# RFGo: A Seamless Self-checkout System for Apparel Stores Using RFID

Carlos Bocanegra
Northeastern University
Boston, MA, USA

Mohammad A. (Amir) Khojastepour
NEC Laboratories America
Princeton, NJ, USA

Mustafa Y. Arslan
NEC Laboratories America
Princeton, NJ, USA

Eugene Chai
NEC Laboratories America
Princeton, NJ, USA

Sampath Rangarajan
NEC Laboratories America
Princeton, NJ, USA

Kaushik R. Chowdhury
Northeastern University
Boston, MA, USA

## ABSTRACT

Retailers are aiming to enhance customer experience by automating the checkout process. The key impediment here is the effort to manually align the product barcode with the scanner, requiring sequential handling of items without blocking the line-of-sight of the laser beam. While recent systems such as Amazon Go eliminate human involvement using an extensive array of cameras, we propose a privacy-preserving alternative, RFGo, that identifies products using passive RFID tags. Foregoing continuous monitoring of customers throughout the store, RFGo scans the products in a dedicated checkout area that is large enough for customers to simply walk in and stand until the scan is complete (in two seconds). Achieving such low-latency checkout is not possible with traditional RFID readers, which decode tags using one antenna at a time. To overcome this, RFGo includes a custom-built RFID reader that simultaneously decodes a tag's response from multiple carrier-level synchronized antennas enabling a large set of tag observations in a very short time. RFGo then feeds these observations to a neural network that accurately distinguishes the products within the checkout area from those that are outside. We build a prototype of RFGo and evaluate its performance in challenging scenarios. Our experiments show that RFGo is extremely accurate, fast and well-suited for practical deployment in apparel stores.

## CCS CONCEPTS

• **Computer systems organization → Real-time system architecture**; • **Hardware → Emerging technologies**.

**ACM Reference Format:**
Carlos Bocanegra, Mohammad A. (Amir) Khojastepour, Mustafa Y. Arslan, Eugene Chai, Sampath Rangarajan, and Kaushik R. Chowdhury. 2020. RFGo: A Seamless Self-checkout System for Apparel Stores Using RFID. In *The 26th Annual International Conference on Mobile Computing and Networking (MobiCom '20), September 21–25, 2020, London, United Kingdom.* ACM, London, UK, 14 pages. https://doi.org/10.1145/3372224.3419211


*MobiCom '20, September 21–25, 2020, London, United Kingdom*

ACM ISBN 978-1-4503-7085-1/20/09…$15.00
https://doi.org/10.1145/3372224.3419211

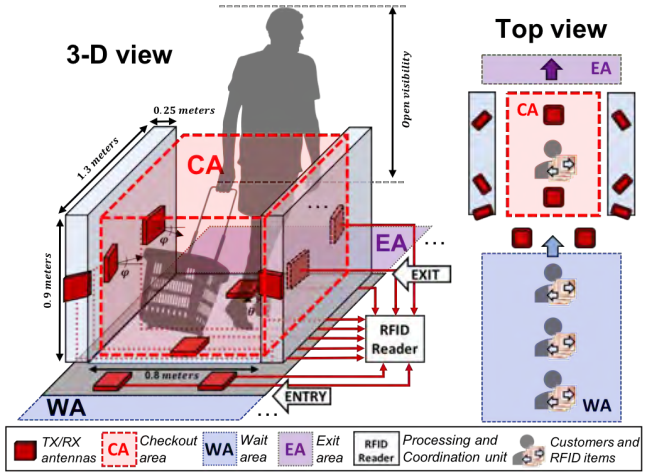



**Figure 1: RFGo's checkout area with RFID antennas**

## 1 INTRODUCTION

Maintaining customer satisfaction, diminishing profit margins and the need for differentiation from competitors have forced retailers to embrace technology to optimize their operations. A major pain point concerns the lengthy checkout process, including both the wait time in the queue as well as the manual process of scanning items in succession [12]. Indeed, this issue dominates the many factors that influence where people choose to shop [10]. We propose RFGo, the first RFID-based seamless self-checkout system for apparel stores that does not require continuous tracking of customers with privacy intrusive cameras, with the potential to impact the multi-billion dollar apparel retail industry.

### 1.1 Problems with State-of-the-art

The problems with the current checkout process stem from inefficient barcode scanning. Since barcodes require line-of-sight (LoS), each item has to be sequentially and manually aligned with the scanner. Recent solutions utilize an extensive camera deployment and various other sensors to eliminate manual product scanning. By using computer vision to continuously monitor the actions of consumers, these sophisticated systems associate items to customers and allow each customer to simply walk out and get billed. As pioneered by Amazon Go, several other companies implement systems centered on the same idea [5]. Using such extensive tracking

with cameras comes at a price, e.g., for Amazon Go, recent third party analysis estimates the need for 30,000 cameras per 100,000 sq. ft. store, and 7,000 GPUs during busy times [32]. Per unit cost, power draw and network planning cost increase quickly with the resulting datacenter-like processing load. In addition, camera-based systems are not well suited to distinguish between apparel items where similar looking items (e.g., two different plain white shirts) may be priced very differently and hence need to be identified as such. Further, continuously monitoring shoppers with an extensive camera network could have privacy implications.

### 1.2 An RFID-based Self-checkout System

RFGo advocates using passive (i.e., battery-less) RFID tags as opposed to identifying products with cameras or barcodes. RFID is prevalent in apparel retail industry, especially for inventory tracking at major brands such as Zara, Uniqlo and Macy's [7]. Efforts such as Japan's Electronic Tag Initiative promote RFID in other retail areas, e.g., convenience stores [29]. Given the decreasing price trend (two cents per tag [11]), RFID is projected to be more widely used in retail.

Even though RFID is poised to become mainstream, there are still technical barriers for it to be widely deployed in some retail sectors. For example, it is well known that RFID tags attached to conductive materials such as metal surfaces (e.g., soda can) or certain liquid containers (e.g., sports drink) may become *inherently unreadable* [3, 33]. Typical apparel items are generally not conductive and hence do not suffer from this problem. For this reason, we focus on the application of RFGo in apparel retail stores. However, we note that the technology behind RFGo is applicable to RFID checkout in general provided that a proper tagging system is in place such that the tags are not inherently unreadable.

Instead of continuously tracking customer activity throughout the store, RFGo identifies purchases in a dedicated area (e.g., right before the store exit) called the checkout area (CA) that is equipped with RFID antennas. The CA has a dedicated entrance and exit. Each customer enters the CA with the products (e.g., in a shopping cart or in hand) and stands briefly inside the CA while other customers with unchecked items line up behind the entrance in the waiting area (WA) as in Fig. 1. While the customer stands in the CA, RFGo scans RFID-tagged products, without requiring the customer to take them out (e.g., if in a bag) and align them with a specific antenna. When signaled by RFGo, the customer leaves the CA to the exit area where no unchecked item is allowed.

RFGo's goal is to *completely* and *solely* identify the products in the CA, meaning no product is missed from the current customer's basket or extra products are identified (e.g., from other's baskets in the WA). We define RFGo's accuracy based on *recall* and *precision* [36]. Recall is the percentage of the products in the customer's basket that RFGo correctly identifies to be in the CA (ideal is 100%). Precision is the percentage of correctly identified products among those identified by RFGo (again ideal is 100%), which may include those inside the CA (correct detection) as well as outside the CA (wrong detection). To avoid making customers stand indefinitely in the CA and have long lines in the WA, RFGo also needs to identify products quickly (e.g., within two seconds).

One may think that RFID tag localization [8, 24, 25, 31, 35, 41–43] may precisely identify the products in the CA. Unfortunately, recent studies show that localization accuracy is impacted by "even minor variations in the environment or tag geometry" [33]. Since customers and products cannot remain absolutely still in and around the CA, high localization accuracy is not realistic in a typical retail store.

RFGo has to address several challenges such as not being able to (1) energize a tag or decode its weak reply, (2) decode a tag reply when multiple replies collide and (3) precisely localize a tag as discussed above. We expand on these challenges in Sec. 3.

### 1.3 Contributions

Our main contribution is the design and implementation of RFGo, a seamless, fast and highly accurate self-checkout system based on RFID. RFGo comprises three components that work in harmony.

(1) The physical structure resembling a checkout lane (defined by two short RF-absorbing walls) accommodates a shopper and her items. It hosts multiple antennas placed in different positions and with different orientations to reliably read tags regardless of their orientation. We also use a dual-area antenna deployment to cover the area within and outside the CA (e.g., the WA in Fig. 1). As we will show later, the dual-area deployment is critical to reach the desired checkout accuracy.
(2) The custom-built RFID reader, implemented in GNURadio and compliant with RFID Gen2 standard, addresses the shortcomings of traditional RFID readers to enable fast checkout. It extracts multiple observations from a single tag's reply by *simultaneously* decoding the reply from multiple antennas that are *synchronized* at the carrier-level. In contrast, traditional RFID readers decode tag replies using one antenna at a time, which requires significantly longer time to capture the same number of observations than our reader. Our reader also implements an intelligent acknowledgment policy (see Sec. 4) to improve the probability of decoding in the absence of standard error-detection such as cyclic redundancy check (CRC).
(3) The software classifier decides if a tag is inside the CA based on the observations from our reader (without explicit tag localization). We train a neural network-based classifier that extracts features from the observations and combines them appropriately using supervised learning.

The rest of the paper is organized as follows. In Sec. 2, we introduce our design principles. Sec. 3 describes the main challenges that RFGo needs to address. Sec. 4 first provides an overview of RFGo's components and discusses how they address the challenges in Sec. 3. Since the custom-built reader is key to achieve extremely fast checkout, the rest of Sec. 4 is devoted to the design of our reader. We describe our implementation in Sec. 5 and evaluate RFGo in Sec. 6. We discuss related work in Sec. 7. Finally, we conclude in Sec. 8.

## 2 DESIGN TARGETS AND PRINCIPLES

In this section, we review current RFID-based self-checkout systems and describe our vision for RFGo.

### 2.1 Current RFID Self-checkout Systems

RFID is widely used for taking inventory where tags are scanned by handheld RFID readers (Fig. 2a). Using handheld readers for

(a) Handheld (b) Cage-based (c) Slot-based

(d) Bin-based (e) Surface-based

**Figure 2: Current RFID-based Self-checkout Systems**

self-checkout leaves the responsibility of scanning the products correctly (i.e., completely and solely) to the customer. Recently, several systems that avoid handheld scanning have been proposed. One example is a cage-based system where the customer loads items into a shielded Faraday cage equipped with antennas and closes the door (Fig. 2b). The system then scans the items in the cage using high transmit power without the worry of erroneous reads from outside the cage. Other systems address unwanted reads by limiting the range of their antennas: examples include slot-based systems (Fig. 2c) where the products are passed through a slot, bin-based systems (Fig. 2d) where products are placed in a container, and surface-based systems (Fig. 2e) where products are laid out on a planar structure. Although these systems reduce the likelihood of unwanted reads by controlling the antenna range, they typically suffer from not being able to read all items due to the limited power.

## 2.2 Our Vision

We envision a self-checkout system where the customers simply walk into and stand in the CA while carrying the products they wish to purchase. In particular, our envisioned system should have the following properties.

- **No manual effort:** Two major factors that cause inconvenience with existing systems is the manual loading/unloading of the products (e.g., opening/closing the cage door) and manual scanning (e.g., handheld readers). Thus, the system should not require any extensive human involvement.
- **Large checkout area:** Small checkout areas limit the items that fit in them and are inconvenient for buying many items (e.g., passing multiple shopping bags through the slot is cumbersome). Thus, we envision the CA to be large enough to comfortably accommodate a customer and her products. The CA should also allow some freedom of movement to the customer without forcing them to stand still or in another particular orientation.
- **Unbarricaded checkout area:** In addition to being large, the CA should be unbarricaded e.g., not enclosed by sliding doors or other large physical obstacles. This is desirable for both, aesthetics, i.e., the CA does not abruptly stand out as a major structure in an open-floor store design, and physical comfort, i.e., customers do not feel trapped. To this end, we envision the CA to form a "lane" demarcated by two waist-level side walls without entry or exit barriers.
- **Speed:** In our vision, the time required for (i) the cart identification process, and (ii) the payment authentication should be comparable. Hence, we envision a fast identification process which takes only a few seconds that is typical for credit card, phone-based (QR code, NFC) payment authentication and facial-recognition based payments. Note that since these two tasks are independent, it could be possible for them to be carried out in parallel to further speed up the overall process.

Our vision has several implications on the design of RFGo. Having no manual scanning and a large CA results in considerable distance between the antennas and tags. To ensure that all tags are read within the CA (i.e., there is no blind spot), we need to employ multiple antennas with sufficient power. Although an increased power helps reading the tags, it may cause signal spillover to areas outside the CA resulting in unwanted reads. This problem is exacerbated by our vision for an unbarricaded CA. Consequently, it is quite challenging to determine whether a tag is inside the CA based on a single reading. For this reason, achieving the desired speed is challenging too since RFGo needs to collect enough readings from tags to accurately make a determination.

# 3 CHALLENGES

In this section, we describe in detail the challenges of designing a self-checkout system using RFID.

## 3.1 Blind Spots

We say a stationary tag in a particular position and orientation is in a *blind spot* when it is not readable. Blind spots occur either due to weak illumination, where the tag is not even activated by the reader; or to weak response, where the tag responds but its reply is not decodable by the reader. Major factors impacting blind spots are as follows.

- **Illumination**: For a passive RFID tag, being illuminated refers to the ability to harvest enough energy to wake up and respond to reader commands. The harvested energy is a function of the transmit power of the reader, the distance between the tag and the reader and a few other factors (e.g., temperature) [33, 49]. In addition to waking up, a tag needs to receive reader commands with enough signal-to-noise ratio (SNR) to decode and respond to them.
- **Orientation**: Tags are best read when they are oriented in a way that maximizes the received power at the tag's antenna. For a flat tag, this happens when the signal from the reader is orthogonal to the plane of the tag's antenna. Any deviation from this ideal *orientation* may greatly reduce the signal strength of the tag's response [15, 33, 37].
- **Coupling**: The signals backscattered by other tags also impact a particular tag's readability. For example, a tag that can be detected when it is the only tag in the reader's view may not be readable when other tags are brought in close proximity (without changing

the observed tag's position and orientation) due to the mutual coupling effect [28, 37].
- **Multipath**: A tag may not be illuminated by the reader when it is situated in a particular spot with deep fading, i.e., when multiple reflected copies of a signal destructively combine. Similarly, destructive combining may also happen at the reader antenna for the reflected copies of a tag's response affecting its decodability by the reader.

Note that we define blind spots only from the perspective of the physical layer and exclude MAC layer factors such as collisions. While collisions in the tag responses may eventually be resolved by a random access protocol, a stationary tag in a blind spot may not be read at all. For this reason, mitigating blind spots inside the CA is crucial as we expect 100% recall in the checkout process.

### 3.2 Collision

A response from a tag may not be decodable by the reader when other tag(s) also respond at the same time due to *collision* [1, 2, 4, 6, 13, 14]. Although a random access protocol may resolve collisions if given enough time, a long resolution time is not desirable. Since a typical retail store may have a large population of tags near the CA (e.g., other tags on the shelves or tags carried by other customers in the WA), the tags inside the CA may experience high number of collisions. This reduces the probability of decoding a tag in a given time slot and extends the time required to discover all the tags within the CA [17–23, 27, 30, 38–40, 44–48]. Therefore, mitigating collisions is also important to achieve 100% recall and fast scan time.

### 3.3 Position Uncertainty

RFID Gen2 standard does not have a built-in mechanism for localization of tags. Prior works use multiple antennas to precisely locate tags (within a decimeter error) based on signal phase: by using angle-of-arrival [34], multi-lateration [26] and synthetic aperture [43]. However, recent studies show that phase-based localization is impacted by even minor variations in tag geometry and by changes in the environment (e.g., people moving) [33]. Retail environments impact localization accuracy as explained below.

- **Tag Mobility:** Since phase is a function of the physical distance between a tag and an antenna, customers moving the items in their shopping bag (e.g., swinging back and forth) introduce unpredictable phase variations. Since customers cannot be forced to stand still, it is imperative for the self-checkout system to work despite the presence of such unpredictable movements.
- **Non-stationary environment:** Scattering objects in the environment create multiple copies of the tag signals that combine at the reader antenna. Movement of scatterers changes the combined signal and thus the phase measurement from it. The self-checkout system needs to work within a non-stationary environment such as people passing nearby the CA or moving within the CA.

Without precise localization, addressing position uncertainty is paramount to achieving 100% precision and recall for the self-checkout application.

## 4 DESIGN

In this section, we first describe basic RFID communication followed by RFGo's design overview. We dedicate most of this section to our custom-built reader and detail other components in RFGo in Sec. 5.

### 4.1 Background on RFID Communication

RFID Gen2 communication involves passive (i.e., battery-free) tags and an active reader device. The reader energizes the tags and sends commands to them. After harvesting enough energy, tags wake up to listen for the reader commands and react to them, e.g., by sending a response. In the so-called inventory mode, each tag responds with its unique ID called Electronic Product Code (EPC). A simple inventory session has four phases orchestrated by a Slotted Aloha protocol: 1) **Query:** the reader broadcasts a query and indicates the number of available slots, 2) **RN16:** if a tag decodes the query, it chooses a random slot and later responds with a 16-bit random sequence (called RN16) using FM0 modulation in the selected slot, 3) **ACK:** in each slot, if the reader decodes an RN16, it sends an acknowledgment (ACK) containing the same decoded RN16, and 4) **EPC:** each tag decoding the ACK matches the included RN16 to the RN16 it chose earlier, and replies with its EPC when there is a match. Since tags randomly choose slots to reply, a slot may have zero, one or more RN16s that collide at the reader. In contrast, since it is unlikely that tags generate the same RN16, EPC responses do not usually collide. RN16s do not have built-in error detection, making it difficult for the reader to know whether an RN16 was decoded correctly or not.

### 4.2 RFGo's Design Overview

As described before, RFGo includes three main components.
**(1) The physical structure** clearly demarcates the CA and hosts multiple RFID antennas placed to mitigate blind spots. We place the antennas in three main orientations defined by three planes that are mutually orthogonal (see Fig. 1). We also place the antennas sufficiently apart from each other to form uncorrelated communication channels; e.g., if a tag is in a deep fading spot with respect to one antenna, it may not be for another antenna. By utilizing antenna diversity for transmit (TX) and receive (RX) paths, RFGo reduces the occurrence of blind spots. *Reading* a tag not only obtains its EPC but also reveals important features that may be extracted from its physical signal (e.g., signal strength). In order to collect better observations for the tags within and outside the CA, we employ a dual-area deployment, where two distinct sets of antennas are placed in such a way that the "inside" ("outside") antennas primarily illuminate and read the tags "inside" ("outside") the CA.
**(2) The custom-built reader** leverages antenna diversity for the TX path by employing Time Division Multiple Access (TDMA) where it activates each antenna in turn to address illumination issues. For the RX path, it implements two novel features: (1) *simultaneous* decoding of tag replies from multiple antennas via carrier-level synchronization of the RF chains and (2) an intelligent acknowledgment policy that sends an ACK for the RN16 that is most likely to be correct (detailed later in this section)[1]. Simultaneous decoding not only increases the likelihood that a tag reply will

[1]The astute reader may notice that multiple RFID readers can be used to simultaneously query the tags and receive responses from them on orthogonal channels (i.e., frequency

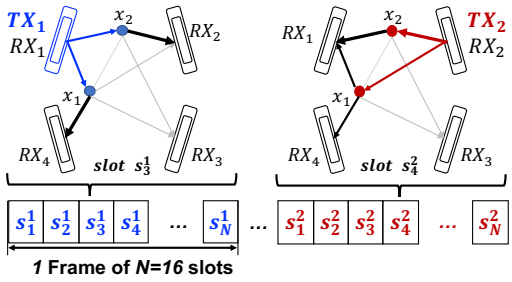



**Figure 3: Exemplary RFGo multi-antenna operation**

be decoded through receive diversity, but it also provides multiple readings (or observations) for the tag (same EPC but different signal features). Recall that collisions only happen during the RN16 phase where tags randomly choose the same time slot. Due to the lack of built-in error-detection for RN16s, the reader needs to somehow choose which RN16 to acknowledge among the colliding ones. By leveraging parallel decoding of the received RN16 signal at multiple antennas and analyzing low-level signal metrics, our reader picks the RN16 with the highest probability of being correctly decoded. If a "poor" RN16 is ACKed in the absence of such an intelligent policy, the subsequent EPC response from the tag would not be decoded and the slot would be wasted.

**(3) The software classifier** detects items that a customer brings to the CA. Since tags cannot be precisely localized in typical retail environments, we rely on a neural network to distinguish the tags in the CA from others outside the CA by learning from the signal features exposed by our reader. Recall that our reader simultaneously decodes the same EPC on all RX antennas (say $N$) for each TX antenna. In contrast, traditional readers have only one RX antenna for each TX antenna. Hence, after $N$ TX antennas are activated, our reader extracts $N^2$ features for each tag from all TX/RX antenna pairs compared to $N$ collected by traditional readers. In addition to the rich feature vector, the dual-area antenna deployment helps us capture observations from both tags inside the CA and the ones in the WA, giving the model even more information from which to learn.

## 4.3 Design of RFGo's Custom-built Reader

As discussed before, we employ multiple TX antennas to mitigate blind spots. Since the TX antennas are activated one by one in TDMA fashion, we connect the TX paths of all antennas to a single TX RF chain through a multiplexer (see Fig. 6). Recall that our reader decodes a tag reply simultaneously from multiple RX antennas; this may seem redundant at a first glance. However, receiving a single reply at different antennas does not mean receiving *the same* reply and has two important benefits. First, even though it has the same EPC, the tag reply yields different observations: the observations include the signal features (e.g., signal strength) from each decoding antenna and the set of antennas which have been able to decode the reply. Second, leveraging the spatial diversity across antennas allows us to better decode a valid RN16 even when collisions happen. This is important due to the lack of error-detection (CRC) for RN16s.

division multiple access or FDMA). However, this is not possible with RFID tags since a concurrent Query command resets any ongoing response on the tag [9].

To enable each antenna both for transmission and reception, we separate its TX and RX paths using a circulator. Hence, tags are illuminated by a single TX antenna at a time (TDMA), while their replies are collected simultaneously through all antennas by using one RF chain per RX antenna (see Fig. 6). To decode a tag's backscattered signal, the coherent detection is far superior than non-coherent detection particularly due to the short communication intervals typical for RFID. Traditional readers implement coherent detection by employing a pair of TX/RX RF chains that share one local oscillator. To use coherent detection for simultaneous reception across *all* antennas, we need to synchronize all the RX RF chains at the carrier-level with the single TX RF chain.

Transmission from an antenna comprises transmit frames where each frame contains several time slots for querying the tags. Fig. 3 shows an example transmission for 2 frames. The slot $i$ of frame $k$ is denoted by $s_i^k$. During the first frame, the first antenna is the active transmitter and triggers responses from tags $x_1$ and $x_2$ in slot $s_3^1$. In the second frame, the second antenna is the active transmitter and triggers responses from the same tags in slot $s_4^2$. Higher decoding probability is achieved for stronger transmission links, denoted by thicker lines in Fig. 3. In time slot $s_3^1$, our reader decodes two RN16s transmitted by $x_1$ and $x_2$ using the received signal at antenna $RX_2$ and $RX_4$, respectively. In contrast, traditional readers use a single receiver, e.g., $RX_1$, while $TX_1$ is the transmitter. In this example, there is a high chance that the received signal from both tags collide at $RX_1$ and neither one is decoded correctly. Moreover, our reader achieves better fairness in terms of selecting tags more uniformly, and discovering the new tags faster by acknowledging one of the decoded RN16s at random, e.g., in two different time slots where the same set of tags reply to the query and the active TX is the same. In contrast, when the same antenna is used both as TX and RX, the tag with stronger received reply (say due to its proximity to the antenna) is always favored where the same set of stationary tags collide in the same time slot.

Hence, leveraging the diversity in RN16 decoding is one of the important aspects of our custom built reader. In the next section, we describe different policies for selecting an RN16 for acknowledgment. We defer detailed evaluation of these policies to Sec. 6.

## 4.4 RN16 Selection Policies

Our reader tries to decode RN16s by collecting the I/Q samples from each of the synchronized RX RF chains in parallel. Since there is no explicit error detection, each RX RF chain cannot determine whether it decoded an RN16 correctly. To decide which RN16 to acknowledge in each slot, we have to rely on the collective information from all RX RF chains. This includes the set of decoded (either correctly or incorrectly) RN16s and their associated signal attributes. We next define three policies for selecting an RN16 for acknowledgment, based on various degrees of the information they use.

(1) Fixed antenna Policy (*FP*): FP relies on information from a single (fixed) RX RF chain to decode an RN16, which is then used by the reader to generate the ACK.
(2) Majority Voting Policy (*MVP*): MVP uses the knowledge of the set of decoded RN16s. MVP selects the RN16 that has the maximum number of detection events among all the ones decoded by different RF chains. If two or more RN16s have been decoded

equal number of times (i.e., majority does not exist), MVP picks one at random.

(3) Interference Metric Based Policy (IMP): IMP uses the signal attributes for the set of decoded RN16s. We compute a specially designed interference metric (IM) for each of the simultaneously received signals. IM provides a measure of interference in the received signal. Thereby, IMP selects the RN16 that has the minimum IM across all decoded RN16s, corresponding to the RN16 with highest decoding probability (i.e., lowest interference). Similar to MVP, IMP randomly selects one RN16 among all RN16s whose IM are below a threshold. We provide the detailed description of IM in Sec. 4.5.

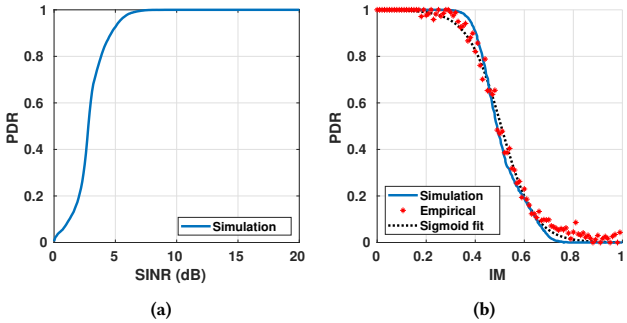



**Figure 4: (a) SINR vs. PDR (b) IM vs. PDR**

## 4.5 Interference Metric

IMP selects one RN16 from the set of RN16s based on their decoding probabilities. The goal of IM is to estimate the probability of correctly decoding an RN16 (i.e., packet decoding rate or PDR). We define IM as the fraction of interference and noise power in the received signal without explicit knowledge of the received signal, the interferers, or their powers. The relation between IM and PDR becomes clear later in this section. Since there is no CRC for an RN16, one has to decide if the decoded RN16 is "correct" or "not correct", which is a binary classification problem. A binary classifier may compare a measure with a threshold to perform the classification. The accuracy of a binary classifier is defined as the ratio of the number of the correct decisions over total number of decisions [36]. The accuracy of a classifier depends on its threshold. Naturally, the *effectiveness* of a measure (in a classification problem) is the maximum accuracy of such binary classifiers taken over all possible thresholds.

A simple and readily available measure is the received signal strength (RSS), which can be used to coarsely estimate PDR. Through a simulation with 32 tags and a frame size of 16 slots, we see that RSS has a low effectiveness of 0.76, which may be explained as follows. Since RSS merely represents the total received energy, it cannot distinguish between a strong signal without interference (which is decodable) vs. the superposition of many low power interferers (which is not decodable), since both cases yield a high RSS.

A more sophisticated measure is signal-to-interference and noise ratio (SINR), which can be obtained if prior knowledge of the signal and the interferers is available. Indeed, SINR has an effectiveness of 0.97 in the same simulation setup since it directly captures the power of the useful signal relative to interference and noise. Unfortunately, SINR is not available for a short packet, such as RN16, with few preamble bits. Hence, we propose IM, which is computed as follows.

Our custom-built reader decodes a single RN16 packet with highest received power among all interfering RN16 signals by treating the interfering RN16s as noise. After equalization and passing the received samples through a matched filter, we reconstruct an estimation of the FM0-modulated RN16 symbols which is then used for decoding. We define a difference sequence by subtracting the even numbered symbols from the preceding odd numbered ones, e.g., (1st symbol - 2nd symbol) followed by (3rd symbol - 4th symbol) and so on. IM is then the ratio of the standard deviation to the mean of the absolute value of the difference sequence. Consequently, the RN16 bits can be retrieved by comparing the absolute value of the difference sequence with a threshold. IM is obtained in an intermediate step of the decoding and hence its computational complexity is negligible. In the same simulation setup, we see that IM achieves an effectiveness of 0.95, very close to that of SINR, without using any prior knowledge.

Next, we present the relationship of PDR vs. both SINR and IM and discuss why they are effective. A measure with 100% effectiveness would have a step-wise behavior. Fig. 4a shows PDR vs. SINR and confirms that SINR's high effectiveness is indeed due to its sharp drop closely resembling a step-wise function. Fig. 4b shows PDR vs. IM and confirms a similar step-wise behavior. However, note that the step direction is reversed since lower interference means lower IM and higher SINR. Fig. 4b also shows that the empirical PDR closely follows the simulation. To find the empirical PDR vs. IM using our reader, we record the value of IM computed in RN16 phase along with the outcome of the CRC check obtained in the EPC phase for each slot. We quantize the IM range into bins. For all the slots in the same IM bin, the PDR is the percentage of correctly decoded EPCs in those slots. The empirical PDR may be modelled as a sigmoid function PDR = $1/(1+exp((\text{IM}-\alpha)/\beta))$ for faster computation, which is obtained in Fig. 4b with $\alpha = 0.5057$ and $\beta = 0.06766$.

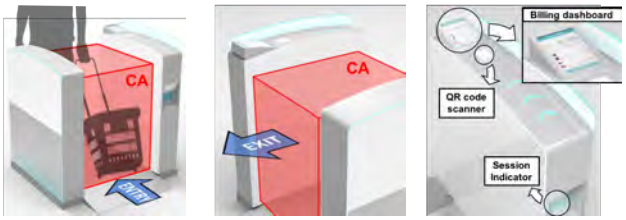
# 5 SYSTEM IMPLEMENTATION

In this section, we detail how we implement each component of RFGo to realize a fully functional self-checkout system.

## 5.1 RFGo - Physical Structure

The physical structure of RFGo's self-checkout lane is defined by the two side walls that are placed sufficiently apart from each other to accommodate a customer along with the items to be purchased (see Fig. 5a). Customers walk in the CA from a defined "entrance" (as indicated by the arrow direction in Fig. 5a), stay inside for a very brief time and leave from the opposite side defined as the "exit" (see Fig. 5b). During this time, RFGo utilizes the custom-built reader to collect tag readings from several RFID antennas and uses the neural network-based classifier to determine which tagged items are in the CA. The sequence of operations including the customer entry, scanning and output of the classification is collectively referred to as a checkout "session."

RFGo employs control sensors and visual indicators to guide the operation of each session. When a customer enters the CA, a pair




(a) (b) (c)

**Figure 5: RFGo's checkout lane showing (a) the entrance, (b) the exit, and (c) the user interface**

of infrared (IR) sensors detect the motion and trigger the custom-built reader to start collecting tag readings for the new session. After analyzing the readings collected within two seconds, the neural network identifies the tagged items in the CA and displays them on the billing terminal (see Fig. 5c). The customer then pays using the QR code scanner and leaves the CA upon seeing the payment confirmation on the billing terminal. Another pair of IR sensors detect the customer's exit, allowing RFGo to accept the next customer and start a new checkout session. The presence/absence of an ongoing session is displayed to an approaching customer by an indicator mounted around the entrance of the lane (see Fig. 5c). If there is an ongoing checkout session, the approaching customer is discouraged from entering the lane by displaying a "busy" sign. Since there are no barricading doors enforcing the boundaries of the CA, the IR sensors are used to detect several undesirable scenarios. Examples include a customer entering the lane during an ongoing session (i.e., ignoring the busy indicator) or a customer leaving the lane too soon (before the scanning is complete).

To read the tags within the CA, we deploy six antennas as follows: (a) two antennas in each side wall facing the inside area which are placed almost at 45° angle to the face of the wall and (b) two antennas under the floor mat placed parallel to the floor facing upward. We also deploy four antennas to read the tags in the WA: (a) one antenna mounted within each side wall at approximately 20° angle facing the WA, and (b) two upward-facing antennas under the floor mat away from the entrance of RFGo (see Fig. 1). The side walls have a sturdy metal frame to minimize signal spillover from the inside antennas. However, RFGo still has to cope with the spillover from the entrance/exit openings between the side walls since having an unbarricaded area was a deliberate design choice.

## 5.2 RFGo - Communications

The block diagram in Fig. 6 details our custom-built reader implementation. The core RFID signal processing is implemented in GNURadio v3.7.13.5 by significantly extending the code in [16]. The system runs on a Linux machine equipped with 12 cores and 128 GB RAM. The radio front-end is emulated using four USRP X310 units which are time-synchronized by means of an Octoclock CDA-2990. Each USRP is equipped with either a UBX or TwinRX daughterboard that allows for communication within the 902-928 MHz UHF RFID band. RFGo is equipped with six antennas in the CA and four antennas in the WA for a total of ten antennas. Each antenna separates the TX and RX signals with a circulator. We use five TwinRX daughterboards to enable ten RX RF chains, where each

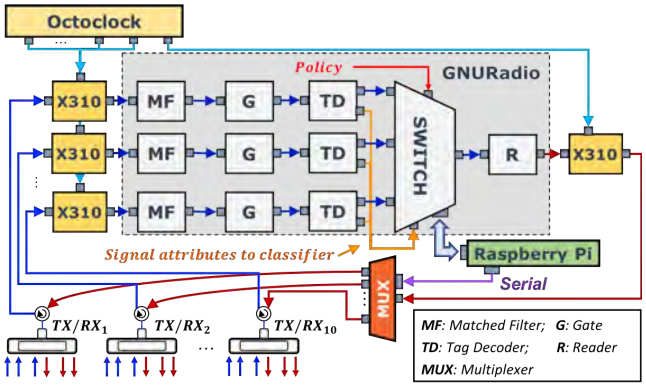



**Figure 6: Implementation of our custom-built reader.**


(a) (b) (c)

**Figure 7: (a) Smartrac tags used in our experiments, (b) dense tag arrangement covering multiple X-Y-Z orientations and (c) unified cube comprising 27 smaller 6" cubes.**

RF chain is connected to one antenna. Since one USRP supports two daughterboards, we use three USRPs to host the five TwinRX daughterboards. The fourth USRP hosts one UBX daughterboard which is the only TX RF chain in the system. The output of the UBX is connected to a set of multiplexers controlled by a Raspberry Pi to select the active TX antenna.

In GNURadio, the received I/Q samples pass through an FIR averaging filter (*MF*) to mitigate the impact of noise. Then, the Gate block (*G*) performs pulse detection to do coarse frame synchronization. The frame is forwarded to the Tag Decoder block (*TD*) which aligns symbols using preamble-based correlation and decodes the RN16 and EPC packets. We extend the *TD* block to compute the IM (see Sec. 4.4) and estimate the RSS. The *TD* blocks forward the decoded RN16s along with the IM to a central block called the *Switch*. The *Switch* block implements the RN16 selection policies (see Sec. 4.4) and forwards the selected RN16 sequence to the reader block (*R*), which generates the ACK based on the forwarded RN16. The *R* block is also responsible for generating the Query packets broadcast to the tags. The *Switch* block also keeps track of current slots to switch to another TX antenna in time. To do so, it establishes a UDP connection with a Raspberry Pi that functions as a micro-controller to activate specific ports of the multiplexers. At the end of a frame, the *Switch* encapsulates the next activated port within the UDP packet and sends it to the micro-controller.

## 5.3 RFGo - Classification Software

Recall that our goal is to correctly determine whether a given tag is inside the CA or not. To do so, we use "soft" features such as RSSI

and number of readings. Although soft features measured from one antenna do not reveal the tag's position, combining several such features from multiple antennas could yield a reliable estimation of the tag's position.

To determine how soft features should be combined from multiple antennas, we train a neural network using the TensorFlow library in Python. Our neural network model uses three fully-connected hidden layers to learn the complex non-linear relationships between the features. To train the model, we deploy several tags in both inside and outside of the CA in arbitrary positions and orientations (see Fig. 7b). We use the commercially available Smartrac tags (see Fig. 7a). We then extract features from the tag readings and feed each tag's feature vector (specified in Sec. 6) to the neural network by labeling it as either corresponding to an "inside" tag or an "outside". The neural network then computes and stores the optimal weights for each of its connections by minimizing the decision error as defined by a loss function. After training, the model applies the weights to an input feature vector and outputs the probability of a tag belonging to the "inside" class. This is achieved by using one Sigmoid-activated output neuron that emits a value between 0 and 1; if the output is $\geq 0.5$ for a tag's feature vector, RFGo considers the particular tag as inside the CA.

## 6 PERFORMANCE EVALUATION

In this section, we first evaluate our custom-built reader based on three performance metrics. **Packet Decoding Rate (PDR)**: the fraction of slots that results in a correctly decoded EPC over the total number of slots. **Multi-decoding Efficiency (MDE)**: the total number of decoded EPCs over the total number of slots. Note that the number of decoded EPCs includes the simultaneous decoding of the same EPC from multiple antennas. **Discovery rate**: the percentage of the unique EPCs that have been decoded, as a function of time.

We then evaluate RFGo's performance for the self-checkout application based on **Recall** and **Precision** metrics, previously introduced in Sec. 1. Note that the "customer basket" and "all products in the CA" refer to the same set of tags, and the terms "tag" and "product" are used interchangeably.

### 6.1 Custom-built Reader Evaluation

Since the tag reading performance within the CA is of utmost importance for the self-checkout application, we evaluate the aforementioned performance metrics within the CA. For this set of experiments, we utilize the subset of antennas (six out of ten) positioned inside the CA to read tags that comprise multiple orientations and are placed in close proximity with respect to one another (see Fig. 7b).

*6.1.1 Packet Decoding Rate.* To evaluate the effect of simultaneous reception, we first activate a single TX antenna while receiving from all 6 antennas. In Fig. 8, we show the impact of different RN16 policies on PDR with varying number of tags. The Slotted Aloha curve in Fig. 8 assumes a single RN16 reply (i.e., no collision) is always decodable and any collision is not decodable. Interestingly, FP using the same antenna for TX and RX performs considerably better than the one predicted by the slotted Aloha curve (see FP empirical curve). The reason being that some collisions can be recovered, e.g., when the RSSI of a reply is high enough to be decoded despite other, weaker colliding replies[2].

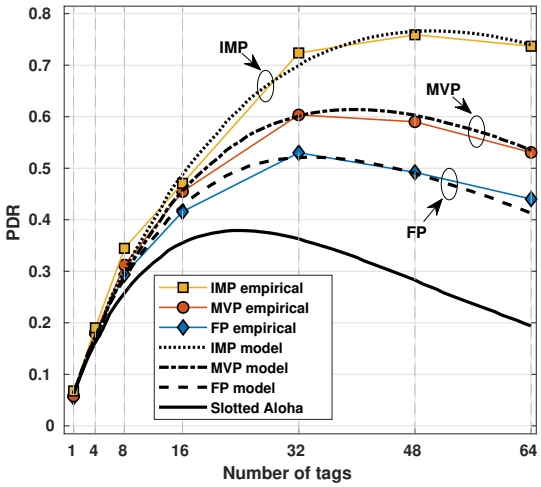



**Figure 8: PDR under different RN16 selection policies**

In addition, Fig. 8 reveals that MVP and IMP outperform FP. This improvement comes from the diversity of decoding the collided RN16s at multiple RX RF chains. Intuitively, the improvement of MVP over FP suggests that the existence of a majority is often a sign of the correctness of the RN16 sequence. On the other hand, IMP's higher performance evinces that the proposed metric IM accurately estimates the decoding probability and smartly selects the RN16 sequence to acknowledge. From Fig. 8, we note that a small tag population results in slot underuse and thus the low PDR, while a large population leads to an increased number of collisions.

In the following, we discuss our probabilistic model and analyze its results in more detail.

*6.1.2 The Proposed Probabilistic Model.* To formulate the impact of recoverable collisions, we develop a probabilistic model based on the probability of decoding at least one RN16 reply out of $k$ colliding RN16 replies in each slot by a policy $P$ (denoted by $\alpha_k^P$), for $1 \leq k \leq M$, where $M$ is the number of tags. Note that when $k = 1$, there is no collision. For $N$ slots, the PDR of policy $P$ is given by

$$\mathrm{PDR}(P) = \sum_{k=1}^{\lfloor \gamma M \rfloor} \alpha_k^P \binom{\lfloor \gamma M \rfloor}{k} \left(\frac{1}{N}\right)^k \left(1 - \frac{1}{N}\right)^{\lfloor \gamma M \rfloor - k} \quad (1)$$

where $\lfloor \gamma M \rfloor$ is the average number of tags that reply in a time slot. Note that $0 \leq \gamma \leq 1$ may be interpreted as the average fraction of the tag population that reply in a frame.

The probabilistic model of FP can be extended to MVP and IMP as follows. Given $\alpha_k^{FP}$, the probability of not being able to correctly decode a single RN16 out of $R$ received signals is $(1 - \alpha_k^{FP})^R$, assuming decoding independence across the RF chains. Therefore, the probability that IMP selects a correctly decoded RN16 is given

[2]It is reasonable to assume that if a tag's RN16 reply is decodable by an antenna (despite possible collisions), then its EPC would also be decodable by the same antenna due to the absence of collisions in the EPC reply stage.

by $\alpha_k^{IMP} = 1 - (1 - \alpha_k^{FP})^R$. Indeed, with this extension, we see a very close fit between the IMP model and the IMP empirical curves.

The probability that MVP selects a correct RN16 $\alpha_k^{MVP}$ may also be approximated based on $\alpha_k^{FP}$ given independence across RF chains. Consider different cases denoted by $CASE(k, j)$ where $k$ tags are colliding in a given time slot and $j$ out of $R$ receivers are able to correctly decode an RN16. The probability of $CASE(k, j)$ is given by

$$P(CASE(k,j)) = \binom{R}{j} (\alpha_k^{FP})^j (1 - \alpha_k^{FP})^{R-j}. \quad (2)$$

The probability $\alpha_k^{MVP}$ is then given by

$$\alpha_k^{MVP} = \sum_{j=0}^{R} \lambda_{k,j} \cdot P(CASE(k,j)) \approx \sum_{j=0}^{R} \underline{\lambda}_j \cdot P(CASE(k,j)) \quad (3)$$

where $\lambda_{k,j}$ is the probability that MVP picks an RN16 among the correctly decoded RN16s. We consider a coarser approximation by removing the dependency to $k$, leading to $\lambda_{k,j} \approx \underline{\lambda}_j$. We differentiate the following cases for $\underline{\lambda}_j$.

- $\underline{j \geq 3}$: There is a high probability that a majority exist in this case and the RN16 with highest popularity has very high chance of being correctly decoded. Hence, we approximate ($\underline{\lambda}_j \approx 1, \forall j \geq 3$).
- $\underline{j = 2}$: In this case, there is good chance that the decoded RN16 by two different RF chains are not the same. Hence, a majority does not exist and MVP selects one RN16 at random with probability of it being correct $1/R$. However, for smaller number of interferers, there is less chance that these RN16s are different. For example if $k = 1$, we have $\lambda_{1,2} = 1$. Hence, we use approximated value $\underline{\lambda}_2 \approx 0.5$ taken over all $j$.
- $\underline{j = 1}$: When only one RF chain correctly decodes an RN16, there is almost no chance to have majority. Hence, MVP selects one RN16 at random and the chance of being correct is only $\underline{\lambda}_1 = 1/R$.

The FP, IMP, and MVP model curves in Fig. 8 is obtained by setting $\gamma = 0.63$, i.e., 63% of tags are in average involved in a given frame, $\alpha_k^{FP} = 0$ for $k \geq 5$, i.e., more than five collisions is not salvageable and $(\alpha_1^{FP}, \alpha_2^{FP}, \alpha_3^{FP}, \alpha_4^{FP}) = (1.00, 0.49, 0.25, 0.20)$. The corresponding values of $\alpha_k^{IMP}$ and $\alpha_1^{MVP}$ is obtained from the same values for $\alpha_1^{FP}$. Fig. 8 shows that our probabilistic model fits very well with the empirical performance.

*6.1.3 Multi-Decoding Efficiency (MDE).* Improving the number of observations for every single tag reply by a multiplicative factor is one of the main advantages of our reader. Fig. 9 illustrates our evaluation results for MDE using IMP, MVP, and FP by having a single fixed TX antenna and 6 RX antennas within CA. Fig. 10 shows the ratio of the MDE to the PDR for different tag populations. Once again, IMP achieves the best MDE in terms of the overall number of observations due to its superior RN16 resolution capability. In addition, the ratio of MDE to PDR remains almost constant at a value of 4.5-5 regardless of the RN16 selection policy and tag population. Recall that the EPC packet never experiences collisions, for only one tag is acknowledged in a slot. Therefore, the ratio of MDE to PDR only depends on the average number of RX chains that can decode a tag's EPC packet, irrespective of the number of colliding tags or RN16 selection policy.

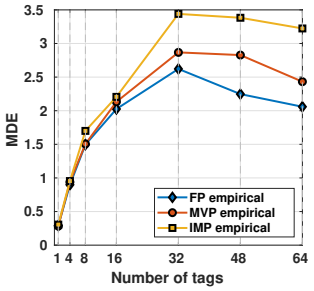



**Figure 9: Empirical MDE**

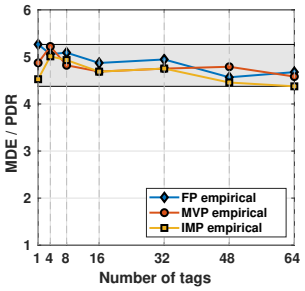
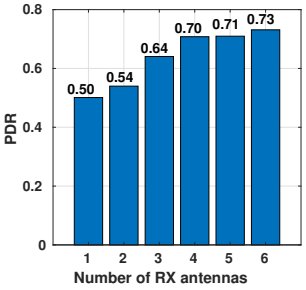
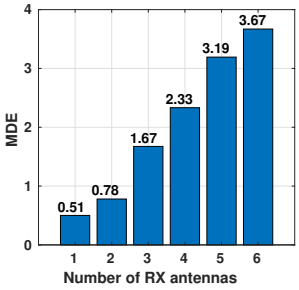



**Figure 10: MDE/PDR**




**(a) var. RX & 1 TX**




**(b) var. RX & 1 TX**

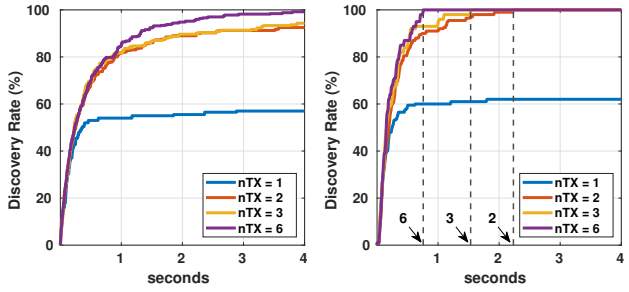
**Figure 11: Effect of the number of RX antennas on (a) PDR and (b) MDE**

Our evaluation so far has confirmed that IMP has superior performance over MVP and FP. Consequently, we use IMP for the rest of our evaluations.

*6.1.4 Effect of Receive Dimension.* We evaluate PDR (see Fig. 11a) and MDE (see Fig. 11b) for a single active TX antenna and a varying number of RX antennas, ranging from one to six. By increasing the number of simultaneously decoding RF chains, the PDR sharply increases between 1 to 4 RX antennas before exhibiting saturation for a larger receive dimension. On the contrary, MDE keeps increasing even beyond the saturation point of 4 RX antennas for PDR as a result of two main factors: the improvement in PDR and the possibility of decoding from more antennas. Although PDR experiences no significant improvement for 6 RX antennas with respect to 4 or 5 RX antennas, a larger number of antennas still increases the total number of observations by a considerable factor.

*6.1.5 Discovery rate.* In order to achieve 100% recall, any tag within the CA must be read, i.e., the discovery rate must be 100% within a session. For this set of experiments, we place 64 tags randomly within the CA. One of the major obstacles in tag discovery is the weak illumination. We activate different TX antennas in TDMA fashion to address issues related to illumination. Fig. 12a illustrates the discovery performance by employing a single RX antenna and different number of TX antennas. We see that increasing the number of TX antennas from 1 to 6 considerably improves the discovery rate from about 60% to 98% over a 4s interval.

Next, we analyze the benefit of our simultaneous decoding by increasing the number of RX antennas from 1 to 6. Our reader achieves perfect 100% discovery rate within 2.2s, 1.55s, and 0.75s using 2 (one antenna on each wall), 3 (one antenna on each wall and




(a) var. TX & 1 RX (b) var. TX & 6 RX

**Figure 12: Effect of the number of TX and RX antennas on the Discovery rate**

one on the floor), and 6 TX antennas as shown in Fig. 12b. Hence, simultaneous decoding increases the discovery rate and reaches the desired checkout speed.

## 6.2 Evaluating RFGo's Checkout Accuracy

In this section, we evaluate RFGo in detail for its checkout accuracy where the goal is to solely identify all the tags present inside the CA. Recall that RFGo uses a neural network-based classifier. We investigate the need for such a sophisticated classifier by first considering a simpler one that classifies tags as inside the CA when they are read by at least one of the inside antennas.

We define two regions for any classifier: (a) **core region**, a region inside the CA where tags are *almost surely* classified to be inside, and (b) **spillover region**, a region outside the core region where tags *may be* misclassified as inside. An ideal classifier would satisfy two conditions: (C1) the core region fully covers the CA, and (C2) the spillover region is negligible. The transmit power at TX antennas greatly impacts the ability to meet these conditions. For instance, a low TX power violates C1 while a high TX power works against C2. To reduce the spillover region, a system should use the minimum TX power that ensures C1.

*6.2.1 Evaluating Core Region and Spillover Region.* To evaluate the performance of the simple classifier, we conduct experiments in a grid of 6"×6" cells and locate tags by means of a unified cube composed of 27 identical smaller cubes. Each of the smaller cubes is attached one tag, covering various heights and orientations within each column (see Fig. 7c) and representing an approximation of possible tag placements within a 3-D volume. By deploying the unified cube in different positions, we identify the number of classified tags as inside for each cell among the three tags (on the three 6" cubes) in the same vertical slice of the unified cube. The readings are collected using the 6 antennas inside the CA for 2 seconds. As previous experiments show, RFGo discovers all the tags within 2 seconds.

The result for a snapshot of a typical 2 second experiment is shown in Fig. 13. The four shades from light to dark indicate the number of classified tags as inside (from 0 to 3). The result reveals that the physical structure of the walls effectively eliminates unwanted readings on the left and right sides of the CA. However, there is significant spillover extending far (e.g., 54" in this snapshot) from the entrance (and towards the exit, not shown for brevity). Our experiments show that even after using the minimum TX power to meet C1, the WA has to be pushed far away from the CA to guarantee acceptable precision (i.e., to meet C2). Such large separation between the CA and WA is not practical in most retail stores.

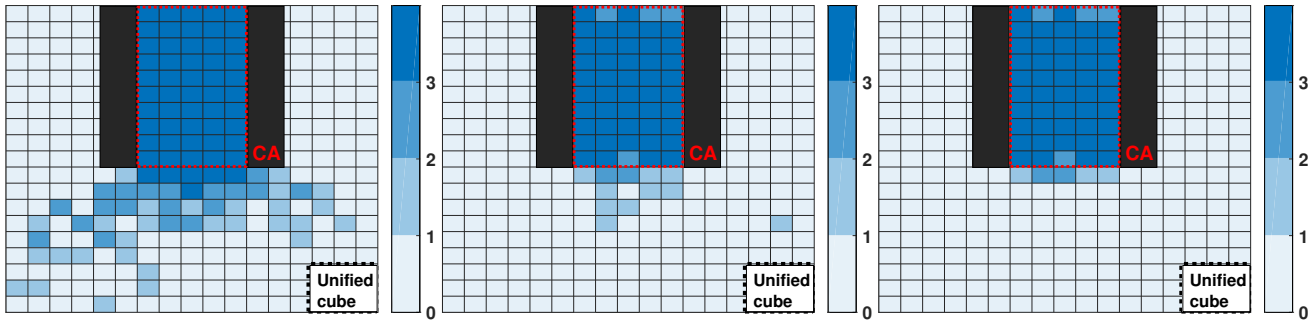
During this experiment, we observed the following: as the unified cube is placed further from the CA, some tag features such as the number of readings and their measured RSSIs "generally" decrease. However, there is no clear relation between RSSI and distance due to multipath. Nevertheless, it could be possible for a classifier to utilize these features to distinguish between tags inside and outside the CA. Using the setup in Sec. 5.3, we train a classifier with a 72-element feature vector for each tag comprised of the number of readings and average RSSI for each of the $6 \times 6 = 36$ TX/RX antenna pairs inside the CA. When a tag is not read by a given antenna pair, we set the corresponding number of readings to 0 and the average RSSI to -80 dBm (lowest observed in our system). After training, we run the classifier using the readings from the same snapshot to obtain the number of tags classified as inside in each cell.

Comparing Fig. 13b to Fig. 13a, we see that the classifier indeed shrinks the spillover region (to 24" from the entrance) but does not entirely eliminate it. Another undesirable effect is that the classifier also shrinks the core region. The problem comes from the misclassification of tags that are inside as outside (and vice versa) due to the similarity of their feature vectors at the boundary of the CA.

While the core region in Fig. 13b is acceptable, the spillover region needs further improvement. To further shrink the spillover region, RFGo enhances the classifier using our dual-area antenna deployment. The rationale is that the collective readings from both inside and outside antennas enhances the features, particularly for the tags along the physical boundary. We next train a classifier with the feature vector incorporating readings from all 10 antennas in RFGo (previous 6 inside plus 4 outside). The dual-area antenna deployment yields $10 \times 10 = 100$ TX/RX pairs instead of 36. We then apply the classifier on the same snapshot and present the results in Fig. 13c. We see that with the addition of outside antennas, the classifier aggressively shrinks the spillover region without sacrificing more from the core region. With this classifier, the spillover region is only up to 6" away from the entry and exit boundaries of the CA.

*6.2.2 Evaluating Self-checkout Accuracy.* In this section, we evaluate RFGo's checkout accuracy by attaching RFID tags to a set of packaged and loose apparel items such as shirts, socks and pants as in Fig. 15 placed inside a shopping basket. We consider different scenarios defined in terms of the number, position and orientation of tags placed inside the CA. The basket is placed in one of three different heights- 20cm, 40cm, and 60cm- and is measured from the ground to the bottom of the basket. The items are either (1) stacked (Fig. 15a) where the tags are parallel to the ground, i.e., horizontally orientated (H), (2) placed side-by-side (Fig. 15b) where the tags are vertically orientated (V), or (3) randomly oriented (R) (Fig. 15c). We run each scenario 50 times and randomly shuffle the items in the basket between the runs.

For all scenarios, we place a total of $26 \times 2 \times 16 = 832$ outside tags in the vicinity of the CA. There are 16 tags attached in different orientations to boxes of dimension 18"(L)×18"(W)×24"(H) as illustrated in Fig. 7b. Fig. 14 illustrates the CA, WA and the location of the outside boxes in a 6" grid. Fig. 14 also illustrates a guard area




(a) Untrained detection (b) Inside features (c) Inside/Outside features

**Figure 13: Snapshot of tags identified as in the CA for different classifiers using the unified cube in Fig. 7c**

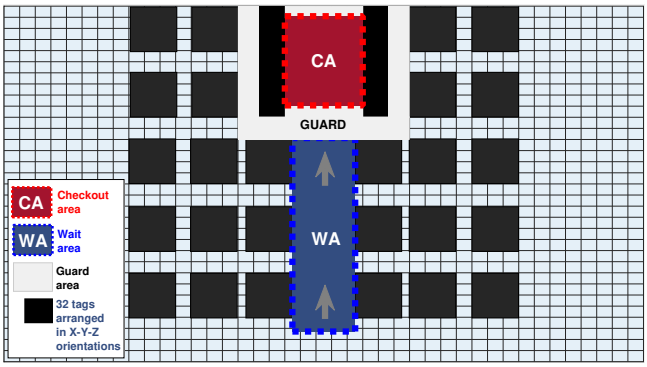



**Figure 14: Outside tag placement**

| Orientation | Height | No. of tags = 20 | | | No. of tags = 10 | | |
|---|---|---|---|---|---|---|---|
| | | I→I | I→O | O→I | I→I | I→O | O→I |
| V | 60 | 997 | 3 | 2 | 500 | 0 | 2 |
| H | 60 | 996 | 4 | 3 | 500 | 0 | 1 |
| R | 60 | 998 | 2 | 1 | 500 | 0 | 1 |
| V | 40 | 997 | 3 | 3 | 500 | 0 | 1 |
| H | 40 | 997 | 3 | 1 | 500 | 0 | 1 |
| R | 40 | 998 | 2 | 2 | 500 | 0 | 1 |
| V | 20 | 996 | 4 | 2 | 500 | 0 | 1 |
| H | 20 | 996 | 4 | 3 | 500 | 0 | 2 |
| R | 20 | 997 | 3 | 2 | 500 | 0 | 1 |

**Table 1: RFGo's self-checkout accuracy**

(a) Vertical (b) Horizontal (c) Random

**Figure 15: Item arrangement used in our experiments**

which is 6" from the side walls and 12" from the entrance. Each dark square of size 3×3 grid cells represent two boxes stacked together for a total of 32 tags.

There are 4 possible groups of tags in each run: (a) The tags placed inside and classified as inside. (b) The tags placed inside and classified as outside. (c) The tags placed outside and classified as inside. (d) The tags placed outside and classified as outside. In the checkout process, the relevant groups are (a), (b) and (c) denoted by I→I, I→O, and O→I in Table 1, respectively.

Table 1 shows I→I, I→O, and O→I for each scenario aggregated over 50 runs. Out of 13500 total number of tags that were placed inside the CA across 900 runs, the total number of tags in I→I is equal to 13472 which means the recall is $13472/13500 \approx 0.9979$. The total number of tags in O→I is 30 which yields a precision of $(13472 - 30)/13472 \approx 0.9977$. We noticed that in each run the number of the tags in the group I→O, or O→I were either zero or one which means that at most one item was misclassified.

*6.2.3 Checkout Experiments with Volunteers.* In this section, we evaluate RFGo with 5 volunteers each carrying a bag of 16 items that are randomly placed in different orientations for each run. One volunteer enters the CA and remain there until the scanning is complete (about 2 seconds), while the other four volunteers line up in the WA. The volunteers do not need to stand still, they can move or swing their bag of items and hold the bags in arbitrary heights.

Each volunteer completes 20 checkout sessions for a total of 100 sessions. Out of the 1600 items brought inside the CA, 1595 belong to the group I→I yielding a recall of $1595/1600 \approx 0.9968$. In addition, only 3 items are in the group O→I yielding a precision of $(1598 - 3)/1598 \approx 0.9981$.

We also compare the checkout performance of our system using a traditional reader instead of the custom-built reader in RFGo. Within the 2 seconds checkout time, the traditional reader suffers considerably with a recall of 92.9% due to two main shortcomings. (i) We observe that the traditional reader only discovers around 97% of the tags in the CA in comparison to 100% for our reader which is due to its simultaneous multi-point decoding capability. (ii) The classification performance based on the features from the traditional reader is lower since the number of features are much smaller. Using our custom-built reader and 10 antennas, each of 100 possible antenna pairs $(i, j), 1 \leq i, j \leq 10$ provides a feature, while traditional readers have 10 features only from the $(i, i)$ pairs.

*6.2.4 Impact of Our Custom-built Reader.* As demonstrated earlier, the custom-built reader in RFGo allows us to quickly collect

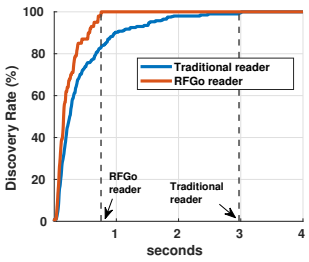



**Figure 16: Discovery rate for traditional vs. RFGo's reader.**

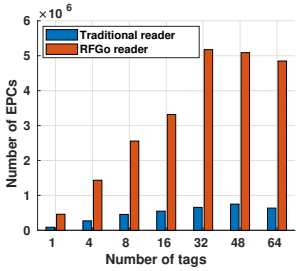



**Figure 17: # of EPCs reading, traditional vs. RFGo's reader.**

several tag readings which are then utilized by our model to accurately classify tags. To understand the importance of our reader, we evaluate RFGo with a traditional RFID reader using the same number of antennas in TDMA fashion. We emulate the operation of a traditional reader by decoding tag responses from the same TX antenna sending the query. The operation of the traditional reader has slower discovery of tags since spatial diversity across multiple RX antennas is not leveraged. Further, the total number of tag readings is significantly less since simultaneous decoding of tag replies is not implemented. Fig. 16 shows a typical example of how the discovery progresses over time with both readers. Within 2 seconds, the traditional reader can only discover around 97% of the tags in the CA while our custom-build reader achieves 100% discovery well before 2 seconds.

The mono-static operation also impacts the feature vector. Instead of containing features from all 100 antenna pairs, the traditional reader has only 10 TX/RX antenna pairs resulting in a much smaller set of features for training the model (Fig 17).

## 7 RELATED WORK

In this section, we present some of the recent advances in RFID tag localization and addressing tag collisions. As explained in Sec. 3, high RFID tag localization accuracy is not realistic in a typical retail environment where customers and products are constantly moving in unpredictable patterns. Also, studies addressing collisions do not tackle issues leading to blind spots such as orientation, coupling, multipath etc. that also impact the reading probability of a tag's response.

• **Preventing collisions with hashing:** Hashing-based solutions, where tags are allocated particular response slots based on their EPC, have been proposed to cope with the reading inefficiencies that Slotted Aloha incurs when the number of tags increases [1, 6, 17–19, 44–46, 48].

• **Estimating the dimension of colliding sets**: Estimating the number of tags in range based on a probabilistic analysis in every slot has also been of interest [13, 20–22, 30, 38–40, 47]. However such approach relies on a known tag set used as a reference. In addition, it does not provide a solution to fast and reliable extraction of EPC from the unknown tags.

• **Resolving collisions**: Other works have leveraged the information at the PHY layer, e.g., RSSI, phase, modulation, to assess whether a collision is present in a slot [13, 14]. Specifically, the FM0 bit flipping property can be used to classify bits from multiple tags and resolve collisions. Some reading improvement of 40% in [4] to 60% in [2] has been reported for parallel decoding of two colliding tags. The work in [27] attempts to resolves up to 5 colliding tags by assessing the linearity of the I/Q clusters to decode the RN16s.

• **Enhancing RFID tags:** In [25], the authors propose a non-linear backscattering mechanism to deal with the coupling issue. The multiple harmonics generated by the reader are backscattered by the tag using an extra module referred to as NLTL. The enhancement of tags is at the expense of standard incompatibility and higher costs, which may hinder deployment in practice.

• **Hologram-based Localization:** Recent studies leverage the diversity offered by the mobility of a tag to create a hologram that shows the maximum likelihood for each plausible location in a map [41–43]. The reported accuracy is within few centimeters. Alternatively, the authors in [31] emulate a virtualized antenna array by moving the transmit and receive antennas achieving an accuracy of 3-6 cm. The assumption of free antenna mobility may not fit certain deployments. While above studies show impressive localization accuracy, they do not fully address issues such as tag coupling since they demonstrate localization with relatively small tag populations.

• **Multi-antenna Localization:** In [8], authors use moving reference tags in a plane and localize a target tag by computing the interference patterns between them, with location error less than 4 cm. In [24], a phase-wrapping method is presented, where multiple reference tags are deployed a priori to characterize the phase dispersion. The goal is to detect misplaced items on shelves in retail spaces. Aside from relying on reference tags and their ground truth, these studies do not explicitly tackle dense tag deployments where tag coupling could impact the reported localization accuracy.

• **Computer Vision-assisted Localization:** The work [35] uses computer vision to enhance RFID tag localization and work despite unknown moving antenna trajectories. Although the authors report a few centimeters localization accuracy, the proposed technique works only for localizing static RFID-tagged objects.

## 8 CONCLUSIONS

In this paper, we present the design and implementation of RFGo, a first-of-its-kind self-checkout system based on RFID. Unlike camera-based approaches that require continuous monitoring of customers throughout a retail store, RFGo conveniently identifies purchases in a dedicated area where customers walk in with RFID-tagged products they wish to purchase. RFGo enables a seamless and highly accurate self-checkout experience where transactions are completed within two seconds without the customers having to manually place products in a dedicated container. Such a fast checkout speed is achieved by the custom-built reader that is specially designed to implement simultaneous decoding of a single tag response from multiple antennas synchronized at the carrier-level. In addition, our detailed evaluations demonstrate that RFGo achieves impressive accuracy in challenging scenarios and thus is ready for practical deployment.

## ACKNOWLEDGMENTS

This work was supported through an internship for Carlos Bocanegra at NEC Labs America and also by the US National Science Foundation under research grant CNS1452628.

# REFERENCES


[1] Zhenlin An, Qiongzheng Lin, Lei Yang, and Wei Lou. 2019. Embracing Tag Collisions: Acquiring Bloom Filters across RFIDs in Physical Layer. In *IEEE International Conference on Computer Communications (INFOCOM)*, Vol. 2019-April. IEEE, 1531–1539. https://doi.org/10.1109/INFOCOM.2019.8737432

[2] Christoph Angerer, Robert Langwieser, and Markus Rupp. 2010. RFID reader receivers for physical layer collision recovery. *IEEE Transactions on Communications* 58, 12 (2010), 3526–3537. https://doi.org/10.1109/TCOMM.2010.101910.100004

[3] Supreetha Rao Aroor and Daniel D. Deavours. 2007. Evaluation of the State of Passive UHF RFID: An Experimental Approach. *IEEE Systems Journal* 1, 2 (2007), 168–176. https://doi.org/10.1109/JSYST.2007.909179

[4] Aggelos Bletsas, John Kimionis, Antonis G. Dimitriou, and George N. Karystinos. 2012. Single-antenna coherent detection of collided FM0 RFID signals. *IEEE Transactions on Communications* 60, 3 (2012), 756–766. https://doi.org/10.1109/TCOMM.2011.020612.110212

[5] CBInsights. 2018. Beyond Amazon Go: The Technologies And Players Shaping Cashier-Less Retail. *[Blog post]* (10 2018). https://www.cbinsights.com/research/cashierless-retail-technologies-companies-trends/

[6] Honglong Chen, Guolei Ma, Zhibo Wang, Feng Xia, and Jiguo Yu. 2017. Probabilistic detection of missing tags for anonymous multicategory RFID systems. *IEEE Transactions on Vehicular Technology* 66, 12 (12 2017), 11295–11305. https://doi.org/10.1109/TVT.2017.2726005

[7] Raghu Das. 2018. RFID 2018-2028: RAIN and NFC Market Status, Outlook and Innovations. *IDTechEx* (3 2018). https://rainrfid.org/wp-content/uploads/2018/03/IDTechEx-RFID-March-_Distribute.pdf

[8] Han Ding, Jinsong Han, Chen Qian, Fu Xiao, Ge Wang, Nan Yang, Wei Xi, and Jian Xiao. 2018. Trio: Utilizing Tag Interference for Refined Localization of Passive RFID. In *IEEE International Conference on Computer Communications (INFOCOM)*, Vol. 2018-April. IEEE, 828–836. https://doi.org/10.1109/INFOCOM.2018.8486313

[9] Dobkin. 2008. *The RF in RFID: UHF RFID in Practice.* Elsevier. 1–493 pages. https://doi.org/10.1016/B978-0-7506-8209-1.X5001-3

[10] Forrester. 2018. Consumers Cringe At Slow Checkout. *Forrester Opportunity Snapshot: Digimarc August 2018* (8 2018). https://www.digimarc.com/resources/forrester-study

[11] HIROMITSU GOTO. 2020. Materials maker Toray's new tech to slash smart tag price by 80%. *Nikkei Asian Review* (1 2020). https://asia.nikkei.com/Business/Technology/Materials-maker-Toray-s-new-tech-to-slash-smart-tag-price-by-80

[12] Craig Guillot. 2019. Reducing customers' biggest pain point. *National Retail Federation (NRF)* (11 2019). https://nrf.com/blog/reducing-customers-biggest-pain-point

[13] Yuxiao Hou, Jiajue Ou, Yuanqing Zheng, and Mo Li. 2015. PLACE: Physical layer cardinality estimation for large-scale RFID systems. In *IEEE International Conference on Computer Communications (INFOCOM)*, Vol. 26. IEEE, 1957–1965. https://doi.org/10.1109/INFOCOM.2015.7218579

[14] Yuxiao Hou and Yuanqing Zheng. 2018. PHY-Tree: Physical Layer Tree-Based RFID Identification. *IEEE/ACM Transactions on Networking* 26, 2 (4 2018), 711–723. https://doi.org/10.1109/TNET.2018.2791938

[15] Chengkun Jiang, Yuan He, Songzhen Yang, Junchen Guo, and Yunhao Liu. 2019. 3D-OmniTrack: 3D tracking with COTS RFID systems. In *ACM Information Processing in Sensor Networks (IPSN)*. ACM Press, New York, New York, USA, 25–36. https://doi.org/10.1145/3302506.3310386

[16] Nikos Kargas, Fanis Mavromatis, and Aggelos Bletsas. 2015. Fully-Coherent reader with commodity SDR for Gen2 FM0 and computational RFID. *IEEE Wireless Communications Letters* 4, 6 (2015), 617–620. https://doi.org/10.1109/LWC.2015.2475749

[17] Yuan Cheng Lai, Riyanto Jayadi, Sheng Chi Huang, and Yen Hung Chen. 2018. Query tree with knowledge-based splitting for RFID tag identification. In *ACM International Conference Proceeding Series*. ACM Press, New York, New York, USA, 24–28. https://doi.org/10.1145/3290589.3290592

[18] Jia Liu, Bin Xiao, Shigang Chen, Feng Zhu, and Lijun Chen. 2015. Fast RFID grouping protocols. In *IEEE International Conference on Computer Communications (INFOCOM)*, Vol. 26. IEEE, 1948–1956. https://doi.org/10.1109/INFOCOM.2015.7218578

[19] Xiulong Liu, Jiannong Cao, Yanni Yang, Wenyu Qu, Xibin Zhao, Keqiu Li, and Didi Yao. 2019. Fast RFID Sensory Data Collection: Trade-off between Computation and Communication Costs. *IEEE/ACM Transactions on Networking* 27, 3 (6 2019), 1179–1191. https://doi.org/10.1109/TNET.2019.2914412

[20] Xiulong Liu, Keqiu Li, Alex X. Liu, Song Guo, Muhammad Shahzad, Ann L. Wang, and Jie Wu. 2017. Multi-Category RFID Estimation. *IEEE/ACM Transactions on Networking* 25, 1 (2 2017), 264–277. https://doi.org/10.1109/TNET.2016.2594481

[21] Xiulong Liu, Bin Xiao, Keqiu Li, Alex X. Liu, Jie Wu, Xin Xie, and Heng Qi. 2017. RFID Estimation with Blocker Tags. *IEEE/ACM Transactions on Networking* 25, 1 (2 2017), 224–237. https://doi.org/10.1109/TNET.2016.2595571

[22] Xiulong Liu, Xin Xie, Keqiu Li, Bin Xiao, Jie Wu, Heng Qi, and Dawei Lu. 2017. Fast Tracking the Population of Key Tags in Large-Scale Anonymous RFID Systems. *IEEE/ACM Transactions on Networking* 25, 1 (2017), 278–291. https://doi.org/10.1109/TNET.2016.2576904

[23] Xiulong Liu, Xin Xie, Shangguang Wang, Jia Liu, Didi Yao, Jiannong Cao, and Keqiu Li. 2019. Efficient Range Queries for Large-Scale Sensor-Augmented RFID Systems. *IEEE/ACM Transactions on Networking* 27, 5 (10 2019), 1873–1886. https://doi.org/10.1109/TNET.2019.2936977

[24] Jiaqing Luo and Kang G. Shin. 2019. Detecting misplaced RFID tags on static shelved items. In *ACM International Conference on Mobile Systems, Applications, and Services (MobiSys)*. ACM Press, New York, New York, USA, 378–390. https://doi.org/10.1145/3307334.3326085

[25] Yunfei Ma, Xiaonan Hui, and Edwin C. Kan. 2016. 3D real-time indoor localization via broadband nonlinear backscatter in passive devices with centimeter precision. In *ACM International Conference on Mobile Computing and Networking (MobiCom)*. ACM Press, New York, New York, USA, 216–229. https://doi.org/10.1145/2973750.2973754

[26] Yunfei Ma, Nicholas Selby, and Fadel Adib. 2017. Minding the billions: Ultra-wideband localization for deployed RFID tags. In *ACM International Conference on Mobile Computing and Networking (MobiCom)*, Vol. Part F1312. 248–260. https://doi.org/10.1145/3117811.3117833

[27] Jiajue Ou, Mo Li, and Yuanqing Zheng. 2015. Come and be served: Parallel decoding for COTS RFID tags. In *ACM International Conference on Mobile Computing and Networking (MobiCom)*. ACM Press, New York, New York, USA, 500–511. https://doi.org/10.1145/2789168.2790101

[28] Swadhin Pradhan, Eugene Chai, Karthikeyan Sundaresan, Lili Qiu, Mohammad A. Khojastepour, and Sampath Rangarajan. 2017. RIO: A pervasive RFID-based touch gesture interface. In *ACM International Conference on Mobile Computing and Networking (MobiCom)*. Association for Computing Machinery, 261–274. https://doi.org/10.1145/3117811.3117818

[29] NXP Semiconductors. 2018. Big in Japan: 100 Billion UHF RFID Tags per Year by 2025. *RFID and Wireless IoT research* (3 2018). https://www.rfid-wiot-search.com/big-in-japan-100-billion-uhf-rfid-tags-per-year-by-2025

[30] Muhammad Shahzad and Alex X. Liu. 2015. Probabilistic Optimal Tree Hopping for RFID Identification. In *IEEE/ACM Transactions on Networking*, Vol. 23. IEEE Press, 796–809. https://doi.org/10.1109/TNET.2014.2308873

[31] Longfei Shangguan and Kyle Jamieson. 2016. The design and implementation of a mobile RFID tag sorting robot. In *ACM International Conference on Mobile Systems, Applications, and Services (MobiSys)*. ACM Press, New York, New York, USA, 31–42. https://doi.org/10.1145/2906388.2906417

[32] Focal Systems. 2019. Analysis of the amazon go platform and its implications on large format grocery stores. *[Blog post]* (7 2019). https://medium.com/focal-systems/analysis-of-the-amazon-go-platform-and-its-implications-on-large-format-grocery-stores-727d9b25f04a

[33] Ju Wang, Liqiong Chang, Omid Abari, and Srinivasan Keshav. 2019. Are RFID sensing systems ready for the real world?. In *ACM International Conference on Mobile Systems, Applications, and Services (MobiSys)*. ACM Press, New York, New York, USA, 366–377. https://doi.org/10.1145/3307334.3326084

[34] Jue Wang and Dina Katabi. 2013. Dude, where's my card? RFID positioning that works with multipath and non-line of sight. *ACM Computer Communication Review (SIGCOMM)* 43, 4 (8 2013), 51–62. https://doi.org/10.1145/2534169.2486029

[35] Zhongqin Wang, Min Xu, Ning Ye, Fu Xiao, Wang none Ruchuan, and Haiping Huang. 2019. Computer Vision-assisted 3D Object Localization via COTS RFID Devices and a Monocular Camera. *IEEE Transactions on Mobile Computing (TMC)* (2019), 1–1. https://doi.org/10.1109/TMC.2019.2954830

[36] Claude Webb and Geoffrey Sammut I. 2010. *Encyclopedia of Machine Learning*. 1–1031 pages. https://doi.org/10.1007/978-0-387-30164-8

[37] Teng Wei and Xinyu Zhang. 2016. Gyro in the air: Tracking 3D orientation of batteryless Internet-of-Things. In *ACM International Conference on Mobile Computing and Networking (MobiCom)*. ACM Press, New York, New York, USA, 55–68. https://doi.org/10.1145/2973750.2973761

[38] Qingjun Xiao, Min Chen, Shigang Chen, and Yian Zhou. 2015. Temporally or spatially dispersed joint RFID estimation using snapshots of variable lengths. In *ACM International Conference on Mobile Ad Hoc Networking and Computing (MobiHoc)*. ACM Press, New York, New York, USA, 247–256. https://doi.org/10.1145/2746285.2746289

[39] Qingjun Xiao, Shigang Chen, and Min Chen. 2016. Joint Property estimation for multiple RFID tag sets using snapshots of variable lengths. In *ACM International Conference on Mobile Ad Hoc Networking and Computing (MobiHoc)*. ACM Press, New York, New York, USA, 151–160. https://doi.org/10.1145/2942358.2942377

[40] Qingjun Xiao, Bin Xiao, Shigang Chen, and Jiming Chen. 2017. Collision-Aware Churn Estimation in Large-Scale Dynamic RFID Systems. *IEEE/ACM Transactions on Networking* 25, 1 (2017), 392–405. https://doi.org/10.1109/TNET.2016.2586308

[41] Huatao Xu, Dong Wang, Run Zhao, and Qian Zhang. 2019. AdaRF: Adaptive RFID-based Indoor Localization Using Deep Learning Enhanced Holography. *ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)* 3, 3 (9 2019), 1–22. https://doi.org/10.1145/3351271

[42] Huatao Xu, Dong Wang, Run Zhao, and Qian Zhang. 2019. FaHo: deep learning enhanced holographic localization for RFID tags. In *ACM International Conference on Embedded Networked Sensor Systems (SenSys )*. ACM Press, New York, New York, USA, 351–363. https://doi.org/10.1145/3356250.3360035

[43] Lei Yang, Yekui Chen, Xiang-Yang Li, Chaowei Xiao, Mo Li, and Yunhao Liu. 2014. Tagoram: Real-Time Tracking of Mobile RFID Tags to High Precision Using COTS Devices. In *ACM International Conference on Mobile Ad Hoc Networking and Computing (MobiHoc)*. ACM Press, New York, New York, USA, 237–248. https://doi.org/10.1145/2639108.2639111

[44] Jihong Yu, Wei Gong, Jiangchuan Liu, and Lin Chen. 2018. Fast and Reliable Tag Search in Large-Scale RFID Systems: A Probabilistic Tree-based Approach. In *IEEE International Conference on Computer Communications (INFOCOM)*, Vol. 2018-April. IEEE, 1133–1141. https://doi.org/10.1109/INFOCOM.2018.8485912

[45] Jihong Yu, Wei Gong, Jiangchuan Liu, Lin Chen, and Kehao Wang. 2019. On Efficient Tree-Based Tag Search in Large-Scale RFID Systems. *IEEE/ACM Transactions on Networking* 27, 1 (2 2019), 42–55. https://doi.org/10.1109/TNET.2018.2879979

[46] Jihong Yu, Jiangchuan Liu, Rongrong Zhang, Lin Chen, Wei Gong, and Shurong Zhang. 2019. Multi-Seed Group Labeling in RFID Systems. *IEEE Transactions on Mobile Computing (TMC)* (2019), 1–1. https://doi.org/10.1109/tmc.2019.2934445

[47] Ziling Zhou, Binbin Chen, and Haifeng Yu. 2016. Understanding RFID Counting Protocols. *IEEE/ACM Transactions on Networking* 24, 1 (2 2016), 312–327. https://doi.org/10.1109/TNET.2014.2361149

[48] Weiping Zhu, Xing Meng, Xiaolei Peng, Jiannong Cao, and Michel Raynal. 2019. Collisions are Preferred: RFID-based Stocktaking with a High Missing Rate. *IEEE Transactions on Mobile Computing (TMC)* (2019), 1–1. https://doi.org/10.1109/tmc.2019.2911586

[49] Simone Zuffanelli, Gerard Zamora, Pau Aguila, Ferran Paredes, Ferran Martin, and Jordi Bonache. 2016. Analysis of the Split Ring Resonator (SRR) Antenna Applied to Passive UHF-RFID Tag Design. *IEEE Transactions on Antennas and Propagation* 64, 3 (3 2016), 856–864. https://doi.org/10.1109/TAP.2015.2513084