0,2}, ..., n_{0,s})$. We calculate the detection region with the known tag density and reader’s power. Based on training data set $T_a$, we can get the the minor region radius $r$. Dotted line circles represent detection ranges of other candidate positions. Through the linear interpolation method based on training data set $T_b$, we can get the estimated tag density $\rho$. We calculate the vectors through the geometric method. As shown in Fig.7, the dotted line circles cover a series of categories, the location information and the estimated tag densities are known, then we can get the number of tags in each category. We denote them as vector $V_i(n_{i,1}, n_{i,2}, ..., n_{i,s})$. Finally, by calculating the similarity between $V_0$ and $V_i$, we can get $k$-nearest vectors’ positions $x_1, ..., x_k$. Then, using an inverse distance weighted average with the $k$-nearest multivariate neighbors, we can get the target’s position $\hat{x}$. As shown in Algorithm 2.

3) Analysis: This method is the improvement of the Tag size weighted average based method. It fully uses the known conditions and the results of the scanning to compute the tag density and the read rate of each category. By using vector matching method to get the $k$-nearest neighbours, the accuracy and stability of the method is guaranteed. However, there are some limitations too. It needs to get parameters of the scene-setting, such as the distributions of detection regions. It also needs to compute the tag density at first, which will cause some errors.

C. Distance Voting based Protocol (DVP)

1) Motivation: To solve the limitation of the RSSI weighted average based solution, we get the relationship between RSSI value and distance to the center of the detection region in different tag density. Then we can compute the distance between each tag and the center of the detection region. Take the position of the category’s midpoint as the each category’s position, then we can get a series of potential position intervals. The intersection of these intervals is the estimation region we want.

2) Algorithm: Through algorithm 1, we can get the optimal value of the power $p_0$. Meanwhile, we can get tag density $\rho$, it
is an intermediate result in algorithm 1. By setting the power to \( p_0 \) and scan the tags, we get the identified categories \( C_i \) and the RSSI value \( s_{i,j} \) of each tag in the category \( C_i \). Through the training data set \( T_d \), we can get the distance \( d_{i,j} \) of the tag to the center of the detection region. Take the position \( x_{i,s} = \frac{x_{i,e} + x_{i,e}}{2} \) of the category’s midpoint as the tag’s position, then the estimation position \( \hat{x}_i \) calculated by the category \( C_i \) is \( \hat{x}_i = \frac{x_{i,s}}{n_i} + \sum_{j=1}^{n_i} d_{i,j} \). The category contains the minimal \( d_{i,j} \) is the most possible category which contains the real position \( x_0 \), we can ensure the relative positions of the categories. Then the estimation position \( \hat{x} \) is \( \frac{\sum_{i=1}^{n_k} \hat{x}_i}{k} \).

**Algorithm 3 Distance Voting based protocol**

1. Call Alg.1 to compute the tag density \( \rho \) and the optimal value of power \( p_0 \).
2. Get the identified categories \( C_i = \{x_{i,s}, x_{i,e}, l_i\} \) and the RSSI value \( s_{i,j} \) of each tag in the category \( C_i \).
3. Through the training data set \( T_d \), get the distance \( d_{i,j} \) of the tag to the center of the detection region.
4. Find the category \( C_0 \) which has the minimal \( d_{i,j} \).
5. for \( i \in \{1,2,..,k\} \), \( j \in \{1,2..n_i\} \) do
   6. if \( x_{i,s} > x_{0,s} \) then \( \hat{x}_i = \frac{x_{i,s} + x_{i,e}}{2} - \sum_{j=1}^{n_i} d_{i,j} \) endif
   7. if \( x_{i,e} < x_{0,s} \) then \( \hat{x}_i = \frac{x_{i,s} + x_{i,e}}{2} + \sum_{j=1}^{n_i} d_{i,j} \) endif
8. end for
9. compute the \( \hat{x} \) as follows: \( \hat{x} = \frac{\sum_{i=1}^{n_k} \hat{x}_i}{k} \).
10. OUTPUT: the estimation position \( \hat{x} \).

3) Analysis: This method improves the reliability of the baseline RSSI weighted average based method. By getting the corresponding distance value of the RSSI, we make the algorithm more robust.

**VII. PERFORMANCE EVALUATION**

This section describes the experimental settings and results. The experimental data in different conditions have been used to analyse the performances of the above algorithms. We use the overall execution time and the accuracy of localisation as the performance metrics.

**A. Experiments Setup**

As shown in Fig.8, we deploy the tags on the shelf. The width and the height of the deployment area equal to 5m and 2m, respectively. The tags are deployed in 5 rows and the tag category information is randomly generated. The tags used in the experiment are ALN-9662(HiggsGN) and the reader is ALR-9900+, which runs on global EPC Gen 2 platform. The antenna is Alien ALR Antenna. The antenna faces to the shelf, and the distance between the antenna and the shelf is 1m. In default situation we fix the reader towards the center of the shelf, and scan the tags for 50 query cycles repetitively.

**B. Comparison of execution time**

Based on Fig. 9, we can conclude that our proposed protocols (CMP, DVP) have better performances than the baseline protocols (CCP, RP). It is because that our proposed protocols scan the tags with the optimal power instead of the default power used by the baseline protocols. Usually, the larger power means the larger detected tag size, which leads to more execution time. In Fig.9, the tag density is 10tags/m and the average width of categories is 0.5m, the execution time of CMP is only 62% of that of CCP, while the execution time of DVP is only 57% of that of RP. In regard to the proposed protocols, DVP has larger execution time than CMP, because DVP needs to get the tag IDs and the RSSI value of each tags as well.

**C. Accuracy of Localisation**

From Fig.10, we can find that all our proposed protocols have higher accuracies than the baseline protocol. Besides, the proposed protocols are more reliable than the baseline protocols with lower localization error. This is because our proposed protocols use the optimal power to scan the tags almost remains the same. Therefore, the execution time of CMP is only 57% of that of CCP, while the localization error of DVP is only 73% of that of RP.

**D. Varying the tag density**

In Fig.11 and Fig.12, we respectively show the execution time and the localization error under different tag densities. In Fig.11, we can find that as the tag density increases, each protocol’s execution time increases. This is because the larger tag density \( \rho \) leads to the larger number of tags identified by the reader, which increases more execution time. However, when the tag density is large enough (eg. 8tags/m, 10tags/m), the execution time of CMP and DVP almost remain the same, because our proposed protocols use the optimal power to scan the tags. When the tag density increases, the effective interrogation region decrease, while the number of identified tags almost remains the same. Therefore, the execution time of CMP is only 58% of that of CCP, while the execution time of DVP is only 60% of that RP. Based on Fig.12, as the tag density increases, the localization error decreases. This is because the variance of the obtained tag information (tag size in each category, RSSI of the tags) decreases, when the tag density decreases.

**E. Varying the distribution of categories**

Fig.13 and Fig.14 show the execution time and the localization error under different category’s widths. As shown in
Fig. 13. Execution time under different category widths

Fig. 14. Localization error under different category widths

This category’s width has little effect on the execution time. Because the number of identified tags almost stay the same, as the tag density does not change in this situation. However, as the category’s width increases, the localization error increases. When the category’s width becomes smaller, the same area will cover more categories. More categories will improve the performance of the protocols, especially for CMP and DVP. CMP needs more categories for matching to select the k nearest categories, while DVP needs more categories for voting to select the k nearest categories.

VIII. CONCLUSIONS

In this paper, we propose an efficient localization scheme based on imprecise anchor nodes in RFID systems. We presented some observations from the realistic experiments and described the relationships among the read rate, detection region, identified tags size, power, RSSI, delay, etc. This paper presented four localization methods from different perspectives. To show the effectiveness of the proposed methods, we have evaluated the estimated position error by some experiments.

ACKNOWLEDGMENTS

This work is supported in part by National Natural Science Foundation of China under Grant No. 61100196, 61073028, 61321491, 91218302; JiangSu Natural Science Foundation under Grant No. BK2011559.

REFERENCES


