

# See Through Walls with COTS RFID System!

Lei Yang\*, Qiongzhen Lin\*, Xiangyang Li\*<sup>†‡</sup>, Tianci Liu\*, Yunhao Liu\*

\*School of Software, Tsinghua University, China

<sup>†</sup> Department of Computer Science and Technology, Tsinghua University, China

<sup>‡</sup> Department of Computer Science, Illinois Institute of Technology, USA

young@tagsys.org, qiongzhen@tagsys.org, xli@cs.iit.edu, tian@tagsys.org,  
yunhao@greenorbs.com

## ABSTRACT

Through-wall tracking has gained a lot of attentions in civilian applications recently. Many applications would benefit from such device-free tracking, *e.g.* elderly people surveillance, intruder detection, gaming, *etc.* In this work, we present a system, named *Tadar*, for tracking moving objects without instrumenting them using COTS RFID readers and tags. It works even through walls and behind closed doors. It aims to enable a *see-through-wall* technology that is *low-cost*, *compact*, and accessible to *civilian* purpose. In traditional RFID systems, tags modulate their IDs on the backscatter signals, which is vulnerable to the interferences from the ambient reflections. Unlike past work, which considers such vulnerability as detrimental, our design exploits it to detect surrounding objects even through walls. Specifically, we attach a group of RFID tags on the outer wall and logically convert them into an antenna array, receiving the signals reflected off moving objects. This paper introduces two main innovations. First, it shows how to eliminate the flash (*e.g.* the stronger reflections off walls) and extract the reflections from the backscatter signals. Second, it shows how to track the moving object based on HMM (Hidden Markov Model) and its reflections. To the best of our knowledge, we are the first to implement a through-wall tracking using the COTS RFID systems. Empirical measurements with a prototype show that Tadar can detect objects behind 5'' hollow wall and 8'' concrete wall, and achieve median tracking errors of 7.8cm and 20cm in the X and Y dimensions.

## Categories and Subject Descriptors

C.2 [Computer Systems Organization]: Computer Communications Networks

## Keywords

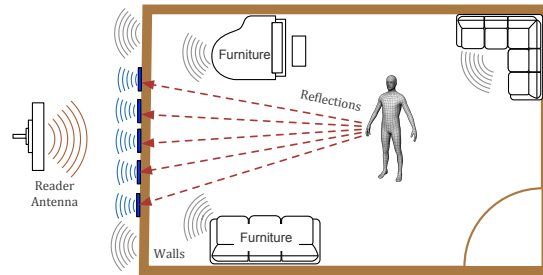
RFID, See-through walls, Tracking, Tadar

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

MobiCom'15, September 7–11, 2015, Paris, France.

© 2015 ACM. ISBN 978-1-4503-3619-2/15/09\$15.00

DOI: <http://dx.doi.org/10.1145/2789168.2790100>.



**Figure 1: Through-wall tracking of moving object.** Tadar logically transforms the tags attached on the outer wall into an antenna array receiving the reflections off the human behind the wall.

## 1. INTRODUCTION

Motion tracking and localization systems have received a lot of attentions recently. A number of innovative systems using different technologies are developed for improving the accuracy and robustness, *e.g.*, [?, ?, ?, ?]. However, these systems typically require the objects to carry an RF device in order to be identified and localized. This paper explores the potential of using COTS RFID systems to build a system that can capture the motion of mobile object behind a wall and in closed rooms without tagging it. Such system is able to provide life-saving benefits for emergency responders, police and military personnel arriving at dangerous situations by letting them know the locations of people inside a building without instrumenting the people in advance. Many other applications would also benefit from this system, such as elderly, people surveillance, intruder detection, and so on.

The concept underlying through-wall tracking is to emit RF signals to a closed room facing with a non-metallic wall. The signal traverses through the wall, reflects off objects and humans, and returns back, imprinted with a glimpse of what is inside a closed-room. Through-wall tracking has been extensively studied in radar community, which mainly employs ultra-wideband and MIMO radar technique [?, ?] to detect moving objects behind a wall. Their devices are almost inaccessible to non-military purpose due to ultra-widebandwidth, voracious power consumption, and unaffordable. For example, a typical radar device consumes 2 ~ 4 GHz bandwidth [?, ?, ?]. The price of the through-wall radar sold online is about 80,000 US dollars [?]. Even worse, the through-wall radar systems in the market are usually prohibited from selling to civilians in most countries in the world. To address these issues, recently several breakthrough systems [?, ?, ?, ?, ?] are developed

for seeing through the walls with WiFi. However, these past systems still need either customized devices (e.g. USRP), specialized signals (e.g. FMCW), or additional receivers inside the room.

In this work, we turn our attentions to a mature technology, RFID. RFID is evolving as a major technology enabler for identifying and tracking objects all around the world. The reason for this widespread deployment is the simplicity of tags, which enables very low cost at high volumes. One of the benefits of RFID systems is that it can identify objects without needs of line-of-sight (LOS) owing to the multi-path effect. The challenging question is “Can RFID systems enable us to identify and track objects through walls?”. At the first instinct, it seems like an impossible mission because the backscatter signals off tags are too weak (typically at -55 dBm) to traverse a wall (with  $\sim 18$  dB attenuation) or even a thick glass (with  $\sim 3$  dB attenuation) [?].

It is well known that RFID systems adopt an asymmetric architecture in order to reduce the complexity and costs of tags. Such design has three important features. First, the reader is able to provide 30 dBm transmitting signals and connected to higher-gain directional antennas. Such strong signals are sufficient for two-way traversal through a wall. Second, the tag has no transceiver and modulates its ID on the backscatter signals through changing the impedance on its antenna. Its modulation is vulnerable to the interference from ambient reflections [?]. Unlike past approaches [?, ?], which consider this vulnerability as detrimental, we exploit it as an opportunity to monitor the surrounding objects. Third, the RFID technology, stemming from the radar, has strong ability to resolve very weak signals from the background. Inspired by these features, we redefine RFID system to extend our senses. We leverage the reader to supply RF signals lighting up objects and treat tags as the antennas receiving reflections. To illustrate Tadar’s approach, Fig. 1 shows a toy example. We attach an array of tags on the outer wall and use a reader to read them. The signal propagation between reader and tags is kept in line-of-sight. Apart from tags, the transmitting signals from reader are also reflected off other objects like walls, furnitures and mobile objects. These reflections plus the original signals from reader provide a combined signal to each tag for their modulations. As a result, the backscatter signal that carries tag’s ID implicitly contains the reflections off human. If we can extract the human’s reflections from the backscatter signals, his locations would be pinpointed.

Transforming this high-level idea into a practical system, however, requires addressing multiple challenges. First, it is quite challenging to separate a *weaker* signal (reflections off object) from another weak signal (backscattered signal). In particular, this task should be accomplished using the COTS RFID reader, which does not support any low-level signal access or modifications. The second challenge stems from ‘Flash Effect’ (see [?]) that the signal power after traversing the wall twice (in and out of the room) is reduced by three to five orders of magnitude, such that the reflections off objects inside the room will be drown in the reflections off the wall itself (call *flash*). Third, the inherent diversity of devices and objects would introduce many unknown impacts on the magnitude attenuations and phase rotations, leading to *unpredictable* tracking errors.

To address these issues, we present *Tadar* for tracking moving object through walls and behind closed doors using COTS RFID readers and tags. Our objective is to enable a *see-through-wall* technology that is *low-cost*, *compact*, and accessible for *civilian* purpose. We propose two main techniques to address the aforementioned challenging issues. First, Tadar views each read of tag as one channel estimation and recovers the signal with RSS and phase reported by reader. Since the wireless signals linearly combine over

the medium, Tadar eliminates the flash by subtracting the signals learned when the room is empty. It also separates the reflections from the backscatter signals by dividing the signal learned in LOS propagation (see §5). Second, Tadar models the object’s motions using Hidden Markov Model (HMM), which considers the object’s locations as hidden states and its reflections as observations. Then it uses Viterbi decoding to recover the object’s trajectory (see §6).

**Summary of results:** We build a prototype of Tadar using one Impinj Reader and 45 Alien tags, and evaluate it in two office buildings (see §7). Our main results are as follows:

- Tadar can detect object moving behind 8’’ concrete walls, 5’’ hollow walls, and 1.5’’ wooden doors.
- Tadar can achieve detecting ranges of 4m and 6m on average respectively when pointing at a closed room with 8’’ concrete walls and 5’’ hollow walls supported by steely frames.
- Tadar can track the moving object in a closed room with median errors of 7.8cm and 20cm in X and Y dimensions when monitoring the room from one side, offering about 3 $\times$ , 3.5 $\times$  and 1.3 $\times$  improvement compared with NRES [?], TagTrack [?] and Wision [?].
- Tadar can correctly decode 93% gestures in our application study where 5 volunteers encode the messages through his/her movements.

**Contributions:** In contrast to the state-of-the-art work, Tadar introduces a low-cost solution to the through-wall tracking of moving objects, enabling civilians to use this technology. Tadar is the first to introduce through-wall tracking using RFID systems by transforming a group of tags attached on the wall into a receiving antenna array. It demonstrates a working system implemented purely based on COTS devices and made accessible to the general public.

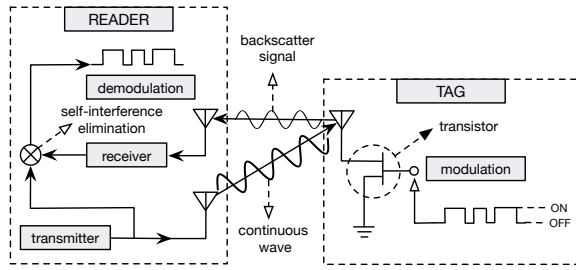
The rest of the paper is organized as follows. We review the related work in §2 and the technical background in §3. The main design is overviewed in §4. The technique details of reflection extraction and object tracking are respectively elaborated in §5 and §6. The implementation is described in §7 and the system is evaluated in §8. We also conduct an application study in §9. Lastly, we conclude our work in §11.

## 2. RELATED WORK

In this section, we briefly review the related literature in see-through-wall technology.

**See through walls with radar.** Through-wall imaging or tracking has gained attentions for a decade, especially in radar community [?, ?, ?, ?, ?, ?, ?]. The past systems eliminate the flash effect and localize the moving object by estimating the time of flight of signals. Specifically, the pulse reflected off the wall arrives earlier in time than that reflected off moving object behind it. To differentiate such slight time difference ( $< 1ns$ ), the systems would transmit a linear frequency chirp at cost of ultra-wide bandwidth of the order of 2 GHz [?]. The products developed in the industry [?, ?] also require multiple GHz bandwidth. Adib *et al.* track user’s 3D motion or heartbeat [?, ?]. These works utilize the FMCW signals (sweeping from 5.46-7.25 GHz) to estimate the time-of-flight, so it is actually UWB based. UWB would continuously occupy a wide band and may impose interference on other devices that share the same frequency. Tadar employs the COTS RFID devices communicate in 50 KHz narrow band channels [?], overcoming the need of UWB.

**See through walls with Wi-Fi.** Some efforts proposed recently attempt to break the limitations of UWB systems and exploit the potentials of using Wi-Fi for through-wall technologies [?, ?, ?, ?, ?]. Chetty *et al.* [?] propose to deploy the transmitter and a reference receiver in the room, which cooperates with other receivers



**Figure 2: Block diagram of backscatter communication used in UHF RFID system.** The tag modulates its ID on the backscatter signal using ON-OFF keying through changing the impedance on its antenna.

outside to see objects in the room. This method is limited by its deployment. Adib *et al.* introduce a system, called Wi-Vi, that borrows the MIMO interference nulling to eliminate the flash effect [?]. It creatively treats the motion of a human body as an antenna array and tracks the resulting RF beam. Our work differs from Wi-Vi in that Wi-Vi focuses on the relative motions of human (*e.g.* walking direction) while Tadar aims to exactly recover object's trajectory. Huang *et al.* explore the feasibility of achieving computational imaging using Wi-Fi signals. Their system works without obstruction [?]. Depatla *et al.* proposed a see-through imaging based on WiFi measurements with unmanned vehicles [?, ?]. In contrast, Tadar does not need any customized devices or specialized signals and really operates at the narrow-band.

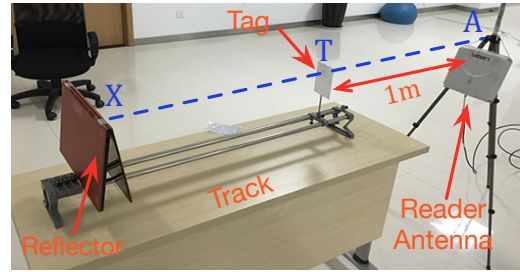
**See through walls with WSN.** Some past works in radio tomography employ networks or hundred sensors to track a person even if he/she does not carry any wireless device. They can be categorized into following groups, *link-based* and *location-based*. (1) *Link-based schemes* measure the RSSI for each of the resulting  $n^2$  links among  $n$  sensor nodes, and attribute the variation of RSSI on a link to a human crossing that link [?, ?, ?, ?, ?]. Wilson and Patwari find that the changes in RSSI due to human motion can be modeled by the skew-Laplace distribution [?]. Kleisouris *et al.* exploit channel diversities to enhance motion detection accuracy [?]. Nannuru *et al.* develop a measurement model for multi-target tracking using RF tomography, applied for tracking up to three targets [?]. (2) *Location-based schemes* are also known as the radio fingerprint, which is a popular approach for RF-based localization [?, ?, ?]. Zhang *et al.* firstly collect a radio map when the subject presents in a few predetermined locations, and then map the test location to one of these trained locations [?]. Moustafa extends RASS to a much large-scaled deployment [?]. Xu *et al.* employ camera to trigger recalibration and update the radio signal data to improve the localization accuracy [?]. These two groups mainly focus on the metric of RSSI, which is not a stable location indicator in RFID systems.

### 3. UNDERSTANDING REFLECTIONS

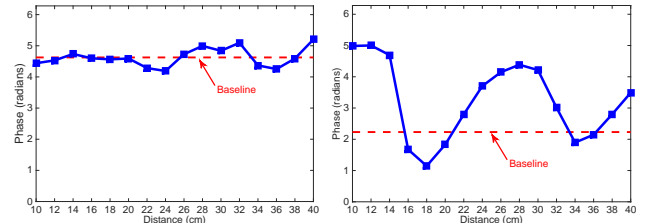
In this section, we firstly introduce the technical background of RFID system, and then conduct a series of empirical studies on the backscatter communications.

#### 3.1 Technical Background

Passive RFID system communicates using a backscatter radio link, as shown in Fig. 2. The reader supplies a Continuous Wave (CW) signal, a periodic signal that persists indefinitely without



(a) Experiment setup



(b) Indirect impact

(c) Direct impact

**Figure 3: Empirical studies on measured RF phase.** (a) The default setup of our experiments. (b) The indirect impact on backscatter signals. (c) The direct impact on backscatter signals.

changing its amplitude, frequency or phase, for tags. An RFID reader communicating with a passive tag must operate in full-duplex mode, in the sense that it must supply CW for the tag to backscatter while listening to the tag response at the same time. The tag with no battery equipped, purely harvests energy from the reader's signal. A simple way to understand backscatter radio is as follows [?]. The current flowing on reader's transmitting antenna leads to a voltage induced on the tag's antenna. This induced current also leads to radiation, producing a signal that can be detected: a *backscattered signal*. The tag modulates its ID on the backscatter signal using ON-OFF keying through changing the impedance on its antenna to transmit a '1' bit and remaining in its initial silent state for '0' bit. Referring to [?] published by the ImpinJ, a reader might employ open-loop estimation techniques such a preamble correlation or closed-loop estimation for acquiring and/or tracking *carrier phase* and *RSS*. For simplicity and saving cost, most of modern RFID readers use one single antenna both for transmitting and receiving.

#### 3.2 Empirical Studies

We conduct empirical studies to better understand the characteristics of backscatter communications, using a COTS ImpinJ reader and an Alien tag. The experiment setup is shown in Fig. 3(a). By default, the reader antenna *A*, tag *T* and reflector *X* are deployed in a straight line. The tag attached on a plastic card, locates between the reader and the reflector. The reader and tag are deployed with 1m apart from each other. We horizontally move the reflector (*i.e.* the notebook) to observe how its reflections take impacts on the phase values of the backscatter signal<sup>1</sup>. In an ideal environment, there are two basic channels, forward channel ( $A \rightarrow T$ ) and backward channel ( $T \rightarrow A$ ), between the reader and tag. In our scenario, however, there emerge two additional signal propa-

<sup>1</sup>The phase is chosen as our metric due to its high resolution and sensitivity to reflections [?].



gations,  $A \rightarrow X \rightarrow A$  and  $A \rightarrow X \rightarrow T$ , due to the existence of reflector. We categorize the impacts of these two propagations into two types based on whether the reflections *directly* or *indirectly* affect on the tag's backscatter signal.

- *Indirect impact*: The transmitting signal arrives at  $X$  and is reflected back to  $A$  (i.e.  $A \rightarrow X \rightarrow A$ ). As a result, the reflection mixes with the tag's backscatter signal at reader end, which may *indirectly* affect the identification of the backscatter signal.

- *Direct impact*: The transmitting signal reflected to  $T$  (i.e.  $A \rightarrow X \rightarrow T$ ) adds up with the original signal through forward channel (i.e.  $A \rightarrow T$ ), producing a combined signal at tag end. Since the tag uses this combined signal to modulate its ID, the received backscatter signal contains a *direct* impact from this propagation.

**HYPOTHESIS 1.** *The indirect impact of reflections off surrounding objects can be negligible.*

We firstly focus on the indirect impact incurred by the reflections propagated from  $A \rightarrow X \rightarrow A$ . To isolate the indirect impact, we prevent the reflections from traversing to the tag by placing an RF absorber (with approximately 35 dB attenuation) between the tag and the reflector. The reflector is much larger than the absorber, so there are still many reflections getting back to the reader antenna. We move the reflector to adjust the distance to tag between 10cm and 40cm. The phase measurements are repeated 100 times for a particular distance. The final phase values reported by the reader are illustrated in Fig. 3(b). The dotted line indicates the baseline, i.e. the phase value measured when removing the reflector but keeping the absorber. Surprisingly, the figure suggests that the phase value vibrates around the baseline with a very small deviation ( $< 1$  radians)<sup>2</sup> no matter what distance is adjusted. In other words, the impact of propagation from  $A \rightarrow X \rightarrow A$  is reduced to be minimal.

Why the indirect impacts are removed? The RFID systems are *originally* designed for identification. Its main task is to resolve the weak backscatter signal from the CW and environment noise. The reflection from  $A \rightarrow X \rightarrow A$  is a kind of *self-interference* because it does not carry backscatter preamble and highly correlates to the original CW that the reader emits. It can be easily removed by learning the environment before interrogating tag. Removing the self-interference is a standard procedure in RFID reader [?, ?] (see Fig. 2), and also becomes an important metric to evaluate the reader's performance. Any modern COTS RFID reader has the strong ability to eliminate the self-interference, so the indirect impact can be negligible in practice.

**HYPOTHESIS 2.** *The backscatter signals are vulnerable to the direct impact of the reflections off surrounding objects.*

This hypothesis is our foundation of leveraging tag to sense the reflections off surrounding objects. In this experiment, we remove the RF absorber and re-collect the phase values. The results are plotted in Fig. 3(c). In contrast to the last experiment, the phase value periodically changes as the distance increases. Apparently, the backscatter signals are vulnerable to the direct impact. The tag modulates its ID using the combined signal (through  $A \rightarrow T$  and  $A \rightarrow X \rightarrow T$ ). The received backscatter signal contains a standard preamble so the reader would not filter it. Intuitively, the phase non-linearly changes and gradually approaches to the baseline as the distance increases. This is because the direct impact from  $A \rightarrow X \rightarrow T$  gradually weakens as the distance increases, resulting in the signal through  $A \rightarrow T$  gradually dominates the backscatter signal. We will quantitatively model the reflections in §5.

<sup>2</sup>The vibrations are mainly caused by the thermal noise [?].

## 4. TADAR OVERVIEW

Tadar is a fine-grained UHF RFID through-wall tracking system that can locate and track moving object including human, and can provide a resolution on the order of a few to tens of centimeters, much smaller than read range of reader.

**System Scope.** Ultra-low cost UHF tags (5-10 cents each) become the preferred choice of many industry applications. Following the common practices, we concentrate on the UHF tags. We assume that tags are well-positioned so their locations are all known. We also assume that the surveillance region is confined within a limited space. Today's UHF RFID has an operating range of around 10m. Thus, the assumption is acceptable and can be effectively satisfied by UHF RFID systems. To avoid the tag collisions, the reader adopts TDMA fashion to read them and discard the collided signals [?]. Thus, the effects among tags are negligible. Our systems works in the application layer of reader with the reported RSS and phase values. Our system is general and can capture any mobile objects like robot or drone.

**Problem Definition.** Suppose  $n$  tags,  $\{T_1, T_2, \dots, T_n\}$ , are attached on the outer wall. We use the term  $T_i$  to denote the  $i^{th}$  tag as well as its coordinate. One reader antenna, denoted as  $A$ , is employed to light up the objects and read tags. The signals propagated between reader and tags are always kept in LOS. During the surveillance, the geometric relationships between tags remain unchanged. The reader is configured to continuously and cyclically read tags, which resembles to using a camera to continuously take snapshots of the room. One *read frame* is finished after  $n$  tags are all read once. Each frame contains  $n$  tags' RSS and phase values reported by a COTS RFID reader. *Our problem is to recover the trajectory of the moving object behind walls and closed doors given a sequence of read frames.*

**Solution Sketch.** At a high level, Tadar decomposes the problem into two subproblems.

- *Extracting reflections*: Since Tadar uses the reflections off the moving object for tracking, the first subproblem is how to extract the reflections while eliminating the flash effect. This subproblem is addressed in §5.

- *Tracking objects*: The second subproblem that Tadar addresses in §6 is how to recover the mobile object's trajectory using the extracted reflections.

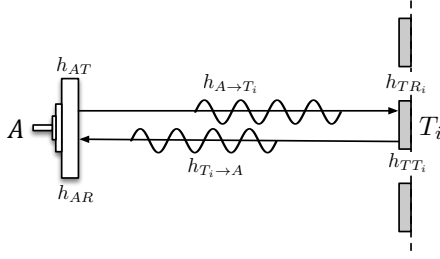
The next two sections elaborate on these two subproblems, providing the technical details.

## 5. EXTRACTING REFLECTIONS

In order to distinguish the signals reflected off moving object, off static objects like wall and furniture, and off tags, we respectively call these signals *reflections*, *flash*, and *backscatter signals*. The problem of extracting reflections off objects behind walls is exacerbated by two issues.

First, the signal power after traversing the wall twice (in and out of the room) is reduced by three to five orders of magnitude [?, ?]. The wall is much larger than the object of interest and has a higher reflection coefficient, leading to stronger reflections than that off objects behind the wall. It is like placing a mirror in front of a camera that reflects the camera's flash and prevents it from capturing objects in the scene, which is called *flash effect* [?].

Second, previous through-wall systems all directly receive the reflections through the real antennas connected to the transceiver. In contrast, Tadar logically converts the RFID tags attached on the wall to a virtual antenna array receiving the reflections. These virtual antennas are *wirelessly* connected to the reader and carry the reflections through the backscatter signals. Thus, an additional step



**Figure 4: Signal propagation without reflections.** There are two basic channels between the reader and the tag.

is required to separate the reflections off the moving object from the backscatter signals reflected by the tags. This is an extremely challenging task as the backscatter signal is already very weak.

To address the above challenges, we start by modeling the backscatter signals through LOS propagation. We then show how to eliminate the flash and derive the reflections from the backscatter signals. Lastly, we sketch the whole algorithm.

### 5.1 Modeling the Backscatter

A wireless signal is typically represented as a stream of complex numbers. The transmitter modulates the bits into complex symbols to transmit a packet over the wireless channel. The received signal is also represented as a stream of complex symbols. Generally, if the transmitted symbol is denoted as  $S$ , then the received symbol  $S'$  can be approximated as

$$S' = h \cdot S \quad (1)$$

where  $h = \alpha e^{j\delta}$  is also a complex number, called channel parameter, and  $\alpha$  and  $\delta$  are referred to as *channel attenuation* and its *phase rotation*. The channel can be well approximated by a linear time-variant system. If Alice and Bob concurrently transmit, their signals add up and the received signal can be represented as

$$S' = S'_A + S'_B \quad (2)$$

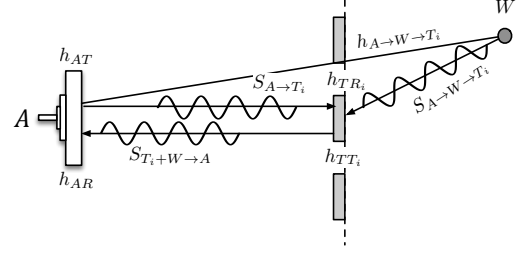
where  $S'_A = h_A S_A$  and  $S'_B = h_B S_B$  refer to Alice's and Bob's signals after traversing their corresponding channels to the receiver. These two general features in wireless communications are also hold in backscatter links.

Now we consider the simple scenario that the signal propagates through the LOS between reader and tags without any surrounding objects (*e.g.* no reflections and flash.). In this situation, there are only two basic channels between reader  $A$  and each tag  $T_i$ : forward channel ( $A \rightarrow T_i$ ) and backward channel ( $T_i \rightarrow A$ ), as shown in Fig. 4. Let  $h_{A \rightarrow T_i}$  and  $h_{T_i \rightarrow A}$  be the parameters of these two basic channels. In an ideal channel model, the channel parameter is defined as follows [?]:

$$h_{A \rightarrow T_i} = h_{T_i \rightarrow A} = \frac{1}{(d_{A \rightarrow T_i})^2} \cdot e^{j\theta_{A \rightarrow T_i}} \quad (3)$$

where the term  $j$  denotes the imaginary number and the term  $e^{j\theta}$  represents a complex exponential signal. The  $d_{A \rightarrow T_i}$  is the Euclidean distance between  $A$  and  $T_i$ . The  $\theta_{A \rightarrow T_i}$  is the phase shift over the distance  $d_{A \rightarrow T_i}$ , which can be calculated for channel wavelength  $\lambda$  as follows.

$$\theta_{A \rightarrow T_i} = 2\pi \frac{d_{A \rightarrow T_i}}{\lambda} \bmod 2\pi \quad (4)$$



**Figure 5: Signal propagation with flash.** We treat the flash off wall and other static objects as if they were from a virtual reflection point  $W$ .

Apparently,  $d_{A \rightarrow T_i} = d_{T_i \rightarrow A}$  and  $h_{A \rightarrow T_i} = h_{T_i \rightarrow A}$  due to the geometric symmetry.

However, the real signal propagations always do not conform to the ideal model due to many factors such as gains of transmitter and receiver, reflectivity, impedance, *etc.* These factors may all introduce additional magnitude attenuations and phase shifts. We abstract their impacts as four unknown parameters,  $h_{AT}$ ,  $h_{AR}$  and  $h_{TT_i}$ ,  $h_{TR_i}$ , respectively denoting the impacts from RFID reader  $A$ 's transmitter, RFID reader  $A$ 's receiver and tag  $T_i$ 's transmitter, tag  $T_i$ 's receiver. Notice that (1) These four parameters are unknown in practice and vary over tags and readers. (2) A tag in fact does not have obvious transceiver [?]. We use  $h_{TT_i}$  and  $h_{TR_i}$  to denote its internal reflection characteristics for uniform mathematical expression. Let  $S_0$  be the transmitting signal from the reader. With regard to the above additional impacts, the real signal propagations of forward and backward channels can be represented as follows.

$$\begin{cases} S_{A \rightarrow T_i} = h_{TR_i} h_{A \rightarrow T_i} h_{AT} S_0 \\ S_{T_i \rightarrow A} = h_{AR} h_{T_i \rightarrow A} h_{TT_i} S_{A \rightarrow T_i} \end{cases} \quad (5)$$

where  $S_{A \rightarrow T_i}$  is the signal received by  $T_i$  and  $S_{T_i \rightarrow A}$  is the backscatter signal received by  $A$ . Merging these two equations, we have

$$S_{T_i \rightarrow A} = h_{AR} h_{T_i \rightarrow A} h_{TT_i} h_{TR_i} h_{A \rightarrow T_i} h_{AT} S_0 \quad (6)$$

How to estimate the received signal? Tadar views each read of tag as one channel estimation. Suppose  $\tilde{a}_{T_i \rightarrow A}$  and  $\tilde{\theta}_{T_i \rightarrow A}$  are the magnitude and phase values about  $T_i$  reported by the reader, then  $S_{T_i \rightarrow A}$  can be estimated by

$$S_{T_i \rightarrow A} \approx \tilde{a}_{T_i \rightarrow A} e^{j\tilde{\theta}_{T_i \rightarrow A}} \quad (7)$$

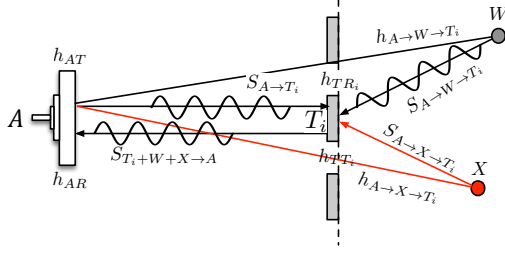
The estimated  $S_{T_i \rightarrow A}$  is useful to recover the reflections from the combined backscatter signal later. Note that the reader reports the RSS in unit of dBm instead of watts so we should take a unit conversion between RSS and magnitude. For example, suppose we received  $(RSS_{T_i \rightarrow A}, \theta_{T_i \rightarrow A})$ , the magnitude of the backscattered signal can be estimated as follows:

$$a_{T_i \rightarrow A} = 10 \sqrt{\frac{RSS_{T_i \rightarrow A}}{1000}} \quad (8)$$

In spite of involvement of physical-layer information, Tadar only uses the common parameters reported by COTS reader without any modification on hardware or firmware.

### 5.2 Eliminating the Flash

To eliminate the impact of flash, we firstly consider the scenario that the tags are attached on the outer wall and no object of interest moves inside the room. The transmitting signals are reflected



**Figure 6: The signal propagation with flash and reflections.** The flash and reflections off the moving object  $X$  add up at the tag end.

off the wall and other static objects like chairs. The key idea underlying Tadar is to treat these reflections as if they were from a virtual reflection point  $W$  due to the principle of linear superimposing of signals. As Fig. 15(a) shows, there is an additional signal (*i.e.* flash), denoted as  $S_{A \rightarrow W \rightarrow T_i}$  traveling from  $A \rightarrow W \rightarrow T_i$  for each tag besides the two basic channels. The  $S_{A \rightarrow W \rightarrow T_i}$  and  $S_{A \rightarrow T_i}$  concurrently arrive at the tag end and linearly add up with each other. Both of them are the continuous wave originated from the reader. The tag  $T_i$  uses this combined signal to modulate its ID and backscatters it to  $A$ . Then the final backscatter signal from  $T_i$ , denoted as  $S_{T_i+W \rightarrow A}$ , is also a combined signal and can be expressed as follows.

$$S_{T_i+W \rightarrow A} = h_{AR} h_{T_i \rightarrow A} h_{TT_i} (S_{A \rightarrow T_i} + S_{A \rightarrow W \rightarrow T_i}) \quad (9)$$

Second, we consider the scenario that a moving object  $X$  appears in the room as shown in Fig. 6. There exists a third signal propagating from  $A \rightarrow X \rightarrow T_i$ . We denote the corresponding signal as  $S_{A \rightarrow X \rightarrow T_i}$ . Similarly, the three signals linearly add up at tag  $T_i$ , resulting in a more complex combined backscatter signal, denoted as  $S_{T_i+W+X \rightarrow A}$ , given by:

$$S_{T_i+W+X \rightarrow A} = h_{AR} h_{T_i \rightarrow A} h_{TT_i} (S_{A \rightarrow T_i} + S_{A \rightarrow W \rightarrow T_i} + S_{A \rightarrow X \rightarrow T_i}) \quad (10)$$

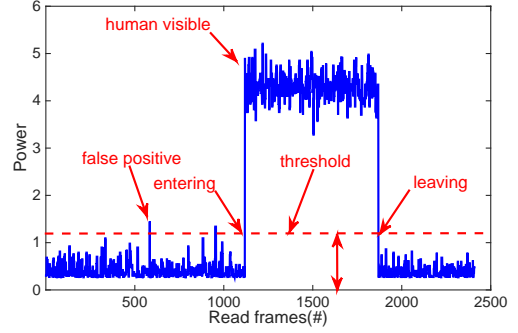
Likewise,  $S_{T_i+W \rightarrow A}$  and  $S_{T_i+W+X \rightarrow A}$  can be estimated using the measured magnitude and phase values. To eliminate the flash, we use Eqn. (10) to subtract Eqn. (9). As a result, we have

$$S_{T_i+W+X \rightarrow A} - S_{T_i+W \rightarrow A} = h_{AR} h_{T_i \rightarrow A} h_{TT_i} S_{A \rightarrow X \rightarrow T_i} \quad (11)$$

In this way, the flash (*i.e.*  $S_{A \rightarrow W \rightarrow T_i}$ ) is removed in the above equation. Hence, the flash is eliminated by the end of this step. Notice if no object moves, *i.e.*  $S_{A \rightarrow X \rightarrow T_i} = 0$ , then the signal difference approximately equals 0. Thus, we can also use the two signals' difference to determine whether moving object appears or disappears in the room, introducing the following hypothesis.

**HYPOTHESIS 3.** *In the case of absence (or presence) of moving object, the tags will suggest a signal difference lower (or higher) than a threshold. Thus, if the tags suggest the signal difference is lower (or higher) than a threshold, it indicates the absence (or presence) of moving object.*

An experiment is conducted to verify this hypothesis. We keep the room empty and then let a person move into it. Fig. 7 shows that the signal difference firstly maintains in a very low level ( $< 1.0$ ), and then sharply increase by  $4\times$  due to the person's reflections. The difference returns back a normal level after the person leaves. Notice that there may occur a small number of discrete false positives shown in the figure. To deal with this issue, we borrow the idea



**Figure 7: The power of signal difference.** The sampled results are shown every  $200ms$  for visual clarity.

of background subtraction in computer vision [?] which is used for foreground segmentation. The main idea behind is to automatically learn and maintain a representation of the background using Gaussian model. When a new frame incomes, it classifies whether a moving object emerges or leaves. It aggregates the determination results of multiple successive frames to filter the false positives.

### 5.3 Deriving the Reflection

As Eqn. (11) shows, the signal difference implicitly contains the reflections off the moving object (*i.e.*  $S_{A \rightarrow X \rightarrow T_i}$ ). On the other hand, the reflection can be further modeled as follows:

$$S_{A \rightarrow X \rightarrow T_i} = h_{TR_i} h_X h_{A \rightarrow X \rightarrow T_i} h_{AT} S_0 \quad (12)$$

where two new parameters,  $h_X$  and  $h_{A \rightarrow X \rightarrow T_i}$ , are introduced. The  $h_X$  indicates the influence caused by the object, varying over its materials. The  $h_{A \rightarrow X \rightarrow T_i}$  is the compound parameter of channel  $A \rightarrow X \rightarrow T_i$ , which equals

$$h_{A \rightarrow X \rightarrow T_i} = \frac{e^{\mathbf{J}(\theta_{A \rightarrow X} + \theta_{X \rightarrow T_i})}}{(d_{A \rightarrow X} + d_{X \rightarrow T_i})^2} \quad (13)$$

It is the parameter that presents a clue about the moving object's location, which will be employed to track moving object in §6. Next, we concentrate on introducing how to derive  $h_{A \rightarrow X \rightarrow T_i}$  from the signal difference. Substituting Eqn. (12) into Eqn. (11), we have

$$S_{T_i+W+X \rightarrow A} - S_{T_i+W \rightarrow A} = h_{AR} h_{T_i \rightarrow A} h_{TT_i} \underbrace{h_{TR_i} h_X h_{A \rightarrow X \rightarrow T_i} h_{AT} S_0}_{S_{A \rightarrow X \rightarrow T_i}} \quad (14)$$

In order to eliminate the unknown parameters, we divide Eqn. (14) by Eqn. (6) as follows.

$$\frac{S_{T_i+W+X \rightarrow A} - S_{T_i+W \rightarrow A}}{S_{T_i \rightarrow A}} = \frac{h_X h_{A \rightarrow X \rightarrow T_i}}{h_{A \rightarrow T_i}} \quad (15)$$

Interesting, all unknown parameters (*i.e.*  $h_{AR}$ ,  $h_{AT}$ ,  $h_{TT_i}$ ,  $h_{TR_i}$ ) relevant to the devices are removed. More importantly, the transmitting signal  $S_0$  are also removed so our tracking results are irrelevant to the transmitting signal. It introduces two benefits: First, we do not have to know the transmitting continuous wave, which actually depends on the internal implementation of reader and varies over the manufactures. Second, we can design a higher-gain antenna boosting the power (see §7) to increase the SNR of reflections, rather than worry about this behavior's influence on the tracking results.

Consequently, the parameter  $h_{A \rightarrow X \rightarrow T_i}$  can be derived as follows.

$$h_{A \rightarrow X \rightarrow T_i} = \frac{(S_{T_i+W+X \rightarrow A} - S_{T_i+W \rightarrow A})h_{A \rightarrow T_i}}{S_{T_i \rightarrow A}} \times \frac{1}{h_X} \quad (16)$$

In this equation,  $S_{T_i+W+X \rightarrow A}$ ,  $S_{T_i+W \rightarrow A}$  and  $S_{T_i \rightarrow A}$  can be estimated using the reported magnitude and phase values. Meanwhile,  $h_{A \rightarrow T_i}$  is the parameter of ideal forward channel, which can be calculated using Eqn. (3). Note that all of these values are complex numbers. Let  $\hat{h}_{A \rightarrow X \rightarrow T_i}$  be the estimated result, namely

$$\begin{aligned} \hat{h}_{A \rightarrow X \rightarrow T_i} &= \frac{(S_{T_i+W+X \rightarrow A} - S_{T_i+W \rightarrow A})h_{A \rightarrow T_i}}{S_{T_i \rightarrow A}} \\ &= \hat{a}_{A \rightarrow X \rightarrow T_i} e^{j\hat{\theta}_{A \rightarrow X \rightarrow T_i}} \end{aligned} \quad (17)$$

and  $h_X = a_X e^{j\theta_X}$ . Then  $h_{A \rightarrow X \rightarrow T_i}$  can be expressed as follows.

$$h_{A \rightarrow X \rightarrow T_i} \approx \frac{\hat{a}_{A \rightarrow X \rightarrow T_i}}{a_X} e^{j(\hat{\theta}_{A \rightarrow X \rightarrow T_i} - \theta_X)} \quad (18)$$

Unfortunately,  $h_X$  is still an unknown variable that cannot be removed in practice. We will discuss  $h_X$  in §6.

## 5.4 Putting Things Together

The discussion so far focuses on how to extract the reflections off the moving object. Now we put all the pieces together and sketch the algorithm as follows.

*Step 1:* Tadar learns the basic backscatter signal  $S_{T_i \rightarrow A}$  through LOS propagation for each tag before attaching tags on the wall.

*Step 2:* Tadar learns the backscatter signal  $S_{T_i+W \rightarrow A}$  that involves flash after attaching tags on the wall.

*Step 3:* Tadar measures the backscatter signal  $S_{T_i+W+X}[t]$  at read frame  $t$ . It starts to monitor moving object by detecting the signal difference  $|S_{T_i+W+X}[t] - S_{T_i+W}|$ . If new object emerges, it estimates  $\hat{h}_{A \rightarrow X \rightarrow T_i}$  using Eqn. (17) and delivers it to the next component (see §6).

The virtue of Tadar lies in its graceful and compact calculation only using the results reported by a COTS reader collected in three steps, despite the involvement of very complex mathematical derivation, physical-layer information and various unknown variables.

Another thing worth-noting is that RFID reader adopts TDMA fashion to successively identify each tag to avoid tag collisions. The collided signals would be discarded. Since our algorithm performs in the application layer of reader, Tadar does not need to worry about the mutual interferences among tags. The question is how can we guarantee the object stays in the same position when reading these  $n$  tags? The modern COTS, e.g. ImpibJ R420, uses a variety of high performance features making it to read more than 1000+ tags per second [?]. Then we can perform 40 ~ 50 frames per second. Such high-speed frame rate allows us to monitor the region very fast. It is similar to continuously capture the snapshot of surveillance region using a high-speed camera. Therefore, we have a good reason to assume that the object approximately stays in the same position during a read frame.

## 6. TRACKING MOBILE OBJECT

There are two main tracking methods employed in most of prior through-wall systems or tracking systems. The first method steers an antenna array's beam to determine the direction of maximum energy. The direction corresponds to the signal's spatial angle of arrival (AoA) [?, ?]. The second method estimates the time of flight by transmitting a specialized signal (e.g. pulse or FMCW)

and measuring the delay between the transmitted pulse and its received echo [?]. Both of them require the customized device or specialized signals to infer the object's location, which cannot be implemented by the ability-limited RFID tags. Alternatively, Tadar adopts a Hidden Markov Model (HMM) to track the mobile object. In this section, we firstly review the HMM and Viterbi decoding, and then discuss how to apply it into in tracking problem.

### 6.1 Hidden Markov Model

Consider a HMM whose latent Markovian state  $X$  is in one of  $N$  discrete states  $\{x_1, x_2, \dots, x_N\}$ . Let the actual state at time  $t$  be denoted by  $X[t]$ . The transition matrix

$$\mathcal{A} = \{x_i \mapsto x_j : i, j = 1, 2, \dots, N\} \quad (19)$$

defines the state transition probabilities where  $x_i \mapsto x_j = \Pr(X[t+1] = x_j | X[t] = x_i)$ . The Markov chain is assumed to be stationary, so  $x_i \mapsto x_j$  is independent of  $t$ . Let the discrete observation space be set  $\{y_1, y_2, \dots, y_M\}$  with the cardinality of  $M$ . Let  $Y[t]$  be the observation symbol at time  $t$ . The observation matrix

$$\mathcal{B} = \{x_i \mapsto y_k : i = 1, 2, \dots, N; k = 1, 2, \dots, M\} \quad (20)$$

defines the emission probability where  $x_i \mapsto y_k = \Pr(Y[t] = y_k | X[t] = x_i)$ . The initial state distribution is given by

$$\Pi = \{\alpha_1, \alpha_2, \dots, \alpha_M\} \quad (21)$$

where  $\alpha_i = \Pr(X[0] = x_i)$ . The HMM is used to solve three kinds of problems. Here we concern the most famous one that given HMM models  $\{\mathcal{A}, \mathcal{B}, \Pi\}$  and a sequence of observations,  $\{Y[1] \rightarrow Y[2] \rightarrow \dots \rightarrow Y[t]\}$ , to find an optimal state sequence for the underlying Markov process. This problem can be resolved by the Viterbi decoding.

### 6.2 Tracking Object via HMM

Inspired by HMM, Tadar treats the object's location as the hidden states and the reflections off the moving object as the observations. Then, our tracking problem is transformed to finding the maximum likelihood sequence of positions it traversed given a sequence of reflections. The crucial task is how to represent our problem in form of HMM (i.e.  $(\mathcal{A}, \mathcal{B}, \Pi)$ ). Below we describe the modeling process.

**Hidden State and Transition Probability.** Tadar divides the whole surveillance room into  $N$  grids at  $cm$ -level. The hidden states are these  $N$  divided grids. We assume that the object's movement during two consecutive frames is within a wavelength (i.e. its speed  $\leq 34m/s$ ). Suppose the object locates at  $x_i$  currently, it may stay at  $x_i$  or move to next grid  $x_j$  with the constraint of  $d_{x_i \rightarrow x_j} \leq \lambda$ . We do not make any assumptions about the motion behavior, and use uniform transition probability between adjacent steps.

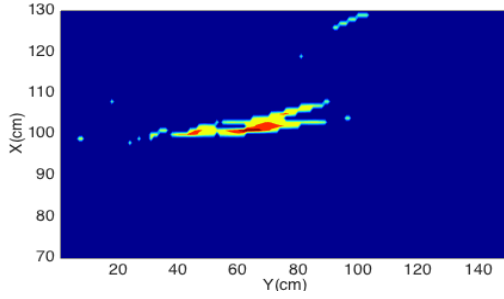
**DEFINITION 1 (TRANSITION PROBABILITY).** Given state  $x_i$  and  $x_j$ , the transition probability  $x_i \mapsto x_j$  is defined as follows.

$$x_i \mapsto x_j = \begin{cases} 1/|\{x_k : d_{x_i \rightarrow x_k} \leq \lambda\}| & \text{if } d_{x_i \rightarrow x_j} \leq \lambda. \\ 0 & \text{otherwise.} \end{cases}$$

where  $k = \{1, \dots, N\}$  and  $|\cdot|$  is the cardinality. The grids in  $\{x_k : d_{x_i \rightarrow x_k} \leq \lambda\}$  are called as the  $x_i$ 's neighbors, which are reachable within  $\lambda$  from  $x_i$ .

Two points are worth-noting. First,  $x_i$ 's neighbors include itself because the object may remain at grid  $x_i$ . Second, the grids near to





**Figure 8: The distribution of emission probability.** The object locates at grid (70, 100).

that in edges of surveillance region have less neighbors than those grids in the center.

**Initial State Distribution.** We also do not make any assumption about the object's initial location (*i.e.* it may emerge at anywhere), so  $\alpha_i = 1/N$ . In practice, if the door is the only way that the person enters the room, we should adjust the grids near to the door with higher probabilities.

**Observation and Emission Probability.** The transmitting signals are scattered to one or more paths after reflected due to the none uniformities of object's surface. Thus, the reflections off the object can be received by multiple tags in the array. On the frame  $t$  incomes, the observation  $Y[t]$  is a vector including  $n$  reflections from  $n$  tags. The observation  $Y[t]$  can be formalized as follows.

$$Y[t] = [h_{A \rightarrow X[t] \rightarrow T_1}, \dots, h_{A \rightarrow X[t] \rightarrow T_n}] \quad (22)$$

where  $h_{A \rightarrow X[t] \rightarrow T_i}$  is the reflection extracted from tag  $T_i$ 's backscatter signal. The emission probability for a given state presents the probability of seeing the object locates at the grid conditioned on the reflections. Correctly modeling the emission probability is a crucial task for trajectory estimation. We adopt the phase value as the metric to depict emission probability. Due to thermal noise, the measured phase value  $\tilde{\theta}_{A \rightarrow X[t] \rightarrow T_i}$  is a random variable following a Gaussian model. Namely,

$$\tilde{\theta}_{A \rightarrow X[t] \rightarrow T_i} \sim \mathcal{N}(\theta_{A \rightarrow X[t] \rightarrow T_i}, \sigma)$$

where  $\theta_{A \rightarrow X[t] \rightarrow T_i}$  is the expectation (*i.e.*  $E(\tilde{\theta}_{A \rightarrow X[t] \rightarrow T_i}) = \theta_{A \rightarrow X[t] \rightarrow T_i}$ ), and  $\sigma$  is the standard deviation. The object is supposed to located at  $X[t]$  so the expectation in theory should equal to

$$E(\tilde{\theta}_{A \rightarrow X[t] \rightarrow T_i}) = \theta_{A \rightarrow X[t]} + \theta_{X[t] \rightarrow T_i} \quad (23)$$

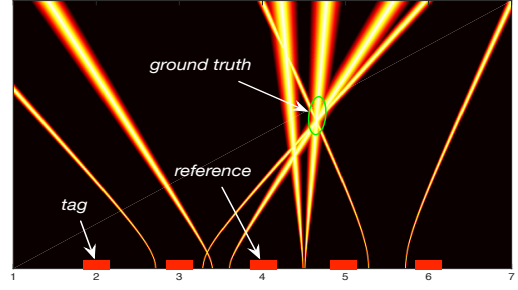
Unfortunately, we can only estimate  $\hat{\theta}_{A \rightarrow X[t] \rightarrow T_i}$  rather than  $\tilde{\theta}_{A \rightarrow X[t] \rightarrow T_i}$  in practice due to the unknown variable  $\theta_X$ , as shown in Eqn. (18). Their relationships are shown as follows:

$$\tilde{\theta}_{A \rightarrow X[t] \rightarrow T_i} = \hat{\theta}_{A \rightarrow X[t] \rightarrow T_i} - \theta_X \quad (24)$$

The fact is that  $\theta_X$  is common term over different tags (because all reflections come from the same object). We select one of them (*e.g.*  $T_j$ ) as a reference and compute the phase difference, denoted as  $\Delta\tilde{\theta}_{i,j}$ , as follows.

$$\begin{aligned} \Delta\tilde{\theta}_{i,j} &= \tilde{\theta}_{A \rightarrow X[t] \rightarrow T_i} - \tilde{\theta}_{A \rightarrow X[t] \rightarrow T_j} \\ &= (\hat{\theta}_{A \rightarrow X[t] \rightarrow T_i} - \theta_X) - (\hat{\theta}_{A \rightarrow X[t] \rightarrow T_j} - \theta_X) \\ &= \hat{\theta}_{A \rightarrow X[t] \rightarrow T_i} - \hat{\theta}_{A \rightarrow X[t] \rightarrow T_j} \end{aligned} \quad (25)$$

Obviously, the unknown variable  $\theta_X$  is removed from the above equation, so phase difference  $\Delta\tilde{\theta}_{i,j}$  can be estimated using Eqn. (17).



**Figure 9: Antenna array beams.** Each pair of tags produces a hyperbola.

Since the mean of the difference of two random variables equals the difference of their expectations, the expectation of  $E(\Delta\tilde{\theta}_{i,j})$  is given by

$$\begin{aligned} E(\Delta\tilde{\theta}_{i,j}) &= E(\tilde{\theta}_{A \rightarrow X[t] \rightarrow T_i}) - E(\tilde{\theta}_{A \rightarrow X[t] \rightarrow T_j}) \\ &= (\theta_{A \rightarrow X[t]} + \theta_{X[t] \rightarrow T_i}) - (\theta_{A \rightarrow X[t]} + \theta_{X[t] \rightarrow T_j}) \\ &= \theta_{X[t] \rightarrow T_i} - \theta_{X[t] \rightarrow T_j} \end{aligned} \quad (26)$$

The variance of the difference of two random variables is the sum of their variances. Thus,  $\text{Var}(\Delta\tilde{\theta}_{i,j}) = \sqrt{2}\sigma$ . The standard deviation is irrelevant to object's location but depends on the reader's transceiver. It can be trained in advance. Totally, we have

$$\Delta\tilde{\theta}_{i,j} \sim \mathcal{N}(E(\Delta\tilde{\theta}_{i,j}), \sqrt{2}\sigma) \quad (27)$$

Then the emission probability is defined as follows.

**DEFINITION 2 (EMISSION PROBABILITY).** *The emission probability governs the distribution of the observations when the object locates at  $X[t]$ , which is defined as follows:*

$$\begin{aligned} X[t] \mapsto Y[t] &= \Pr([h_{A \rightarrow X[t] \rightarrow T_1}, \dots, h_{A \rightarrow X[t] \rightarrow T_N}] | X[t]) \\ &= \frac{1}{N} \sum_{i=1}^N F(\Delta\tilde{\theta}_{i,j}; E(\Delta\tilde{\theta}_{i,j}), \sqrt{2}\sigma) \end{aligned}$$

where  $F(X; \mu, \sigma)$  is the cumulative probability function of Gaussian distribution  $\mathcal{N}(\mu, \sigma)$ .

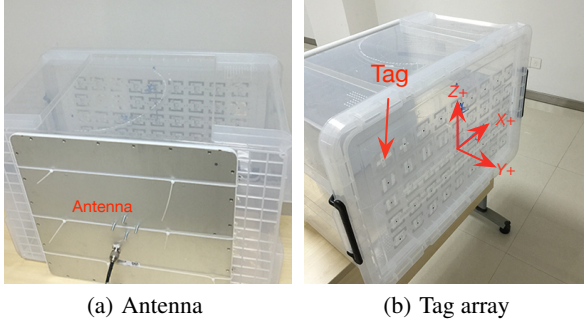
One important reason using HMM and Viterbi decoding for our tracking is that they can reduce the computations for each step considering the principle of moving continuity in time-domain. It is not necessary to search the whole surveillance region every read frame. Fig. 8 shows the distribution of emission probability from one of our experiments where the tag array is deployed in the negative direction of Y-axis. We observe that the the emission probability achieves the highest probability at the ground truth and diffuses to its neighbors as expected.

### 6.3 Further Discussion

It is known that the phase value repeats every a wavelength. How can we ensure the  $X[t] \mapsto Y[t]$  archives the maximum when visiting the correct grid  $X[t]$  given the observation  $Y[t]$ ? Suppose  $\theta_i$  and  $\theta_j$ ,  $\theta_i, \theta_j \in [0, 2\pi]$ , are the phases of reflections off object  $X$  received by tag  $T_i$  and  $T_j$ . The distances and the received signal phases have the following relations due to the phase rotation:

$$\begin{cases} \theta_i = \frac{2\pi}{\lambda} d_{A \rightarrow X \rightarrow T_i} \bmod 2\pi \\ \theta_j = \frac{2\pi}{\lambda} d_{A \rightarrow X \rightarrow T_j} \bmod 2\pi \end{cases} \quad (28)$$





**Figure 10: Tadar's setup.** A plastic box is used to facilitate our tests. (a) An antenna is fixed on the bottom. (b) A group of tags are attached on its top.

Hence, the phase difference,  $\Delta\theta_{i,j} = \theta_i - \theta_j$ , between the received signals at the two tags relates to the difference in their distances from the object,  $\Delta d_{i,j} = d_{A \rightarrow X \rightarrow T_i} - d_{A \rightarrow X \rightarrow T_j}$ , as follows.

$$\Delta d_{i,j} = \lambda \left( \frac{\Delta\theta_{i,j}}{2\pi} + k \right) \quad (29)$$

where  $k$  can be any integer in  $[-\frac{D}{\lambda} - \frac{\Delta\theta_{i,j}}{2\pi}, -\frac{D}{\lambda} + \frac{\Delta\theta_{i,j}}{2\pi}]$  and  $D$  is the distance between two tags. In our scenario, the  $k = 0$  or 1 because all tags are deployed within one wavelength. Thus, the phase difference  $\Delta\theta_{i,j}$  used in emission probability is equivalent to the distance difference  $\Delta d_{i,j}$  in essence.

$$\begin{aligned} \Delta d_{i,j} &= d_{A \rightarrow X \rightarrow T_i} - d_{A \rightarrow X \rightarrow T_j} \\ &= (d_{A \rightarrow X} + d_{X \rightarrow T_i}) - (d_{A \rightarrow X} + d_{X \rightarrow T_j}) \\ &= d_{X \rightarrow T_i} - d_{X \rightarrow T_j} \end{aligned}$$

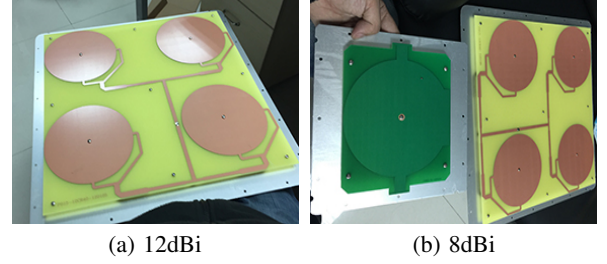
The above equation is a standard equation representing one hyperbola that focuses on  $T_i$  and  $T_j$ . Thus, the emission probability actually indicates how the assumed location (*i.e.* the hidden state) matches the  $n - 1$  hyperbolas. The phase ambiguity is removed by multiple hyperbolas. Fig. 9 shows the 4 hyperbolas including 8 branches produced by the 4 phase differences where the third tag is chosen as the reference. All these curves intersect on the ground truth.

## 7. IMPLEMENTATION

We build a prototype of Tadar using the COTS RFID reader and tags. Fig. 10 shows the main setup of our devices. We use a plastic box to maintain the geometric relationships between reader antenna and tags instead of attaching them on wall for convenience, because we would move the array to different buildings with a same setting. In details, we deploy the reader antenna on the bottom and attach  $5 \times 9$  tags on the top. The box's dimension is  $61 \times 43 \times 33 \text{ cm}^3$ . Tags are separated by  $5 \text{ cm}$ .

**RFID:** We adopt one ImpinJ Speedway modeled R420 reader [?] without any hardware or firmware modification. The reader is compatible with EPC Gen2 standard [?]. The whole RFID system operates in the  $920 \sim 926 \text{ MHz}$  band. We use the Alien tags modeled “ $2 \times 2$ ” ( $2.25 \times 2.25 \text{ cm}^2$ ) [?], which are widely used in supply chain applications. Each of them costs 5 cents.

**Boosting power:** The reflections off the wall are  $30 \sim 40 \text{ dB}$  above the that off moving objects. Even eliminating the flash effect, we can hardly discern the reflections due to human because it is too weak and immersed in the receiver's hardware noise. To boost the



**Figure 11: Antenna design.** We design a high-gain directional antenna with 12dBi to boost power. (a) The 12dBi antenna with four patches. (b) The comparisons of our antenna with 8dBi antenna employed in logistics.

power, we build an antenna with 4 patches, as shown in Fig. 11(a). The antenna has a directional gain of 12 dBi, a 3 dB beam width of  $40^\circ$ . Fig. 11(b) shows the comparisons between the antenna with the ordinary antenna with 8 dBi. Our antenna is composed of a couple of two-layer boards (boards with two copper layers and one substrate layer) that are physically connected to each other by plastic screws.

**Experiment setup:** Majority of our experiments are performed at our conference rooms in one office building. The conference room is  $7 \times 4 \text{ m}^2$ . There are one big table ( $4 \times 0.75 \text{ m}^2$ ), twelve chairs, one white board, *etc.* The interior walls of the building are 5" hollow walls supported by steel frames with sheet rock on top. To discuss the effect of building materials, we also conduct some experiments in a second building with 8" concrete walls. To seize the ground truth, we command an iRobot Create robot with a 12" Lenovo notebook as the reflector to move along a programed trajectory. We also invite 5 human subjects with different heights and builds for the application study. Each experiment is repeated for many times and the average value is reported only. The coordinate system is shown in Fig. 10(b) that the origin locates at the center of the tag array and the signal is emitted to positive direction of Y-axis.

## 8. EVALUATION

In this section, we evaluate Tadar from aspects of reflection acquisition, diversity elimination and object tracking.

### 8.1 Reflection Acquisition

First, we would like to automate the thresholds for the presence of moving objects. To this end, we investigate the SNRs in different scenarios. The SNR is defined as power ratio of signals acquired in presence to absence of moving objects in the room. The reflections with very weak SNR will be overwhelmed by the noise. In this experiment, we place the object at a distance of  $3 \text{ m}$  away from the wall. We choose three typical kinds of materials: 0.98" glass, 1.5" wood door and 5" hollow wall to measure their SNRs. The results in form of CDFs are plotted in Fig. 12. Each CDF contains more than 2000 samples with presence of moving objects. The average SNRs detected are 10.61, 4.69 and 7.01. As expected, the thicker and denser the obstructing material, the harder it is for Tadar to capture reflections from behind it. However, the power level is  $4 \times$  higher than that of noise even in the worst case.

Second, we evaluate Tadar's detection range considering four different building materials including 5" hollow wall (C2), 1.5" wood door (C3), 0.98" glass (C4), and 8" concrete (C5). We

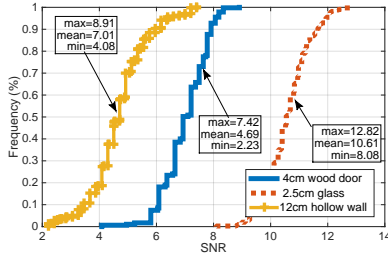


Figure 12: SNR Comparisons

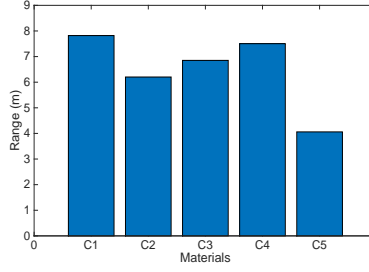


Figure 13: Detecting Range

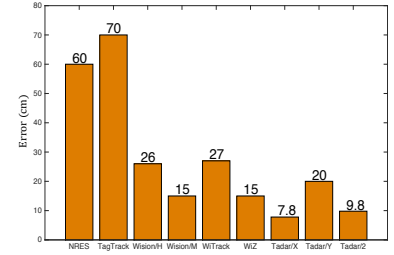
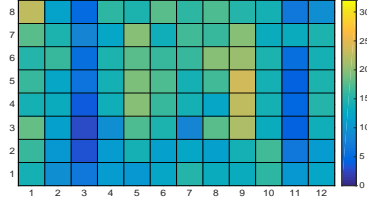
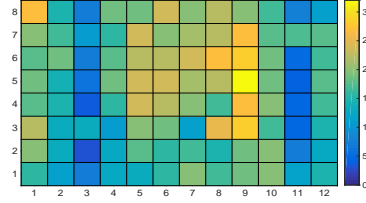


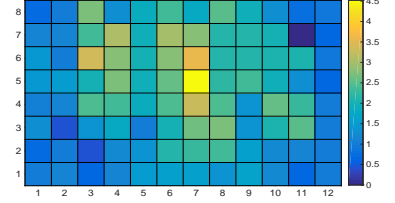
Figure 14: Tracking Accuracy



(a)  $|S_{T_i \rightarrow A}|$



(b)  $|S_{T_i + X \rightarrow A}|$



(c)  $|\hat{h}_{T_i + X \rightarrow A}|$

Figure 15: Magnitude distribution based on tag array. The array includes  $12 \times 8$  tags. Each grid denotes a result collected from the corresponding tag. (a) The signal collected with an empty room. (b) The signal collected when an iPhone 6 Plus is placed in front of the wall. (c) The separated reflections.

also perform experiments in free space (C1) with no obstruction between tag and the object. The range is the maximum distance that Tadar can detect the object. The results are shown in Fig. 13, which are consistent with the comparisons of SNRs. Specifically, Tadar can detect the object with a range of  $7.6m$  in free space, and achieve a median range of  $6m$ ,  $6.77m$ ,  $7.5m$  and  $4.12m$  in other four different buildings materials. RFID system working at  $800 \sim 900$  MHz UHF has a bit weaker penetrability compared with the UWB technologies. After all, RFID system is not designed for see-through-wall originally. However, the benefits of Tadar are obvious in that it is low-cost, compact and without needs of specialized signals or customized devices.

**Summary.** The above two experiments show that Tadar has good behaviors in acquisition of reflections. There are three main reasons for this feature. First, the RFID technology originates from the radar so the RFID devices have the strong ability to resolve the weak signals and remove the self-interference [?]. This feature helps Tadar well suppress the environmental and thermal noise. Second, Tadar utilizes the continuous wave transmitted from the reader to light up the objects. The CW is very stable because it does not carry any information, which facilitates the removal of flash using the learned signals. Third, the tag has no hardware filters for saving cost such that it is hypersensitive to ‘interferences’, enabling us to capture the weak reflections off surrounding objects.

## 8.2 Diversity Elimination

Third, we would like to know whether the unknown parameters related to devices are well removed from the results. To visually understand this issue, let us consider a simple setup with  $12 \times 8$  tags attached on the wall. Fig. 15(a) plots the magnitudes of signals backscattered from these 96 tags with an empty room. A grid in the figure denotes a tag’s result. The antenna is deployed at the center. In the ideal wireless model, the magnitude values would descend from center to edges because the magnitude goes as the inverse

twice power of the distance (see Eqn. 3). However, the distribution is out of order due to the affects from those unknown parameters ( $h_{AT_i}$ ,  $h_{AR}$ ,  $h_{TT_i}$ , and  $h_{TR_i}$ ). Then we place an iPhone 6 Plus in front of the wall with a distance of  $1m$ . Fig. 15(b) again plots the magnitudes collected from those tags. In total, several grids are lightly brightened (thus large RSS values) because their signals are strengthened by the reflections. However, the results are still too ambiguous to pinpoint the iPhone. Fig. 15(c) plots the magnitude of  $\hat{h}_{A \rightarrow X \rightarrow T_i}$ . The result shows quite in agreement with the ideal wireless model. We can even point out the area the phone locates as the tags under the phone have higher values. Specifically, the tag right below the phone has the maximum value. This example demonstrates that  $\hat{h}_{A \rightarrow X \rightarrow T_i}$  indeed well eliminates the impacts from the unknown diversities.

## 8.3 Object Tracking

Fourth, we investigate Tadar’s tracking accuracy compared with other five methods. In the experiments, we let the robot move along a predefined track and then use the HMM and reflections to estimate its trajectory. The accuracy is calculated using the difference of the estimated trajectory and the ground truth. The comparisons are shown in Fig. 14.

**NRES:** NRES [?] deploys a large number of WSN nodes around the room, and tracks humans in building using changes in the signal strength incurred by the human’s motions. It obtains a median tracking error of  $60cm$ .

**TagTrack:** TagTrack [?] is the similar attempt using RFID signals to passively localize the objects. It deploys the tags around the objects and employs the RSS changes as the tracking indicator. It improves the result by introducing classification technique, and achieves a median error distance of  $70cm$ . However, TagTrack is not a see-through-wall technology because the backscatter signals cannot traverse through a wall.

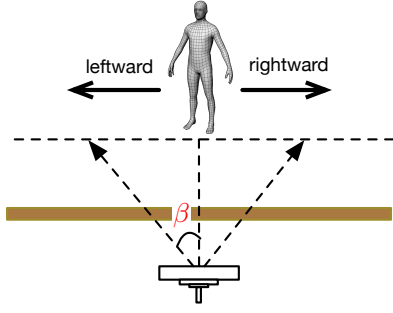


Figure 16: Encoding gesture

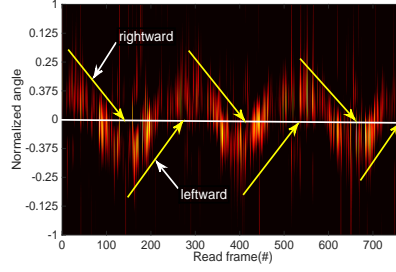


Figure 17: Decoding gesture

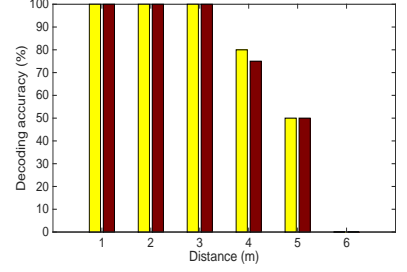


Figure 18: Gesture Accuracy

**Wision:** Wision [?] leverages the multi-path propagations to perform object imaging using two-dimensional antenna arrays. Wision localizes static humans (Wision/H) with a median accuracy of  $26cm$  and metallic objects (Wision/M), like MacBook and HP desktop, with a median of  $15cm$ . Wision suffers from the effects from objects' materials, which is removed in Tadar by comparing with the reference tag. Note that Wision is also not a see-through-wall method.

**WiTrack and WiZ:** WiTrack [?] is a system that localizes the user's motion in three-dimensional space from the radio signals reflected off his/her body. It localizes the center of a human body to within a median of  $27cm$ . WiZ [?] is an incremental work extending WiTrack to multi-person and obtains a median of  $15cm$ . Both WiTrack and WiZ needs a customized hardware to transmit a well-designed radio signal (*e.g.* FMCW) for estimating the traveling time. WiZ generate a signal that sweeps  $5.46 \sim 7.25$  GHz, so both of them, strictly speaking, are UWB based methods.

**Tadar:** Tadar's tracking errors are exhibited in two dimensions. It can achieve a median of tracking error of  $7.8cm$  and  $20cm$  at X and Y dimensions. It outperforms the existing non-UWB technologies by  $3\times$ ,  $3.5\times$  and  $1.3\times$  compared with NRES, TagTrack and Wision, and performs as well as WiTrack and WiZ<sup>3</sup>. Why the results show such a dramatic difference in two dimensions? Reviewing Fig. 9, we see that all hyperbolas of emission probability are towards X-axis because all tags are deployed on one side of the surveillance region only. This incurs a lower discernibility at Y-axis. To coverage the results along Y-axis, we could attach two arrays of tags on two sides of the room. In this way, the combined error can be further reduced to a median of  $9.8cm$ .

**Summary:** Totally, Tadar improves the see-through and tracking technology from three main aspects. First, Tadar well eliminates all uncertain variables (*e.g.* device diversity, reflectivity differences *etc*) in calculations. Second, the low cost of RFID tags allows us to deploy a large number of tags to acquire reflections, which helps improve the accuracy. Specifically, the cost of each antenna, used in Wision and WiTrack, is able to purchase  $40 \sim 50$  RFID tags, while Wision needs to consume 64 antennas (the same price as the sum of 3200 tags). Third, the deployment of Tadar is easier and more compact than other methods, facilitating its application in various domains.

## 8.4 Cost Comparisons

One of the virtues of the Tadar is low cost. Lastly, we investigate the cost comparisons with other through-wall technologies.

<sup>3</sup>There is a setup difference between WiTrack/WiZ and Tadar. WiTrack/WiZ performs the experiments through 6-inch wall while Tadar is with 5-inch wall. This is because thickness standard in China is a little different from that of USA.

**Military radar.** Through-wall imaging and tracking in radar community is very mature. We search the military-level through-wall radars in Alibaba and find the prices of radar devices are ranged between  $80,000 \sim 90,000$  US dollars,  $40\times$  as much as ours on average. Apparently, our solution is very low-cost compared with the military radar. In addition, the through-wall radar systems in the market are usually prohibited from selling to civilians in most countries in the world. Our system targets to civilian purpose with COTS devices, which is totally legal in any country. This benefit cannot be offered by the military radar.

**Wi-Fi based methods.** Wi-Fi based through-wall technologies need a set of software-defined radio like USRP or WARP to emit customized signals [?, ?, ?]. The total price of USRP including antennas is more than 2000 US dollars, a little higher than ours. However, we must claim that our solution totally employs the COTS devices with industry-level design and encapsulation, which can be deployed without any modification on the hardware. The tracking algorithms can run at normal PC computer. With regard to the price and convenience, we insist our solution is prior to the Wi-Fi based.

**WSN based methods.** Some past work in radio tomography employs hundred sensors to track a person passively. No matter the approach is linked-based or location-based, they need deploy a large number of nodes around the room to enable the links cover the entire region. Unlike our approach that deploys the tags in one side, the need cover the room from various angles. So their approach is limited by the room arrangement. With regard to the price, 200 nodes costs about 10,000 US dollars,  $5\times$  as much as ours.

**Summary:** We adopt a powerful reader, ImpinJ R420, to build our prototype. The reader costs about 2,000 US dollars. At first blush, it is not cheap enough. This is because this reader is specially designed for logistics and transportation. There is a very costly industrial computer, serving as a stable HTTP/TCP server, inside the reader to fulfill offline surveillance. In our scenario, we can construct a simpler and cheaper reader only using the reader chip (like Indy RS2000). There is a large space to reduce the system cost. For example, the ThingMagic reader does not embed industrial computer. Its price is around 700 dollars, which is much lower than that of the R420. We will consider the way to further reduce the cost in our future work. Even considering the cost of the current prototype, the low cost is still a major advantage of our solution prior to other methods.

## 9. APPLICATION STUDY

One of the most popular applications in research community recently is to communicate commands or short message to a system by tracking human's movements or gestures, *e.g.* [?, ?, ?, ?]. We also investigate this application using Tadar.

**Gesture encoding and decoding.** Tadar is more sensitive to the X-axis so we employ the rightward (X+) or leftward (X-) movement for gesture encoding. We adopt the Manchester-like encoding as proposed in Wi-Vi [?]: a ‘0’ bit is a step leftward followed by a step rightward; a ‘1’ bit is a step rightward followed by a step leftward, as shown in Fig. 16. This modulation is very simple so that a human finds it easy to perform them and compose them. To support such gesture encoding, Tadar firstly locates the human and then calculates the angle  $\beta$ , defined as the angle related to the vertical line. For example, Fig. 17 shows a sequence of angles detected when a subject repeatedly walks rightward and leftward. The angle equals zero when he locates at the vertical line, and becomes negative when across the vertical line to right. The main reason that we use angle for encoding instead of his absolute positions is that the human may be not know where he is. The input of this application is the angle sequence. Tadar applies two matched filters: one for the step leftward and one for the step rightward. Note that the human can change his direction at anywhere without need of coming back to the vertical line.

**Evaluation.** We evaluate Tadar’s ability to decode the bits associated with the gesture. All the experiments are conducted in our conference room. In each experiment, we ask a volunteer to randomly stand at the conference room, and perform the two gestures. He walks in a way that he finds comfortable. Total 5 subjects take part in experiments. The experiments are repeated with various range from 1  $\sim$  6m. Fig. 18 plots the fraction of experiments that the gestures are exactly decoded. We observe that Tadar can successfully decode the gestures at all distances within 3m. The percentage falls to 75% when the distance increases to the 4m. The decoding accuracy of Wi-Vi [?] is better than that of Tadar when the distance is longer than 4m. However, Wi-Vi cannot exactly pinpoint his location.

## 10. LIMITATIONS

Tadar marks an important contribution by enabling device-free mobile object tracking. It, however, still has some limitations that are left for future work.

**Direction-dependent:** Both the RFID tags and reader antenna are directional, so the whole system is also direction-dependent. Current prototype distributes tags in a fairly large area to enable them to capture different perspectives of the mobile objects in the scene. Our future research may advance at both hardware and algorithmic to reduce the dependence of direction allowing for the tags to be stacked within a smaller area.

**Learning environment:** A second limitation stems from the fact the Tadar needs to train the channel parameters of empty room in

advance to eliminate the flash effects. Not all situations, especially in the case of emergency, allow the user to learn the environment ahead. Our future research may address this issue by detecting the differences of two consecutive frames.

**Tracking one object:** Our algorithm design presented in this paper can track only one mobile object at any point in time. When there are multiple objects present in the scene, the reflections off multiple objects would overlapped together. This limitation is not fundamental to the design of Tadar and can be addressed as the research evolves. For example, considering the case of  $K$  moving objects, we can model the HMM state using a  $K$ -dimensional vector. Each element in the vector indicates each object’s location. Correspondingly, the observations are modeled as the combined signal of  $K$  reflections. The optimized state conforming to all observations is the current objects’ location distributions.

**Beyond FCC power limits:** In order to discern the weak reflections off human, we boost the power using a 12 dBi antenna. Considering 30 dBm transmission power, it would result in 42 dBm EIRP (Equivalent Isotropically Radiated Power) which is 6 dBm more than the FCC limit. To conform the FCC limit, Tadar would be confined to detect the objects with high-reflective metallic surface.

Although there is scope for many improvements, we believe that Tadar inspires us from another different perspective to advance the through-wall technology and other various aspects [?, ?].

## 11. CONCLUSION

In this work, we present Tadar for tracking moving objects without instrumenting them using COTS RFID readers and tags. It works even through walls and behind closed rooms. A key innovation is to logically transform a group of tags into an antenna array receiving the reflections from surrounding objects. We believe that it will open up a wide range of exciting opportunities bridging our physical world and the cyber world.

## ACKNOWLEDGMENT

The research is supported by NSF China Major Program 61190110. The research of Lei Yang is supported by China Postdoctoral Science Foundation funded project under NO. 2015M570100. The research of Xiangyang Li is supported by NSF ECCS-1247944, NSF ECCS-1343306, NSF CMMI 1436786, National Natural Science Foundation of China under Grant No. 61170216 and No. 61228202. We thank all the reviewers and shepherds for their valuable comments and helpful suggestions.