HMO: Ordering RFID Tags with Static Devices in Mobile Environments

Ge Wang, *IEEE Member*, Chen Qian, *IEEE Member*, Longfei Shangguan, Han Ding,*IEEE Member*, Jinsong Han, *IEEE Senior Member,* Kaiyan Cui, *IEEE Member*, Wei Xi*, *IEEE Member*, Jizhong Zhao, *IEEE Member*

Abstract—Passive Radio Frequency Identification (RFID) tags have been widely applied in many applications, such as logistics, retailing, and warehousing. In many situations, the order of objects is more important than their absolute locations. However, state-of-art ordering methods need a continuing movement of tags and readers, which limit the application domain and scalability. In this paper, we propose a 2-dimension ordering approach for passive tags that requires no device movement. Instead, our method utilizes signal changes caused by arbitrary movement of human beings around tags, who carry no device for horizontal dimension ordering. Hence our method is called Human Movement based Ordering (HMO). The basic idea of HMO is that when people pass between reader antenna and tags, the received signal strength will change. By observing the time-series RSS changes of tags, HMO can obtain the order of tags along with a specific horizontal direction. For vertical dimension, we employ a linear programming method that is tolerant of tiny errors in practice. We implement HMO with commodity off-the-shelf RFID devices. The experimental results show that HMO can achieve up to 88.71% and 90.86% average accuracies in the signal- and multi-person cases, respectively.

Index Terms—RFID, relative localization

1 Introduction

Passive Radio Frequency Identification (RFID) tags have been widely applied in many applications [1][2], such as logistics, retailing, and warehousing. Besides identification, object locations have become increasingly important in those applications. Hence a number of solutions have been proposed to localize passive tags [3][4][5][6]. However, most of them require expensive devices and infrastructure (*i.e.*, multiple antennas [6] or synthetic aperture radar(SAR) [5]), or continuous tag movement [3]. These constraints limit the applications of tag localization.

Very recently, researchers have identified the importance of tag ordering or relative localization of RFID tags [7][8][9] [1]. Instead of reporting the physical locations of each tag (called absolute localization), tag ordering or relative localization tells the relative locations of all tags. For example, in a one-dimensional space such as a moving conveyor belt, for any two tags a and b, relative localization must tell whether a is ahead of b or in reverse. In a two-dimensional space such as a shelf, relative localization must tell the relative location of an arbitrary pair of tags a and b, such as "a is below b vertically and to the left of b horizontally". Compared to absolute localization, relative localization is simple and significantly saves the infrastructure cost, when absolute locations are not necessary in the application. We identify the following user cases for relative localization.

- In a library, correct orders of the books or archival brochures are essential for readers to find their targets using the library call numbers. Managers can use relative localization to ensure that the books are placed in the right order.
- A large number of boxes (or baggages) are placed in an order to be loaded to a truck (or cargo aircraft). The truck will unload the boxes to multiple destinations hence we need ensure that the boxes that are unloaded earlier should be placed closer to the gate. In this application the absolute locations of the boxes are not important but the orders should be identified and checked.
- In a retail store, the manager needs to ensure that items in a shelf is correctly placed and consistent to the labels. As long as the manager knows the relative locations of the items are correct and items are tightly placed, s/he can conclude that the placement is good. It is because the shelf and items all have fixed size. If the orders are correct, miss-placement can hardly happen.

However, prior relative localization approaches still have two rigid requirements (or limitations) similar to absolute localization solutions: 1) the continuous movement of the reader or tags, and 2) avoiding the impact of environment noises (*e.g.*, human movements). For example, OTrack [7] is designed only for tracking the order of objects on a moving conveyor belt. STPP [9] recognizes the relative positions of tags by observing the sudden changes of phase values from different tags when the reader antenna moves along a known direction. These methods require a continuous relative movement of the reader or tags, which may be very time-consuming and labor-intensive. In addition, if there are people moving nearby or between the reader and tags, these movements can also incur sharp changes on the received

- *Ge Wang, Han Ding, Wei Xi, and Jizhong Zhao are with Department of Computer Science and Engineering, Xi'an Jiaotong University, China. Email: wangge@stu.xjtu.edu.cn, {dinghan, xiwei, zjz}@mail.xjtu.edu.cn*
- *Chen Qian is with Department of Computer Engineering, University of California Santa Cruz, USA. (Email: cqian12@ucsc.edu)*
- *Longfei Shangguan is with the Department of Computer Science, Princeton University, Princeton, USA. (Email: longfeis@cs.princeton.edu)*
- *Jinsong Han is in Zhejiang University, China. (Email: hanjinsong@gmail.com).*
- *Kaiyan Cui is with school of software, Xi'an Jiaotong University, China. (Email: kaiyan_cui@stu.xjtu.edu.cn)*
- *Wei Xi is the corresponding author.*

1. In previous work [7][8][9], researchers define relative localization as tag ordering in a direction. Hence we follow this definition and use tag ordering and relative localization interchangeably in the paper.

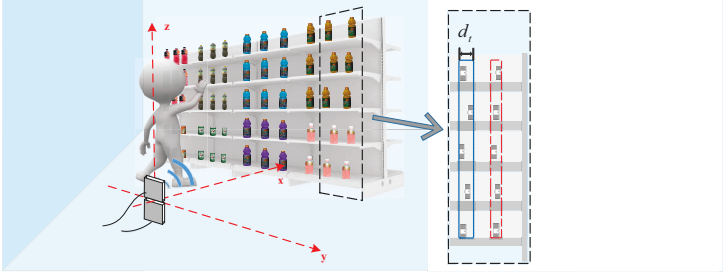


Fig. 1. Overview of HMO

data. As a result, the recognized relative locations are error-prone and environment-dependent.

Unfortunately, these two requirements are hardly to satisfy in practice. Except for some specific cargo systems that are with moving conveyor belts, objects as well as their tags always keep static in most time, such as the books or other items tightly placed in the shelves. Under these circumstances, moving either the reader or tags is impractical, considering the dedicated space for accommodating the moving devices or infrastructures. Holding portal RFID readers to scan all tags may avoid the interference caused by moving objects. But this solution should be manually performed and hence increases the infrastructure and labor cost. In addition, it cannot provide real-time localization.

In this paper, we propose a Human Movement based Relative Localization system, namely HMO, using passive RFID to achieve accurate relative localization. HMO leverages the arbitrary movement of human beings in the region, rather than devices, to explore the spatial information about the tags. With the help of human movement, HMO can acquire two-dimensional relative locations of passive tags on a shelf or rack as shown in Fig. 1. Nevertheless, we should address the following challenges in the design and implementation of a practical HMO system.

- **Arbitrary human movement**. Human movement in the region is uncontrollable. It introduces non-trivial uncertainty and increases the difficulty in analyzing signal changes of tags.
- **Multiple-person impact**. In real-world applications, it is quite common that multiple persons walk within the area of interests. When they move simultaneously, the line-of-sight propagation paths from the reader to several tags may be blocked, causing disturbance to the received signal. This disturbance might incur errors when using the signals for relative localization. Mitigating this impact, however, is challenging in existing RFID technology.
- **Untidy object position**. In practice, some objects are roughly vertical, but not aligned strictly on shelves, like the example shown in the black rectangle in Fig. 1. Our relative localization system should be resilient to such misalignments.
- **Multipath effects**. In indoor environments, multipath effects always play a considerable role in the wireless communication. It may change the profile of received signals, which introduce unpredictable errors in signal processing.
- **Irregularity of moving objects**. As aforementioned, HMO utilizes the mobility of arbitrary objects. However, those objects may not be regular in shape. For example, the obstacle ability of person's body and separate legs may not be the same.

The contribution of HMO is summarized as follows.

- We design HMO, a light-weight relative localization system to extract the two-dimensional sequences of a tag array based on the RSS values and phases of backscatter signals. Different from prior work, HMO utilizes human movements, instead of any device movement, to retrieve the horizontal-dimension ordering of tags. In addition, we propose a linear programming method, which is tolerant to practical noises, to find out the vertical-dimension order. The experiments show that our system provides more flexibility and scalability to modern inventory and logistics applications.
- HMO is resilient to the multiple-person movements. This merit enables HMO to operate and scale well in real deployments.
- HMO is also resilient to multipath effects and irregularity of moving objects, which makes HMO more practical in indoor environments.
- We implement a prototype of HMO using commodity off-the-shelf (COTS) devices and conduct extensive experiments to evaluate its performance. The results demonstrate that HMO is an efficient and accurate system for tag relative localization with static devices.

The rest of paper is organized as follows. We review related work in Section 2. The problem specification of this paper is presented in Section 3. In Section 4 and Section 6, we introduce our system design and analysis. The experiments and evaluation is illustrated in Section 7. Finally, we conclude this paper in Section 9.

2 Related Work

Using RFID for Localization is an attractive research direction. Prior work falls into two categories: absolute localization and relative localization. We list and compare some of related work in Tab. 1.

Absolute localization: Researchers found that the information contained in the RF signal can be leveraged to infer the locations of RFID tags. According to the signal information the authors used, we divide absolute localization methods into three classes, namely fingerprint-based, phase-based and signal-strength-based. Among fingerprint-based localization methods, LANDMARC [4] is one of the pioneering work for RFID tags. It utilizes the RSSI similarity between the target tag and reference tags for localization. However, fingerprint based method have its drawbacks. To setup the fingerprints for all possible locations, fingerprint based approaches [15][16]usually rely on densely pre-deploying tags in the area of interests and pre-collecting their RF signals. Meanwhile, there is a growing interest in using phase differences [10][17][18][19] [20][6][3] or Angel of Arrival (AOA) [21][22] [20][23][21] to estimate the absolute locations of tags. Liu *et al.* [6] utilize the phase differences received at different antennas to localize tags which are on the same horizontal level with those antennas. Their method has been widely adopted by many following solutions. Among these work, RF-IDraw can extract the trajectory of the target tag bonded on a finger [10]. Yang and *et al.*

TABLE 1
Some recent RFID localization methods (Some methods may include trajectory tracking)

Localization category	Method	Hardware requirements	Restrictions	Combat multipath
Absolute	PinIt [5] *SIGCOMM'13*	SDR	mobile antenna, anchor tags	✓
	RF-IDraw [10] *SIGCOMM'14*	2 readers, 8 antennas	limited space	✗
	Tagoram [3] *MobiCom'14*	2 antennas	N/A	✗
	BackPos [6] *Infocom'14*	4 antennas	2D	✗
	RFly [11] *SIGCOMM'17*	Drones	N/A	✓
	RFind [12] *MobiCom'17*	Newly-built and expensive reader	N/A	✓
	BNB [13] *MobiCom'16*	Entirely newly-built devices	N/A	✓
Relative	OTrack [7] *INFOCOM'13*	Conveyor	Moving objects/antenna, 1D	✗
	STPP [9] *NSDI'15*	1 antenna	Moving objects/antenna	✗
	MobiTagbot [14] *MobiSys'16*	Robot	Mobile antennas, robot, training	✓
	HMO (Our work)	2 antennas	few reference tags	✓

propose a real-time tag tracking system called Tagoram [3], which implements hologram-based method on tag tracking and localization. Tagoram has very high precision (*cm-* and *mm-* level). Recently, Ma and *et al.* propose a new localization system called RFly [11], which can scan the tags in a larger range with a drone. It is also common to utilize AoA information for localization [24][25][26][20][21][22]. Besides aforementioned works, there is also some work rely on the received signal [27][28][29]. For example, Wang *et al.* propose a method to find out the position of the target tag by comparing the received signal profile with adjacent reference tags [5]. This method can cope with both line-of-sight and none-line-of-sight cases. In addition, RFind [12] utilizes ultra-wideband RF signals to quary the tag and measure its flight time. It achieves centimeter-scale localization without reference tags. In particular, some researchers employ advanced electromagnetism and communication techniques, such as the synthetic aperture radar [30] [5] and multiple antennas [3][23], to achieve accurate localization. The basic idea of HMO to find out the inner relationship among the measurement of physical information from multiple channels of the reader and then infer the actual value of the clean phase/RSS. Recent year, multi-channel based signal measurement has been widely applied in RFID applications[14]. This method can estimate the dynamic vectors introduced by the movement of hands.

Relative localization: Relative localization is increasingly important for emerging applications of RFID systems. The methods proposed in [7] and [9] are two prominent ones. OTrack [7] leverages a phenomenon that when a tag is close to the reader, the reader can experience higher response reception ratio from the tag. The work proposed in [9] utilizes the spatial-temporal phase profiling to localize objects in two-dimensional space with a moving reader. By analyzing the phase profiles of tags, the system can obtain the spatial order of tags. The major limit of existing RFID based relative localization is the requirement of continuing movement of readers or tags to get necessary signal changes. MobiTagbot [14] is a recent method that can determine the order of a set of tags. The authors propose a hologram-based method and compare the phase profiles between the tranining and testing data. However, MobiTagbot needs extra hardware, *i.e.*, a moving robot, as well as a good deal of training data.

Different from previous work, our approach (HMO) makes use of signal changes caused by the movement of human beings who carry no device, which significantly enlarges the application domain and cost-efficiency of relative localization.

Beside localization, RFID system can also be employed in human-computer interaction and security scenarios. Most human-computer interaction work rely on detecting the signal changes that interferenced by the movement of people or a part of their body. Researchers have studied the relationship between the received phase or signal strength profile and the human activity [31][32][33][34][35][36][37][34]. Among them, Tadar [31] introduces a device-free method for tracking moving objects through a wall. Shopminer [34] and CBID [35] monitors the movement of tagged commodities and infer the customer behavior. RFIPad [38] is a human hand gesture detection system to recognize basic touch-screen operations and English letters. Beside RFID, many wireless techniques are used to improve the user experience [39][40][41], such as Wi-Fi [42], Bluetooth, LoRa *etc.*.

3 Problem Specification

In this section, we specify the problem the paper focuses on.

We assume our system is aiming to determine the relative locations of tags that are deployed in a library, warehouse, or supermarket. Each object is pasted by a passive RFID tag. These tags have different IDs (denoted as EPCs according to the EPC C1G2 Global Protocol [43]). All objects of interests are placed in sequence on layered shelves. The tags in a shelf *roughly* form a grid. In each row of this grid, *i.e.*, each layer of the shelf, the tags are horizontal. But in each column of the grid, the tags are roughly vertical to each other, but not strictly aligned. An RFID reader is used for interrogating the tags. Our goal is to identify the vertical and horizontal orders of these tags in the two-dimensional plane on the shelf.

To achieve this goal, we model the tag group as a two-dimensional array. For simplicity, we only focus on the foremost tags in the shelf, as shown in Fig. 1. We setup a three-dimensional coordinate system, with its origin at the center of two reader antennas. We focus on the two-dimensional plane that contains the tag array. In this plane, the tags with a same y coordinate or z coordinate are in a column or row, respectively. We assume there is one or more persons walking in front of the shelf arbitrary, like customers in a shopping mall or employees. By scanning the tags attached to the objects, we can obtain their IDs and signal information, including the Doppler shift, phase, Received Signal Strength (RSS), etc. To obtain the order that the tags are along the Y dimension, we try to investigate the relationship between human movement and the RF

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2018.2890520, IEEE Transactions on Mobile Computing

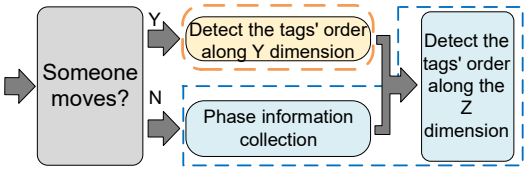


Fig. 2. System work flow

changes of tags. We utilize the hyperbolic positioning to sort tags along Z dimension. We will show above techniques in Section. 4. Note that this model is no simpler than existing relative localization methods [7][9].

4 System Design

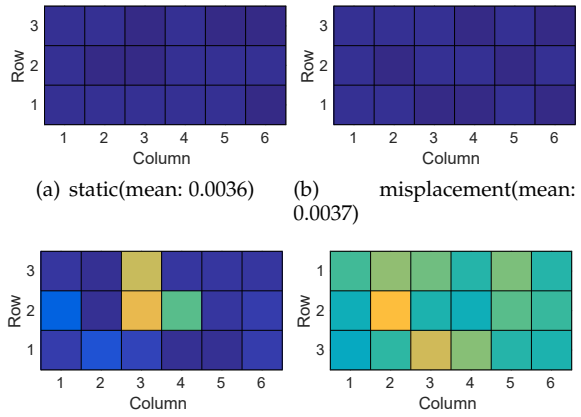
In this section, we elaborate the design of HMO in details. As aforementioned in Section 3, the tags attached to the objects on the shelf actually form a matrix. The matrix is on a two-dimensional space, as illustrated in Fig.1. Relative localization of the tags in this two-dimensional space can be addressed by identifying their order along the Y dimension and Z dimension, respectively.

4.1 System overview

Fig. 2 outlines the work flow of HMO. First, we determine whether there is one or more person moving inside the area of interests. If YES, we detect the tags' order along the Y dimension by analyzing the RSS changes caused by the moving person on these tags. In particular, we propose a new metric, namely 'influenced region', to describe such changes. Utilizing this metric, we can cluster the tags with a roughly similar z coordinate, *i.e.*, the tags are placed in a same column in the matrix. Note that if there is no person moving in the area, we can collect the phase information for localization at a later time. Second, we determine the vertical order of tags in the clustered columns. For each column, we detect the order of tags along the Z dimension. With some essential information (*e.g.*, the vertical distance between two adjacent layers), we can identify the relative locations among the tags in a same column. To do so, we utilize a hyperbolic method by measuring unperturbed phase differences received at two antennas. This step uses two kinds of information, the signal phase information of tags when there is no moving person and the clustering results along the Y dimension. For every time that a person moves in front of the shelf, HMO will report the tag's two-dimensional orders.

4.2 Human movement detection

Before tag ordering, we first determine whether the application area is static. We define the 'static' environment as there are no objects existing between the reader and tags. Our basic idea is to measure the extent of data variability of the tags in an application area. It is obvious that the tags' data collected in static environments is more stable than that collected from dynamic environments. However, it is not an easy task to decide whether the environment is static. The first challenge is the tag diversity. Due to the manufacture imperfection, different tags may have some tiny differences in the initial phases and signal strength, even if they are in a static environment. As a result, it is hard to propose a universal and ubiquitous standard to evaluate all the tags. Another challenge is that the position



(a) static(mean: 0.0036) (b) misplacement(mean: 0.0037)

(c) stand-still(mean: 0.0317) (d) movement(mean: 0.1077)

Fig. 3. The CV results in four cases.

of a tag may also change due to human interferences. For example, a customer misplaces a good on the wrong shelf. In this case, the data of the misplaced tag will not identical to prior ones. To deal with the aforementioned challenges, we employ a metric called *coefficient of variation* (CV), which can be expressed as follows:

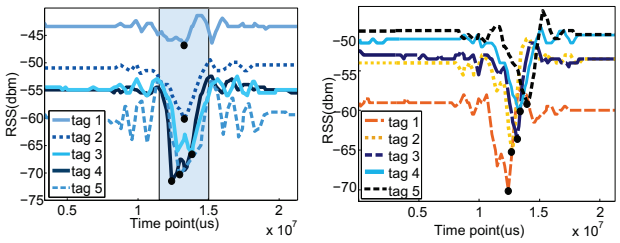
$$CV = \sigma/\mu \tag{1}$$

where σ is the standard deviation and μ is the mean. CV is good for measure the extent of variability in relation to the mean of the population. To exhibit the effectiveness of CV, we conduct an experiment with a 3×6 tag array. Our experiment includes four different cases, namely static, misplacement, people stand-still and people movement. The first two cases are conducted when there is no people stay inside the application area. The only difference is we relocate one tag to another position in the 'misplacement' case. And the last two cases simulate the different activities of a human being in practice. The results of the four cases can be found in Fig. 3. In Fig. 3, each cell represents the CV value of the tag at according position. We find that in static and misplacement cases, the mean CV values are less than 0.0037, while for people stand-still and movement cases, the mean CV value (0.0317 and 0.1077) is about 10 times to 30 times larger than that in static environments. So CV is appropriate for determining whether the application environment is static or not. We use an experimental threshold which equals to 0.01. The environment that the mean CV is less than the threshold will be considered as 'static'.

4.3 Order detection along the Y dimension

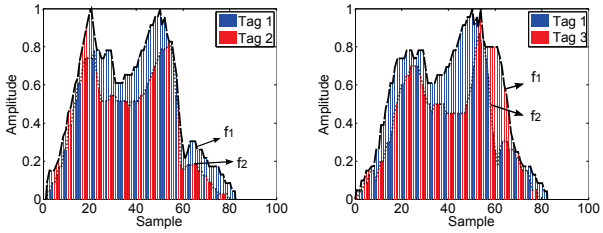
We try to first determine the order of tags along the Y dimension. We resolve the problem of detecting the order along the Y dimension into two problems, namely *'which tags are in a same column'* and *'which column these tags belong to'*. The solution for the former one will classify all tags into several groups. Solving the latter problem then deterministically maps each of these groups to a column in the tag array.

1) Which tags are (roughly) in a same column: We aim to leverage the RSS changes influenced by human movement in this problem. We define the signal segment under this influence as an *influenced region*. In order to thoroughly understand the influenced RSS, we conduct experiments to investigate the appearance of influenced regions in the RSS



(a) tags in a same column (b) tags in different columns

Fig. 4. Comparison of tags in a same and different columns



(a) $similarity = 0.9739$ (b) $similarity = 0.6878$

Fig. 5. Similarity of a pair of tags in a same and different columns

curve. We ask a volunteer to walk from one side to the other side of the shelf. The layout of the tag array on this shelf is a 5×5 tag grid. Each cell in the grid is 10 cm $\times$ 20 cm ($L \times H$). We depict the results of 5 tags in Fig. 4. We find that such a region is like a bathtub. At the start point of this region, there is a sudden decrease in RSS. The influence may take its effect for a short time period, and then end up with a sudden increase in the RSS curve. Intuitively, we have several parameters about the influenced region potentially used for the clustering, including the start/end time point, the lowest RSS point and the duration of the influenced region. Fig. 4(a) shows the RSSes of those tags when they are in a same column and Fig. 4(b) shows their RSSes when they are placed in different columns. From the result, we find that all the parameters mentioned above are distributed arbitrarily no matter what disposition of these tags is. So it is infeasible to utilize these aforementioned parameters for clustering the tags. Therefore, we have to pursue another methodology.

We propose a new metric, called Overlapped Area of Influenced Region (OAIR). It is also derived from our observation on a comprehensible phenomenon. We find that if a group of tags are roughly in a same column, they almost experience the influence from the moving person *at the same time*, considering the person's moving direction is along the Y dimension. On the contrary, if the tags are in different columns, their influenced regions start at different time points and experience different time durations. Meanwhile, the influenced regions are very likely overlapped for the tags with close distances. In particular, the tags in different columns, though their start (or end) time points may be very close to each other, their overlapped regions will be shrunken if increasing the interval distance between their columns. This observation inspires us to use the Overlapped Area of Influenced Region (OAIR) to represent the similarity between two tags' influenced regions, and hence to judge whether they are in a same column. For a fair comparison, we use the absolute values of signals and normalize them to $[0, 1]$. Specifically, we calculate the proportion of the overlapped area to the union of the two tags' influenced regions as the similarity, *i.e.*,

$$similarity = \frac{Overlapped\ areas}{Union\ of\ influenced\ regions} \qquad (2)$$

Clearly, this $similarity$ is suitable for comparing the overlapped influenced regions between two tags. We thereby define the OAIR of a pair of tags as their similarity.

To demonstrate the feasibility of OAIR, we select two tags from the array and observe their influenced regions. As shown in Fig. 5, f_1 (the black dash line) is the boundary of union influenced regions of tag 1 and 2, while f_2 (the gray dotted line) is the boundary of overlapped areas. The start and end points of the influenced region are determined by a simple but effective method. We utilize a moving time window and calculate the signal's variance inside it. Intuitively, the signals inside the influenced region have a much larger variance. The start and end points can be chosen as the time points at which the signal's variance has a sudden increase or decrease. Fig. 5(a) and 5(b) show the RSS overlapped region of tag pair in the same and different columns, respectivley. The OAIR of tags in the same column is much higher than the value of tags in the different columns. It is worth to note that though the start and end time points of the two tags are very close to each other, OAIR can still distinguish them in the latter case. In addition, OAIR utilizes the similarity between influenced regions, which is resilient to the vertical misalignments of the objects in different layers on the shelf. In the following, we define the *tolerant distance*, denoted as d_t, to quantify the distance between misaligned tags in different layers.

To cluster all tags in the array, we calculate the OAIR of each pair of tags and put the result in a matrix $A_{n\times n}$, where n equals to the total number of tags in the array. We still use the 5×5 tag array for illustration. For simplicity, we only choose 6 tags, whose layout is actually a sub-matrix with 3 rows and 2 columns. Specifically, tag 1, 3, 5 are in the first column and tag 2, 4, 6 are in the second column. The volunteer walks through the first column and second column sequentially. Tab. 2 demonstrates the OAIR results of the six selected tags in their $A_{6\times 6}$ array. For each tag in the matrix, our goal is to determine the top M tags that have the highest OAIR, which are most likely in a same column with this tag. In practice, the number of objects in one column can be known in advance. For instance, it is very common that this number is usually equal to the number of layers in the shelf. So we can set M as the number of layers in the shelf.

After choosing the top M tags, we obtain a new $n \times M$ table $T_{n\times M}$ for all n tags. Each element (i, t) in the table $T_{n\times M}$ represents the t-th tag in the tag i's top M list. For example, the element $(4, 2)$ in Fig. 6 is 6, which indicates that tag 6 has the second highest similarity with tag 4 in its top M list. Then we transfer table T into a digraph, in which each edge is correlated an element in T. For example, Fig. 6 can be mapped to a digraph, as plotted in Fig. 7.

From this digraph, we have two observations: *1)* any pair of tags in a same column have a pair of edges with each other; *2)* two tags in different columns only have a single edge or no edge at all. From the digraph Fig. 7, if we remove all single edges among the vertices, we can get

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2018.2890520, IEEE Transactions on Mobile Computing

TABLE 2
OAIR results between all tag pairs

No. / No.	1	2	3	4	5	6
1	1	0.9706	0.9888	0.8759	0.9951	0.9380
2	0.9706	1	0.9079	0.9553	0.9684	0.9918
3	0.9888	0.9079	1	0.7972	0.9875	0.9041
4	0.8759	0.9553	0.7972	1	0.8667	0.9909
5	0.9951	0.9684	0.9875	0.8667	1	0.9506
6	0.9380	0.9918	0.9041	0.9909	0.9506	1

top / index	1	2	3
1	1	5	3
2	2	6	1
3	3	1	5
4	4	6	2
5	5	1	3
6	6	2	4

Fig. 6. Table T

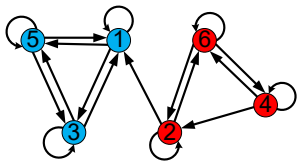


Fig. 7. Digraph for reflecting the tag's top M list

several independent subgraphs. One is formed by tag 1, 3, and 5, while another contains tag 2, 4, and 6.

To use the above idea, we first transform table T into an adjacent matrix B. The value of component b_{ij} in matrix B is defined as follows:

$$b_{ij} = \begin{cases} 1 & j \in \Phi_i \\ 0 & j \notin \Phi_i \end{cases} \quad \Phi_i = \{\forall j | \exists < i, j >\}, \tag{3}$$

where Φ_i is the set of all vertices in the digraph that have an direct edge $< i, j >$ between the pair of vertices, i and j. In other words, set Φ_i contains the top M candidate tags that are in a same column with i. For example, the adjacent matrix B generated from the digraph in Fig. 7 is shown in Eq.4.

$$B = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \quad R = \begin{bmatrix} 3 & 1 & 3 & 0 & 3 & 0 \\ 1 & 3 & 1 & 2 & 1 & 2 \\ 3 & 1 & 3 & 0 & 3 & 0 \\ 0 & 2 & 0 & 3 & 0 & 3 \\ 3 & 1 & 3 & 0 & 3 & 0 \\ 0 & 2 & 0 & 3 & 0 & 3 \end{bmatrix} \tag{4}$$

Based on matrix B, we remove the single edges and identify the column for each tag. We calculate $R = B \times B^T$. Here B^T is the transpose of B. Obviously, only when the arrows $< i, j >$ and $< j, i >$ both exist in the digraph, *i.e.*, the b_{ij} in matrix B and b'_{ij} in matrix B^T are all equal to 1, the product will be equal to 1. Otherwise, the product will be 0. In this way, we remove the single edge in the digraph. The matrix R corresponding to the graph in Fig. 7 is shown in Eq. 4.

Each component r_{pq} in matrix R reflects the number of common members in the Φ_p and Φ_q, *i.e.*, the number of common candidate tags in both the top M lists of tags p and q. That is:

$$r_{pq} = |\Phi_p \bigcap \Phi_q| \tag{5}$$

Here we use the value of r_{pq} as a weight for determining whether tags p and q are in a same column. Specifically, the larger the weight is, the higher probability that the tags p and q are in a same column. In practice, for any pair of tags p and q, the maximum number of their r_{pq} is limited to M. Therefore, we multiple matrix R by $1/M$ and compare the results with a threshold $thre$. This threshold can be empirically determined, as described in Section 7. In this way, we cluster the tags into different groups, and in each group the tags are in a same column.

Algorithm 1 Determining the order of columns

Initialization:
The influenced column sequence set $S = S_1 \rightarrow S_2 ... \rightarrow S_t ... \rightarrow S_T$, $(T \geq N)$ (T is the number of influenced columns, N is the number of columns);
The leftmost and rightmost reference tags ref_l and ref_r.
Output:
The columns' order set from left to right is $O = O_1, O_2, ...O_N$;

$p = 1, q = 2$;
while $q \leq T$ **do**
  $weight(S_p, S_q) ++, weight(S_q, S_p) ++$;
  $p ++, q ++$;
end while
for each $t \leq T$ **do**
  $F_1(S_t) = \{S_t^1 | weight(S_t, S_t^1) \geq weight(S_t, S)\}$;
  $F_2(S_t) = \{S_t^2 | weight(S_t, S_t^2) \geq weight(S_t, (S - S_t^1))\}$;
end for
for each $n \leq N$ **do**
  if n=1 **then**
    $O_1 = \{S_t | S_t \ni ref_l\}$
    O_2=$F_1(O_1)$.
  end if
  if n=N **then**
    $O_N = \{S_t | S_t \ni ref_r\}$
    O_{N-1}=$F_1(O_N)$.
  end if
  if $3 \leq n < N - 2$ **then**
    if $F_1(O_{n-1}) = O_{n-2}$ **then**
      $O_n = F_2(O_{n-1})$
    else
      $O_n = F_1(O_{n-1})$
    end if
  end if
end for
return O;

2) To which column these tags belong: After clustering the tags, we need to determine the relative order of these clusters, *i.e.*, to which column these clusters belong. A straightforward way is to place a number of reference tags in a row on the shelves. The position of each reference tag is known in advance for localizing any tags nearby. However, such a scheme is impractical, since for any tag in a given cluster the system has to search all reference tags to determine which one is nearest the tag.

Instead of using this treatment, we paste two reference tags aligned to the rightmost and leftmost columns of a shelve. It is obvious that in the shelf each column is adjacent to no more than two columns. Note that the rightmost and leftmost columns are only adjacent to one column. According to this observation, we propose Algorithm 1. In this algorithm, we first assign a serial number S_t to each column according to the time sequence that it is influenced. For each column S_t, we find out two columns S_t^1 and S_t^2, which are with the top two weights in weight function $weight(S_t, S)$.

Obviously, S_t^1 and S_t^2 are the two adjacent columns of S_t. After identifying these two columns, the next task is to determine their positions relative to S_t, *i.e.*, on the left-hand or right-hand side of column S_t. We first find out the rightmost and leftmost columns O_1 and O_N by comparing the influenced area of all columns with reference tags ref_l and ref_r. Intuitively, O_1 and O_N are only adjacent to one column *i.e.*, $O_2 = F_1(O_1)$ and $O_{N-1} = F_1(O_N)$. Combining the relative positions between each column and its adjacent columns, we can determine the relative sequence one by one either from left to right or vice versa. For example, since we have known the index of O_1, we can determine O_2. Except O_1, O_3 has the maximum weight in $weight(O_2, S)$ among all other columns, *i.e.*, O_3=$F_1(O_2)$ or $F_1(O_2)$. In this way, we can determine the relative position of all the columns and obtain. the columns' order O. Note that the human movement direction does not impact this algorithm. No matter which trajectory the person chooses, we can extract the ordering from the received signals as well.

4.4 Detect tags order along Z dimension

Compare the Y dimension's ordering, vertically determining the relative locations for tags in a same column is even more difficult.

Due to the multi-path effect, tags with different heights might not show distinct differences on their received signal strength(RSS). In addition, the hardware diversity among tags is non-trivial. As a result, the RSSes of a given tag cannot be deterministically mapped to the distance from the tag to the reader.

We also try some prior solutions, namely Phase Change Rate (PCR) and Received Response Rate (RRR), which have been employed by existing tag ordering work [9] and [7]. However, neither of the method can accurately extract the tags' order along vertical dimension. Therefore, we have to pursue a method that does not rely on aforementioned characteristics for relative localization. Our basic idea is constructing a number of spatial hyperbolas by measuring the phase differences received at two near antennas. There should be some intersections of those hyperbolas and the two-dimensional plane that contains the tag array. These interactions are potential locations of the tags. In this way, we are able to determine the tags' order by identifying the most possible intersections.

We illustrate the above idea in Fig. 8. We employ two reader antennas. One is on the ground. Another is deployed vertically above it. Both of them face towards the tag array. We establish a 3D coordinate system, in which the origin is at the center of demarcation line of two antennas. The x-axis is perpendicular to the shelves and z-axis is vertical to the ground, as shown in Fig. 8. Note that the y-axis is paralleled to the shelf, which is not shown in Fig. 8. We use A_1, A_2, and T to denote the centers of antenna 1, 2 and tag, respectively. The distance between A_1 and A_2 is denoted as d. We adopt two mainstream directive antennas whose size is 25 cm$\times$ 25 cm. So $d = 25$ cm. The distance between A_1 (or A_2) and T is shown in Eq. 6.

$$\begin{aligned} d(A_1,T) &= \frac{\frac{1}{2}\cdot(2k_1\pi+\theta_1+\theta_{t_i})}{2\pi}\cdot\lambda \\ d(A_2,T) &= \frac{\frac{1}{2}\cdot(2k_2\pi+\theta_2+\theta_{t_i})}{2\pi}\cdot\lambda, \end{aligned} \tag{6}$$

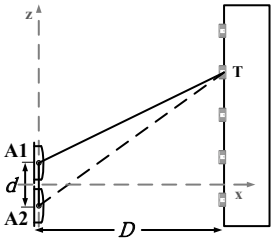


Fig. 8. System deployment

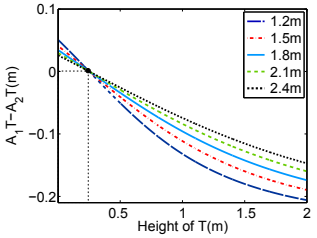


Fig. 9. $d(A_1,T) - d(A_2,T)$

where θ_1 and θ_2 are the received phase values of Tag_i at antenna 1 and 2, respectively. θ_{t_i} is the initial rotation of Tag_i. Note that there is a "round-trip" for the RF signals received at the antenna. Hence we multiply the length path by $\frac{1}{2}$ in Eq. 6. Since the phase is periodic, k_1 and k_2 are two nonnegative integers whose values can be $0, 1, 2, ..., K, (K \to \infty)$.

Locating a given tag requires to determine the value of k_1 and k_2, which implies that we should alleviate the influence of individual differences of tags. We then subtract $d(A_2,T)$ from $d(A_1,T)$ in Eq.6 and find that the initial phase θ_{t_i} can be eliminated. According to the triangle inequality, *i.e.*, the difference (absolute value) between any two sides of a triangle should be less than the length of the third edge. Thus, the absolute value of the difference between $d(A_1,T)$ and $d(A_2,T)$ should be less than the distance between A_1 and A_2. Hence we have:

$$\begin{aligned} &\mid d(A_1,T) - d(A_2,T) \mid < d(A_1,A_2) \\ &\frac{\lambda}{4\pi} \mid 2k\pi + \triangle\theta \mid < d(A_1,A_2), \end{aligned} \tag{7}$$

where $k = k_1 - k_2$ and $\triangle\theta = \theta_1 - \theta_2$. Note that each possible k corresponds to a hyperbola, *i.e.*, a possible solution of tag's position T. In practice, there should be only one k corresponding to the real location of tag T. However, k has 5 possible candidates according to Eq. 7, including $k = 0, \pm1, \pm2$ (here the detail is omitted because of the space limit). We need to identify this feasible k.

Fortunately, in practice the size of shelves facilitates identifying this k. We assume that the distance between the antennas and shelf is D and the height of shelf is h. These two parameters are usually constrained in real implementations. For example, the shelf used in shopping malls or libraries are usually designed for people to take items conveniently. Thus, the shelf height would not exceed the average human height too much. Similarly, the interval distance between two adjacent shelves should be sufficiently long for enabling multiple costumers' passing. Under this circumstance, the value ranges of D and h can be determined via such prior-knowledge, which shrinks the value range of k.

To illustrate how HMO discards the infeasible ks, we take an example of $D = 1.8$m and $h = 2$m, which are two general settings in real-world. For simplicity, we assume tag T's $y = 0$. For the tags in the column, their coordinates range from (1.8,0,-0.25) to (1.8,0,1.75). Therefore, we are aiming to determine the z coordinate of tag T within this range. We can obtain the value of $d(A_1,T) - d(A_2,T)$ according to the Pythagorean theorem. Fig. 9 exhibits the value range of $d(A_1,T) - d(A_2,T)$ for $D = 1.2m \sim 2.4m$. We find that the value of $d(A_1,T) - d(A_2,T)$ is inversely proportion to the height of tag T, *i.e.*, the vertical distance of the tag. In

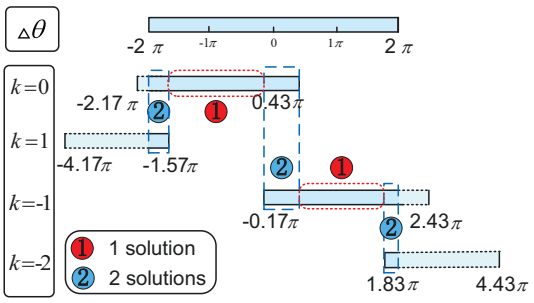


Fig. 10. Possible solutions of k

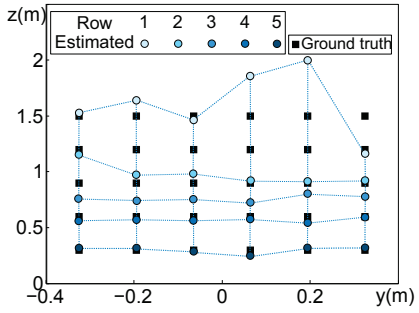


Fig. 11. 2D relative localization for the tag array

addition, the larger the D is, the smaller the possible range of $d(A_1,T)-d(A_2,T)$ is. When the tag is at the center of the two antenna (*i.e.*, $z=0.25m$), the difference is 0. We represent the minimum and maximum values of $d(A_1,T)-d(A_2,T)$ as min and max, respectively. Hence we have:

$$\frac{4\pi}{\lambda}\cdot min \le (2k\pi+\triangle\theta) \le \frac{4\pi}{\lambda}\cdot max, \tag{8}$$

In above example, the range of $d(A_1,T)-d(A_2,T)$ is $[-0.1742, 0.0343]$. To obtain the possible value of k, we exhibit the possible range of $\triangle\theta$ and its corresponding values of k in Fig. 10. Because both θ_1 and θ_2 are received phases of tag T, their difference $\triangle\theta$ is in the range of $(-2\pi, 2\pi)$. We find that for some observations of $\triangle\theta$, only one k conforms all the conditions above. We mark it with a circle 1 in Fig. 10. And we mark the phase difference range that holds two possible values of k with circle 2. As shown in Fig. 10, when the value of $\Delta\theta$ is within $(-1.57\pi, -0.17\pi)\bigcup(0.43\pi, 1.83\pi)$, the number of k is only 1. On the other hand, when the value of $\Delta\theta$ is out of this range, the number of possible k increases to 2. In this way, we can shrink the number of k to 2. Given each k, we have a corresponding height of tag T.

Since there may be two possible intersections, it is still unclear what order the tags are in a same column, because each possible k may introduce a solution of the order. Next, we apply Linear Programming to ultimately find out the unique solution fork. Adopting LP in HMO is based on two constrains: a) some tags have only one possible solution. We named these tags as *anchor tags*. Other tags that are more than one possible solutions are defined as *undetermined tags*. Obviously, each anchor tag has a deterministic z coordinate; b) The vertical distance between two adjacent layers in the shelf is known in advance. Suppose this distance is ℓ. In this example, we assume that ℓ is identical in the shelf, *i.e.*, $z_{upper}-z_{lower}=l$ (where z_{upper} and z_{lower} are the z ordinates of two adjacent layers). Note that in practice l is adjustable. Combined with these two constrains, we can determine the relative order for tags by minimizing the residual errors as follows:

$$min_{h(\hat{i})}\{\sum \parallel h(\hat{i}) - m_i\cdot\ell \parallel^2 + \sum \parallel h(j) - m_j\cdot\ell \parallel^2\} \tag{9}$$

where $h(j)$ represents the z ordinate of anchor tag j and $h(i)$ represents the z ordinate of undetermined tag i. Here m_i and m_j are the relative order of i and j, respectively. For each undetermined tag i, there is an intersection of hyperbola and shelves given a possible value of k. However, the height of this intersection $h(\hat{i})$ may yield different tags' order m_i. The intuition behind Eq. 9 is that we pursue all possible ks for all undetermined tags and find out the one with the minimum residual error, which is correlated to the correct order.

For example, if an undetermined tag has two possible intersections, *i.e.*, $h(\hat{i})=0.1$ and $h(\hat{i})=1.1$. Given that the heights of the anchor tags are $h(j_1)=0.6$ and $h(j_2)=0.3$, there are two estimated results for the order m_i of the undetermined tag. Since the $h(\hat{i})=0.1$, $m_i=1$, indicating that the tag is on the lowest layer. While for $h(\hat{i})=1.1$, $m_i=3$, suggesting that the tag is on the highest layer. Accordingly, the optional orders of the anchor tags m_{j_1} and m_{j_2} are $2,3$ and $1,2$. If $\ell=0.3m$, the residual errors are $[(0.1-0.3)^2]+[(0.3-0.6)^2+(0.6-0.9)^2]=0.22$ when $h(\hat{i})=0.1$ and $[(1.1-0.9)^2]+[(0.3-0.3)^2+(0.6-0.6)^2]=0.04$ when $h(\hat{i})=1.1$. Obviously, the latter is the correct order.

The Linear Programming method can tolerant some errors in practice. For example, the tags in two adjacent layers are not aligned, or their distance is not equal to the layer height l, *etc.*. That is because the errors introduced by the estimation of l (maybe several centimeters) are much less than the distance between two tags in adjacent layers (maybe several decimeters). In other words, the tags of right order still have the minimum residual error in Eq. 9. So Linear Programming is robust in practice.

The effectiveness of above solution for the vertical relative localization is demonstrated in Fig. 11. In this figure, the black squares are the actual positions of tags and the blue circles are the estimated results using our hyperbola based method. We can see that our method is reliable in retrieving the tags' order along the Z dimension.

4.5 Put all together

In this section, we briefly summarize our solution when it comes to 1D and 2D, respectively. For 1D ordering, we can either employ Y-dimension ordering or Z-dimension ordering according to the application cases. While for 2-dimensional ordering, we first detect whether the environment is static or not. If there is somebody moving inside the area, we analyze the horizontal ordering of tags by exploring the RSS variations. If the environment is static, we will update the phase information. However, tags are in a 3D space in the real world. Due to the non-isotropic emission characteristic [44] of RFID tags, we also consider the tags' orientations in practice. To guarantee that the reader can obtain the item's information accurately, we can paste multiple tags on different sides of the same item. In this way, no matter which side of item faces to the antenna, we can successfully order the items.

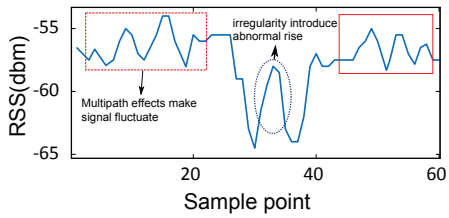


Fig. 12. The impacts on received signals introduced by practical factors

5 Cope with Practical Influence Factors

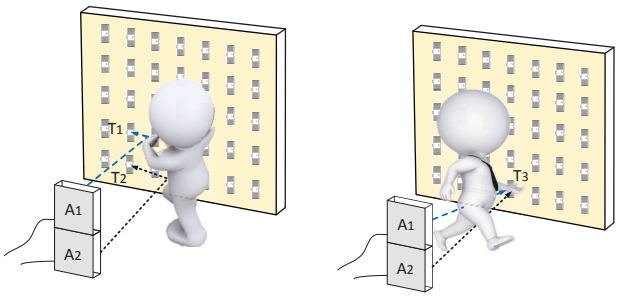
In Section. 4, we specify HMO's design. However, in practice, many factors may introduce unpredictable errors. Among these factors, multipath and human body's irregularity may be two crucial ones. In this section, we try to propose a method to reduce or even eliminate the influence introduced by these two factors.

In practice, the impact of multipath effect on the received signal is not ignorable, and it will induce errors in OAIR calculation. As shown in Fig. 13(a), with the continuously moving of a person, the received signal exhibits an amplitude-increasing(or decreasing) fluctuation. To verify that, we ask a volunteer to walk through a tag and illustrates the received signal in Fig. 12. Obviously, the signal fluctuates when the person goes approach and departs the tag. The fluctuations, which may increase the overlapped area with adjacent tags' signals, may eventually increase the false possibility. Our basic idea to cope with this problem is simple but effective, *i.e.*, the blockage of any objects on a tag is always the same for the two antennas. On the contrary, the multipath effects are not. As shown in Fig. 13(a), as a person walks approach tag T_1 and tag T_2, the multipath effects of T_1 and T_2 are not consistent for the two antennas A_1 and A_2.

Besides multipath effects, the irregular contour of human body may also affect the signals' profile, and introduce errors in overlapped area calculation. For example, as shown in Fig. 13(b), a man blocks tags that on tag T_3's column. However, due to the walking posture, tag T_3 may have a line-of-sight propagation path between person's two legs, which induce a considerable rise inside the influenced area (marked in blue circle). Fortunately, the swing of two legs do not affect the propagation path between T_3 and antenna 1.

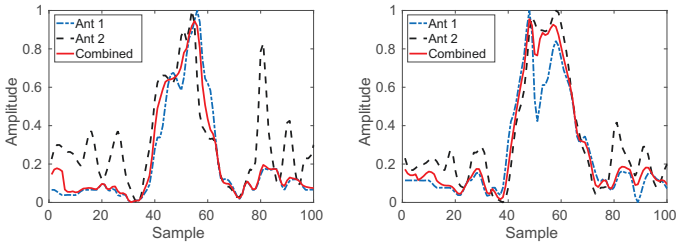
However, we do not know when will these aforementioned factors happen. And we also have no priori knowledge about which antenna would be a better signal collector at each time point. We have to find out a general signal processing method to solve both of these two problems. So we propose a Weighted-based Signal Combination (WSC) method and specify it as follows.

In our system design, we has two antennas, namely A_1 and A_2. That means for each tag in the matrix, we have two observations. One is collected by antenna 1 and another is observed by antenna 2. Due to the differences between the two antennas, such as location, working time, *etc.*, the two observations would be much the same but not identical. Our basic idea is that an obstacle is always making a fall in the received signal strength, while other factors, such as mutipath effects and irregularity of human body, can hardly introduce a similar change for both the two antennas. So



(a) Multipath effects are not identical for the two antennas (b) The human body irregularity may introduce a considerable rise on the received signal

Fig. 13. The impacts of practical factors



(a) cope with multipath effects (b) cope with body irregularity

Fig. 14. The performance of WSC

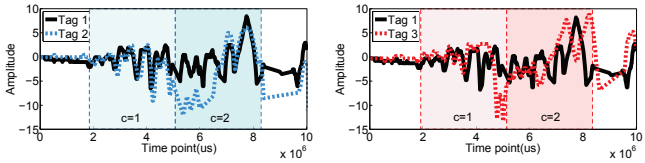
we consider both the two observations and combining them together to get rid of some undesired data. We first enrich the two observations by interpolation and make them have a same length. Then we try to assess the accuracy of the two observations for each time point. To evaluate the data quality at time point p, we employ a weight function ρ, which is determined by the data variance inside a time window δ, *i.e.*:

$$\rho_n^p = 1/var(S_n(p - \frac{\delta}{2} + 1 : p + \frac{\delta}{2})), \tag{10}$$

where ρ_n^p is the weight for the p-th sample of antenna n. And S_n is the observation for antenna n. Intuitively, if a person block a tag, the two weights, namely ρ_1^p and ρ_2^p, are both increase obviously. However, if an antenna observes abnormal signals, its variation is much larger than another one. Keep this in mind, we combine the two observations by the following formula:

$$S_c^p = \frac{S_1^p \cdot \rho_1^p + S_2^p \cdot \rho_2^p}{\rho_1^p + \rho_2^p}, \tag{11}$$

With aforementioned equation, only if a common changes for these two observations will be considered in the combined observation S_c. We exhibit the performance of WSC algorithm in Fig. 14. In Fig. 14(a), the data received at antenna 2 (drawn by black dashed line) is seriously polluted by multipath effects and induce many small waves in the signal. And in Fig. 14(b), the data collected by antenna 1 (marked in blue dotted line) appears an abnormal fall inside the influenced area. We find that WSC can effectively cope with the abnormal changes. The combined observation (plotted in red full line) gets rid of these error-prone samples and gives a more accurate input for OAIR algorithm. We will evaluate the performance of WSC algorithm with experiments in Sec. 7.



(a) Signals of tags in a same column (b) Signals of tags in 2 columns

Fig. 15. Signals in the multi-person case

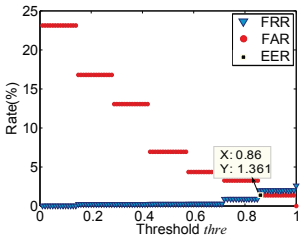


Fig. 16. Threshold selection

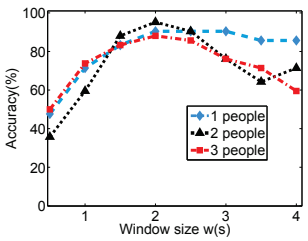


Fig. 17. Window size w selection

6 Scenario of Multiple Moving Persons

It is very common that multiple persons are moving in the deployment area of HMO at a same time. Since HMO does not rely on the human movement to determine the Z dimension order, we only need to detect the Y dimension order under the scenario of multi-persons.

Different from the case of single-person, multiple moving persons may influence multiple columns of tags simultaneously, raising the difficulty of relative localization along the Y dimension. Under this circumstance, the influence from multiple moving persons is much more arbitrary in both the spatial and temporal dimensions. In addition, we have no idea about the real positions of those moving persons, not to mention their influences to the tags.

Fortunately, our OAIR method can still function well in the multi-person case. The basic idea of OAIR is that the tags in a same column will always be influenced at the same time, no matter how many persons move horizontally. While the tags in different columns do not follow this principle. So we can just compare the signals between each pair of tags within a same time duration w, which denotes the length of a time window c ($1 \leq c \leq C$). n other words, the variable w is the length of window c, *i.e.*, $w = c_e - c_s$, where c_e and c_s are the end time point and start time point of the c-th window. If the two tags are in a same column, their OAIR will always be much higher than those pairs of tags that are in two different columns. To exam the effectiveness of OAIR, we observe the RSSes of 3 tags in our tag array when 2 volunteers walk simultaneously. Tags 1 and 2 are in a same column while tags 1 and 3 are in two adjacent columns. The horizontal distance between tags 1 and 3 are 10 cm. We present the variation tendency in Fig. 15. In the first time window $c = 1$, the RSS changes of tags 1, 2, 3 are indistinguishable. However, in the time window $c = 2$, tag 1 and 3 show significant difference in their RSSes.

To realize above comparison, we try to combine estimation results in several time windows. In each window c, we redefine the OAIR similarity as follows:

$$similarity = \frac{\int_{c_s}^{c_e} f_2 dx}{\int_{c_s}^{c_e} f_1 dx}, \tag{12}$$

where f_1 and f_2 are the union and overlapped area of the influenced regions of tag pairs, respectively, and both functions of time. We record the weight matrix R_c with newly generated OAIRs. At the end of this time period, we multiply these R_cs successively as following.

$$R' = \frac{\prod_{c=1}^{C} R_c}{M^C}, \tag{13}$$

where R' is a weight matrix by accumulatively combining multiple estimation results. After such multiplication, the correct result in R_cs will be amplified and strengthened, while the incorrect ones may be suppressed. Note that we can also utilize Algorithm 1 to infer the order of columns under the multi-person situation.

We are further motivated to combine multiple estimation results in the single person case for improving the accuracy. As aforementioned, every time a person passing the shelves offers an opportunity for a relative location estimation. To obtain a more accurate classification result, we try to combine multiple estimations derived from multiple passings of single-person. We screen out the signal segments that influenced by multiple single-person-passings, and divide these signal segments into several time windows. The estimated matrix of the c^{th} time window is denoted as R'_c. Similar to the multi-person case, we multiply these R'_cs like Eq. 13. We evaluate the effectiveness of this combination in Section 7.

7 Experiment and Evaluation

7.1 Implementation

Hardware: We implement an HMO prototype, which consists of a COTS UHF RFID reader model Impinj R420, two directional antennas in model IPJ-A0311, and a set of passive tags. The tags include five different types: Alien-964X, Impinj E41B, E41C, H47, and AZ-E53. Note those 5 types of tags are in different sizes and shapes.

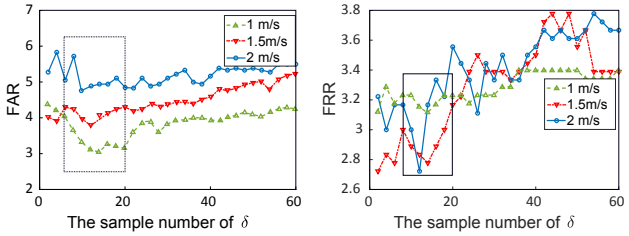
System setup: To be consistent with the real implementation, the horizontal distance between the reader antennas and the shelves is 1.2 m $\sim$ 2.4 m. We utilize a tag array that is with 7 rows and 6 columns, as well as 2 reference tags. In our experiments, we invite 5 volunteers, varying from 1.6 m to 1.8 m in their height and from 46 kg to 77 kg in their weight. The average walking speed of these volunteers is about 1.5 m/s . To account for the hardware diversity, we mix all types of passive tags together in the tag array. The height of each layer ranges from 20 cm to 40 cm. While the distance between two adjacent tags in a same row varies from 5 cm to 40 cm.

Metrics: We mainly use the accuracy to quantify the HMO's performance. The metric *Accuracy Ratio* is defined as follows:

$$Accuracy = \frac{\#\ of\ tags\ ordered\ correctly}{\#\ of\ tags\ in\ total}. \tag{14}$$

We define that a tag is ordered incorrectly in a sequence of tags if its detected order is not equal to its actual location. For example, if a tag sequence 'ABCDE' is wrongly recognized as 'ABCED', the accuracy will be 60%. We also utilize False Reject Rate (FRR) and False Accept Rate (FAR) to further evaluate the accuracy. Note we measure the accuracy based on only one report. In actual applications, the system can output the locations for every person walking through and the accuracy can be improved by considering multiple times of localization.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2018.2890520, IEEE Transactions on Mobile Computing



(a) False accept rate (b) False reject rate

Fig. 18. Window size δ selection

7.2 Parameter Selection

We try to determine two important parameters for HMO in real-world applications, *i.e.*, the threshold $thre$ of weight matrix R and time window's size w.

Threshold $thre$: As aforementioned in Section. 4, the threshold $thre$ plays an important role in judging whether two tags p and q are in a same column. In general, the lower the threshold $thre$ is, the more efficient the top M tags will be identified. However, on the other hand, the lower the $thre$ is, the more inaccurate the system will be. It is necessary to balance two factors while determining a proper $thre$. We utilize FAR and FRR to represent the performance of these two factors. In Fig. 16, we depict the Equal Error Rate (EER) of HMO. EER is a common metric to evaluate the accuracy performance for recognition systems. If we choose the threshold at the EER point, the sum of FAR and FRR reaches its minimum, and these two metrics are optimally balanced. With this EER point, we can properly determine the threshold. The default $thre$ is set as 0.86 in our prototype HMO. In practice, $thre$ can be experimentally adjusted according to users' demands.

Window size δ in WSC algorithm: As described in Sec. 5, we define a weight function ρ, which is determined by the signal variance inside a time window δ. To achieve a good performance, we experimentally choose the value of δ. Intuitively, the signals' variance are highly related to the moving speed of objects. As shwon in Fig. 18, we exhibit both the false Accept Rate (FAR) and False Reject Rate (FRR) of y-axis ordering when varying the sample number of δ from 2 to 60. The results show that when the window is of a size of about 10 to 20, both the FAR and FRR are extremely low, namely $3\% \sim 5\%$ and $2.7\% \sim 3.2\%$, respectively. In our experiments, we set the window size as 10.

Window size w in multi-person case: As aforementioned in Section. 6, we perform a matrix multiplication on the R_cs at the time window c successively. In general, a proper window size w results in high accuracy. As shown in Fig. 17, we exhibit the accuracy of HMO with $1 \sim 3$ persons moving inside the area, respectively. The x axis represents the window size w, which varies from 0.5s to 4s. We find that when $w = 2$s, The accuracy of HMO is high, *i.e.*, about 90%, for all these three cases. With this window size, HMO is also accurate in the single-person case. So we employ $w = 2$s as the setting of time window size.

7.3 Evaluation

After determining critical system parameters, we evaluate the performance of HMO in both the single-person and multi-persons cases. In addition, we discuss the effect of combining multiple estimation results.

Our experiment scene is like Fig. 19. We depict the HMO's accuracy when performing relative localization along the Y dimension in Fig. 20. The x axis is the window index. For the single-person case, we exhibit the accuracy before and after combining multiple estimation results. And for the multi-person case, we only focus on the result after the multiple measurement. We combine the measurement results in five time windows, and each time window lasts for 2 seconds, *i.e.*, w=2s. The result shows that combining multiple measurement results helps HMO to maintain a high accuracy (about $90.24\% \sim 93.11\%$). In addition, in the multi-person case HMO shows a slightly higher accuracy (about $90.48\% \sim 95.24\%$) than in the single-person case.

We also show the results that perform relative localizations along the Y dimension, Z dimension, and in the entire two-dimension space in Fig. 21. We find that the relative localization accuracy along the Y and Z dimension is 91.24% and 90.48%, respectively. These results are higher than that (90%) in single-person case. And for the relative localization in the entire 2D space, the accuracy is 88.71%, 90.48%, 91.24% when there is/are 1, 2, and 3 persons, respectively. The performance in multi-person case is a little better than that of single-person case. That is because multiple person may provide more information in the received signals, which is helpful for HMO to correct miscalculations.

We also evaluate the performance of the WSC algorithm. We ask 3 volunteers to walk through the tag array (7 rows $\times$ 6 columns) and exhibit the matrix R after OAIR algorithm. The larger the value of cell $< i, j >$ in matrix R, the larger probability that tag i is in the same column with tag j. As shown in Fig. 23(a), we find that without performing WSC, most tags may have a high probability with tags in adjacent columns. However, after performing WSC, most tags can be accurately classified (exhibit in Fig .23(b)). After extending the experiment time duration from 30 seconds to 2 minutes, we find that the classification accuracy increase obviously (as shown in Fig. 23(c)). We also discuss the improvement of WSC algorithm on original OAIR. As shown in Fig. 22, we compare the accuracy of employing OAIR only with both OAIR and WSC. By varying the experiment time, the accuracy of them are roughly increasing. In addition, applying WSC on the original OAIR can significant improve the accuracy. So WSC is helpful to improve the accuracy of y-axis ordering and reduce the impact introduced by multipath and human body's irregularity.

In a short summary, HMO can achieve accurate two-dimensional relative localization in practice. Besides, combining multiple estimation results indeed improves the accuracy of HMO.

7.4 The performance of HMO under different settings

We then investigate the performance of HMO by tuning the prototype system's settings, including the layer height, interval distance between adjacent tags, distance between the shelf and reader, and the human moving speed. We show the result of using HMO under these tunings along the Y dimension, along the Z dimension, and in the entire 2D plane. We also compare HMO with the approach utilizing RSS directly and a state-of-the-art relative localization approach, STPP [9].

Fig. 19. Experiment scene

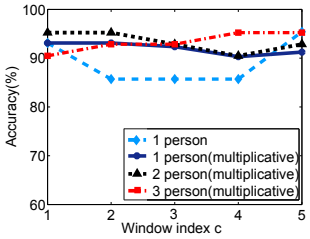


Fig. 20. Accuracy of OAIR

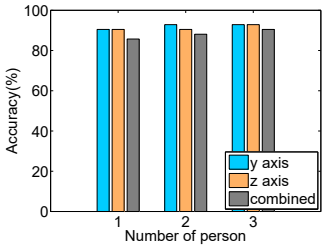


Fig. 21. Localization accuracy when c=5 and w=2s

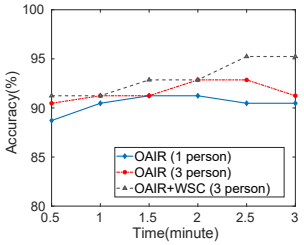
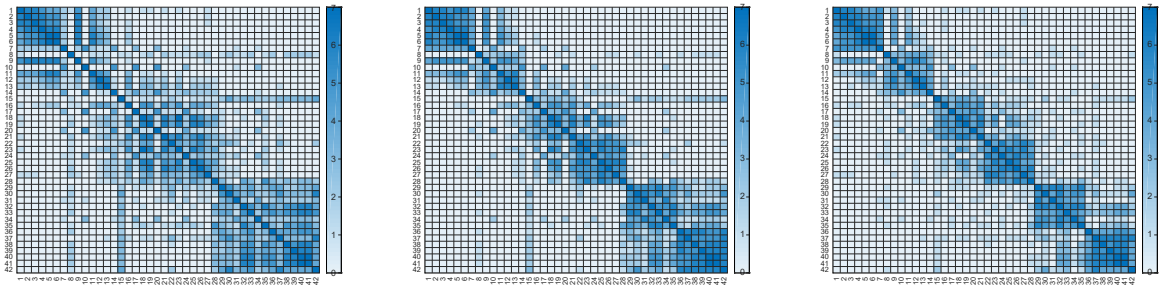


Fig. 22. The performance of WSC



(a) Without WSC, time: 1 minutes (b) With WSC, time: 30 seconds (c) With WSC, time: 2 minutes

Fig. 23. The performance after operating WSC alogorithm

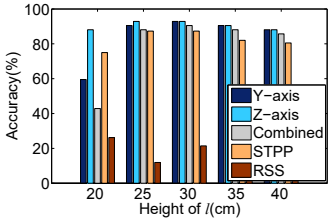


Fig. 24. Layer height

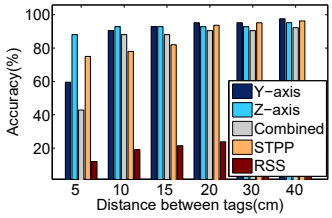


Fig. 25. Interval distance

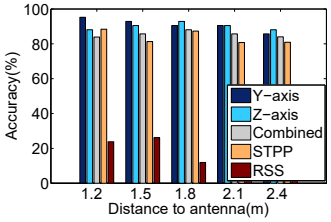


Fig. 26. The distance from the reader to shelves

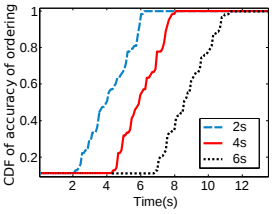


Fig. 27. CDF of the accuracy of ordering if somebody blocks tags for a while

TABLE 3
Accuracy along the Y dimension when varying d_t

d_t(cm)	0.5	1	2	3	4	5
Accuracy	0.881	0.881	0.857	0.809	0.762	0.524

Tolerant distance d_t vs. Accuracy: To detect the tolerant distance d_t, we also conduct an experiment in which we vary the average misaligned distance between the tags in a same column from 0.5 cm$\sim$ 5 cm. The distance between two adjacent columns is 10 cm. The results are listed in Tab. 3. We find that the accuracy is higher than 0.809 even if $d_t \leq 3$ cm, demonstrating that HMO is resilient to misalignments.

Layer height vs. Accuracy: We vary the layer height from 20 cm to 40 cm. The accuracy is shown in Fig. 24. Different layer heights may result in different shelves' heights (the larger the layer height, the higher the shelves), which may influence the performance of HMO. That is because, on one hand, if the shelves are very high, the tags on the top layer may not be blocked at the same time with that on the lower layers. On the other hand, the lower the layer height, the denser the tags in different columns. According to Fig. 24, we find that except $\ell = 20$ cm, HMO outperforms both STPP and RSS in terms of the accuracy of 2D relative localization.

Tag separation distance vs. Accuracy: To simulate the real-world layout of objects, we tune the interval distance between two tags in each layer within the range of [5 cm, 40 cm]. The localization results are shown in Fig. 25. We find that the larger the distance is between tags, the higher the accuracy HMO can achieve along the Y dimension. The 2D relative localization accuracy of HMO is always higher than 88.10%.

Reader distance to shelves vs. Accuracy: We then adjust the distance between the shelf and the reader antennas from 1.2m to 2.4m. The results are plotted in Fig. 26. We find that with the increase of above distance, HMO presents a decrease in its localization accuracy along the Y dimension, but an increase in the one along the Z dimension. Under most situations, HMO performs better than STPP and RSS.

TABLE 4
Accuracy vs. Human moving speed

Human speed (m/s)	1	1.5	2
Accuracy	0.905	0.833	0.810

Human moving speed vs. Accuracy: We also evaluate the influence of human moving speed on the system accuracy. We invite one volunteer to walk through the inventory space randomly with a speed of 1m/s and 1.5m/s, 2m/s, respectively. As shown in Tab. 4, the average accuracy is always higher than 80%.

Blocked tags: In our system, we assume that we can always receive all the tags' data. However, in practice, the tags may be blocked by the customer and can not be read for a little while. HMO addresses this issue by periodically collecting data within each time window and leverage multiple ordering results to reach a final result. We conduct experiments where a stop-and-go customer moves in the region. He first moves inside and block the first column of tags for 2s, 4s and 6s, respectively. Then he passes through the inventory area. As shown in Fig. 27, the x-axis is the

TABLE 5
The process time of HMO

Tags' number	1	9	36	42	56	63
Time(s)	0.13	0.28	0.39	0.42	0.56	0.76

running time of our system. At start, due to the blockage of human being, we can not obtain all the tags' data. However, when the person moves and pass through the shelf, HMO can successfully obtain the tags' right order after several seconds.

Time cost: We also evaluate the processing time of HMO. As shown in Tab. 5, we vary the number of tags and observe the process time of HMO. The results show that even if the population of tags is up to 63, the processing time is still very short (less than 0.8s). HMO is very efficient.

8 Discussion

Different placement of tags: In our system, we assume the tags are placed in a grid so that they can be partitioned into rows and columns. However, in practice, some grids may have more than one tags and some grids may have no tag. If the placement of tags in known as the input of HMO in advance, our algorithm is also effective. Fortunately, the tag placement is mostly stable in most applications. For example, one kind of items in the supermarket may be always placed at the same position, and the books or documents in a library or a chancery should be put at the certain places. With necessary inputs, HMO may cope with the problems introduced by different placement types. We will leave this study to future work.

Tag's orientation: The tags' on the items may face to different directions. So in our system, we employ circular polarization antennas, which can read the tags in different orientations.

The choice of ordering method: Intuitively, the phase-based approach is sufficient for extracting the location of each tag. However, in our work, we only employ the phase-based method along the Z dimension, while choosing an RSS-based method to order the tags in the Y dimension. In fact, either of the RSS-based and phase-based methods has its own advantages. For the RSS-based method, it requires fewer antennas (at least 3 antennas are needed with the phase-based method for 2-dimension ordering). However, it relies on the movement of human beings. On the other hand, the phase-based method does not rely on walking-by persons. However, it only works in static environments. In addition, due to the uncertainty of the phase data, the number of antennas will increase accordingly with the enlargement of the application area. In other words, the larger the application environments are, the more antennas we need in the phase-based algorithm. After comparing their advantages and drawbacks, we employ an RSS-based method along Y-dimension and the experiments show that it works well for horizontal ordering.

9 Conclusion

In this paper, we propose HMO, a relative localization system for passive tags by utilizing human movement. Compared to prior solutions, HMO is more efficient and convenient to deploy, thus it enables more applications. HMO can also cope with many practical factors in real deployments, such as multipath effects, untidy object position, irregularity of moving objects, *etc.*. The experimental results show that HMO is accurate and practical in the relative localization for real-world applications.

10 Acknowledgment

This work was partially supported by National Key R&D Program of China 2017YFB1003000, NSFC Grant No. 6182285, 61572396, 61751211, 61772413, 61672424, 61802299, Alibaba-Zhejiang University Joint Research Institute of Frontier Technologies and Project funded by China Postdoctoral Science Foundation 2018M643663. Chen Qian was partially supported by National Science Foundation Grants CNS-1717948 and CNS-1750704.

References

[1] S. Pradhan, E. Chai, K. Sundaresan, L. Qiu, M. A. Khojastepour, and S. Rangarajan, "RIO: A Pervasive RFID-based Touch Gesture Interface," in *Proceedings of ACM Mobicom*, 2017.
[2] G. Wang, C. Qian, J. Han, W. Xi, H. Ding, Z. Jiang, and J. Zhao, "Verifiable smart packaging with passive RFID," in *Proceedings of ACM Ubicomp*, 2016.
[3] L. Yang, Y. Chen, X. Li, C. Xiao, M. Li, and Y. Liu, "Tagoram: Real-time tracking of mobile RFID tags to high precision using COTS devices," in *Proceedings of ACM MOBICOM*, 2014.
[4] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, "LANDMARC: Indoor location sensing using active RFID," *Wireless Networks*, vol. 10, no. 6, pp. 701–710, 2004.
[5] J. Wang and D. Katabi, "Dude, where's my card?: RFID positioning that works with multipath and non-line of sight," in *Proceedings of ACM SIGCOMM*, 2013.
[6] T. Liu, L. Yang, Q. Lin, Y. Guo, and Y. Liu, "Anchor-free backscatter positioning for RFID tags with high accuracy," in *Proceedings of IEEE INFOCOM*, 2014.
[7] L. Shangguan, Z. Li, Z. Yang, M. Li, and Y. Liu, "OTrack: Order tracking for luggage in mobile RFID systems," in *Proceedings of IEEE INFOCOM*, 2013.
[8] L. Shangguan, Z. Yang, A. X. Liu, Z. Zhou, and Y. Liu, "STPP: Spatial-temporal phase profiling-based method for relative RFID tag localization," *IEEE/ACM Transactions on Networking*, vol. 25, no. 1, pp. 596–609, 2017.
[9] ——, "Relative localization of RFID tags using spatial-temporal phase profiling," in *Proceedings of USENIX NSDI*, 2015.
[10] J. Wang, D. Vasisht, and D. Katabi, "RF-IDraw: Virtual touch screen in the air using RF signals," in *Proceedings of ACM SIGCOMM*, 2014.
[11] Y. Ma, N. Selby, and F. Adib, "Drone Relays for Battery-Free Networks," in *Proceedings of ACM SIGCOMM*, 2017.
[12] ——, "Minding the billions: Ultra-wideband localization for deployed rfid tags," in *Proceedings of ACM Mobicom*, 2017.
[13] Y. Ma, X. Hui, and E. C. Kan, "3D Real-time Indoor Localization via Broadband Nonlinear Backscatter in Passive Devices with Centimeter Precision ," in *Proceedings of ACM Mobicom*, 2016.
[14] L. Shangguan and K. Jamieson, "The design and implementation of a mobile RFID tag sorting robot," in *Proceedings of ACM MobiSys*, 2016.
[15] L. Yang, Y. Qi, J. Fang, X. Ding, T. Liu, and M. Li, "Frogeye: Perception of the slightest tag motion," in *Proceedings of IEEE INFOCOM*, 2014.
[16] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient backscatter: wireless communication out of thin air," *Proceedings of ACM SIGCOMM*, 2013.
[17] D. Arnitz, K. Witrisal, and U. Muehlmann, "Multifrequency continuous-wave radar approach to ranging in passive UHF RFID," *Transactions on Microwave Theory and Techniques*, vol. 57, no. 5, pp. 1398–1405, 2009.
[18] X. Li, Y. Zhang, and M. G. Amin, "Multifrequency-based range estimation of RFID tags," in *Proceedings of IEEE RFID*, 2009.
[19] G. Wang, C. Qian, J. Han, and H. Cai, "Poster: Combating Multipaths to Enable RFID Sensing in Practical Environments," in *Proceedings of ACM Mobicom*, 2017.
[20] C. Hekimian-Williams, B. Grant, X. Liu, Z. Zhang, and P. Kumar, "Accurate localization of RFID tags using phase difference," in *Proceedings of IEEE RFID*, 2010.

[21] S. Azzouzi, M. Cremer, U. Dettmar, R. Kronberger, and T. Knie, "New measurement results for the localization of UHF RFID transponders using an angle of arrival (AOA) approach," in *Proceedings of IEEE RFID*, 2011.

[22] P. V. Nikitin, R. Martinez, S. Ramamurthy, H. Leland, G. Spiess, and K. Rao, "Phase based spatial identification of UHF RFID tags," in *Proceedings of IEEE RFID*, 2010.

[23] R. Miesen, F. Kirsch, and M. Vossiek, "Holographic localization of passive UHF RFID transponders," in *Proceedings of IEEE RFID*, 2011.

[24] R. Kronberger, T. Knie, R. Leonardi, U. Dettmar, M. Cremer, and S. Azzouzi, "UHF RFID localization system based on a phased array antenna," in *Proceedings of IEEE APSURSI*, 2011.

[25] J. Zhou, H. Zhang, and L. Mo, "Two-dimension localization of passive RFID tags using AOA estimation," in *Proceedings of IEEE I2MTC*, 2011.

[26] G. Mao, B. Fidan, and B. D. Anderson, "Wireless sensor network localization techniques," *Computer networks*, vol. 51, no. 10, pp. 2529–2553, 2007.

[27] K. Chawla, C. McFarland, G. Robins, and C. Shope, "Real-time RFID localization using RSS," in *Proceedings of IEEE ICL-GNSS*, 2013.

[28] J. Zhou and J. Shi, "RFID localization algorithms and applicationsła review," *Journal of intelligent manufacturing*, vol. 20, no. 6, pp. 695–707, 2009.

[29] M. Bouet and A. L. Dos Santos, "RFID tags: Positioning principles and localization techniques," in *Proceedings of IEEE WD*, 2008.

[30] A. Parr, R. Miesen, and M. Vossiek, "Inverse sar approach for localization of moving RFID tags," in *Proceedings of IEEE RFID*, 2013.

[31] L. Yang, Q. Lin, X. Li, T. Liu, and Y. Liu, "See through walls with COTS RFID system!" in *Proceedings of ACM MOBICOM*, 2015.

[32] H. Ding, J. Han, A. X. Liu, J. Zhao, P. Yang, W. Xi, and Z. Jiang, "Human object estimation via backscattered radio frequency signal," in *Proceedings of IEEE INFOCOM*, 2015.

[33] F. Adib, Z. Kabelac, and D. Katabi, "Multi-person localization via RF body reflections," in *Proceedings of USENIX NSDI*, 2015.

[34] L. Shangguan, Z. Zhou, X. Zheng, L. Yang, Y. Liu, and J. Han, "Shopminer: Mining customer shopping behavior in physical clothing stores with COTS RFID devices," in *Proceedings of ACM SenSys*, 2015.

[35] J. Han, H. Ding, C. Qian, D. Ma, W. Xi, Z. Wang, Z. Jiang, and L. Shangguan, "CBID: A customer behavior identification system using passive tags," in *Proceedings of IEEE ICNP*, 2014.

[36] H. Ding, L. Shangguan, Z. Yang, J. Han, Z. Zhou, P. Yang, W. Xi, and J. Zhao, "FEMO: A platform for free-weight exercise monitoring with RFIDs," in *Proceedings of ACM SenSys*, 2015.

[37] L. Shangguan, Z. Zhou, and K. Jamieson, "Enabling Gesture-based Interactions with Objects," in *Proceedings of ACM MobiSys*, 2017.

[38] H. Ding, C. Qian, J. Han, G. Wang, W. Xi, K. Zhao, and J. Zhao, "RFIPad: Enabling Cost-efficient and Device-free In-air Handwriting using Passive Tags," in *Proceedings of IEEE ICDCS*, 2017.

[39] Y. Liu and Z. Li, "aleak: Privacy leakage through context-free wearable side-channel," in *Proceedings IEEE INFOCOM*, 2018.

[40] L. Xie, C. Wang, A. X. Liu, J. Sun, and S. Lu, "Multi-touch in the air: Concurrent micromovement recognition using rf signals," *IEEE/ACM Transactions on Networking*, vol. 26, no. 1, pp. 231–244, 2018.

[41] L. Xie, Q. Li, C. Wang, X. Chen, and S. Lu, "Exploring the gap between ideal and reality: An experimental study on continuous scanning with mobile reader in rfid systems," *IEEE Transactions on Mobile Computing*, vol. 14, no. 11, pp. 2272–2285, 2015.

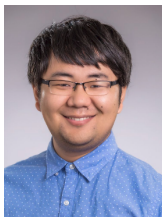
[42] Z. Li, Y. Xie, M. Li, and K. Jamieson, "Recitation: Rehearsing wireless packet reception in software," in *Proceedings of ACM MobiCom*, 2015.

[43] EPCglobal, "EPC^{TM} radio-frequency identity protocols class-1 generation-2 UHF RFID protocol for communications at 860 MHz–960 MHz," 2005.

[44] T. Wei and X. Zhang, "Gyro in the air: tracking 3d orientation of batteryless internet-of-things," in *Proceedings of ACM MobiCom*, 2016.

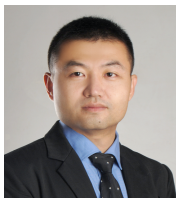
Ge Wang is now a Ph.D student at Xi'an Jiaotong University. She is currently a visiting student at University of California, Santa Cruz. She received her B.S. at Xi'an Technological University in 2013. Her research interests include wireless sensor network, RFID and mobile computing.

Chen Qian received the Ph.D. degree from the University of Texas at Austin in 2013. He is now an Assistant Professor at the Department of Computer Engineering, University of California Santa Cruz. His research interests include computer networking, data-center networks, software-defined networking, and mobile computing. He is a member of IEEE and ACM.



Longfei Shangguan received the B.S. degree from the Scool of Software, Xidian University, Shanghai, China, in 2011, and the Ph.D. degree from the Department of Computer Science and Engineering, the Hong Kong University of Science and Technology, Hong Kong, in 2015. He is currently a Researcher at Microsoft. His research interests are wireless networks, mobile systems, and low-power communication.

Han Ding received her Ph.D. degree in computer science and technology from Xi'an Jiaotong University in 2017. She is currently an assistant professor in Xi'an Jiaotong University. Her research interests focus on RFID system and smart sensing.



Jinsong Han received his Ph.D. degree in computer science from Hong Kong University of Science and Technology in 2007. He is now a professor in Zhejiang University. He is a senior member of ACM and IEEE. His research interests focus on mobile computing, RFID, and wireless network.

Kaiyan Cui is pursuing the Ph.D degree at Xi'an Jiaotong University, Xi'an, China. She received her B.S. at Taiyuan University of Technology in 2016. Her research interests include RFID, mobile computing and smart sensing.

Wei Xi received his Ph.D. degree in computer science computer science and technology from Xi'an Jiaotong University in 2014. He is now an associate professor in Xi'an Jiaotong University. He is a member of CCF, ACM, and IEEE. His research interests focus on wireless networks, smart sensing, and mobile computing.

Jizhong Zhao received his Ph.D. degree in computer science computer science and technology from Xi'an Jiaotong University in 2001. He is a member of CCF, ACM, and IEEE. His research interests focus on computer software, pervasive computing, distributed systems, network security.