

Longitudinal Study of a Building-Scale RFID Ecosystem

Evan Welbourne, Karl Koscher, Emad Soroush,
Magdalena Balazinska, Gaetano Borriello

Department of Computer Science and Engineering
University of Washington, Seattle
Seattle, WA 98195

{evan,supersat,soroush,magda,gaetano}@cs.washington.edu

ABSTRACT

Radio Frequency IDentification (RFID) deployments are becoming increasingly popular in both industrial and consumer-oriented settings. To effectively exploit and operate such deployments, important challenges must be addressed, from managing RFID data streams to handling limitations in reader accuracy and coverage. Furthermore, deployments that support pervasive computing raise additional issues related to user acceptance and system utility. To better understand these challenges, we conducted a four-week study of a building-scale EPC Class-1 Generation-2 RFID deployment, the “RFID Ecosystem”, with 47 readers (160 antennas) installed throughout an 8,000 square meter building. During the study, 67 participants having over 300 tags accessed the collected RFID data through applications including an object finder and a friend tracker and several tools for managing personal data. We found that our RFID deployment produces a very manageable amount of data overall, but with orders of magnitude difference among various participants and objects. We also find that the tag detection rates tend to be low with high variance across the type of tag, participant and object. Users need expert guidance to effectively mount their tags and are encouraged by compelling applications to wear tags more frequently. Finally, probabilistic modeling and inference techniques promise to enable more complex applications by smoothing over gaps and errors in the data, but must be applied with care as they add significant computational and storage overhead.

Categories and Subject Descriptors

H.2 [Information Systems]: Information Interfaces and Presentation; H.4 [Information Systems]: Miscellaneous

General Terms

Measurement

Keywords

RFID, Measurement, Pervasive Computing

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. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MobiSys'09, June 22–25, 2009, Kraków, Poland.

Copyright 2009 ACM 978-1-60558-566-6/09/06 ...\$5.00.

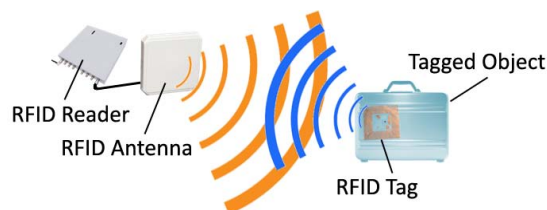


Figure 1: RFID reader antennas read RFID tags carried by people and objects in an RFID deployment.

1. INTRODUCTION

Radio Frequency Identification (RFID) technology has become increasingly popular in the last several years with applications appearing both in industrial [8, 46] and user-oriented settings [35, 40].

In an RFID deployment, an environment is instrumented with RFID readers while users and objects carry RFID tags as illustrated in Figure 1. Readers are equipped with antennas so that when a tag passes in the vicinity of an antenna, the reader detects the presence of the tag and generates a triple of the form: (TAG_ID, LOCATION, TIME), which records the location and time when the tag was sighted. Applications use these location traces to offer various services from simple product tracking [8, 46] to sophisticated activity recognition and workflow analysis [35, 40].

There are different types of RFID technology. In particular, some systems use active (*i.e.*, battery-powered) tags, while others use passive tags. Passive RFID, and particularly the EPC Class-1 Generation-2 standard, promises to enable large tracking applications at much better cost and convenience. This is not only because passive tags cost tens of cents while active tags cost tens of dollars, but because passive tags do not require periodic change of battery.

Although passive RFID technology is successfully used in constrained industrial settings, many questions remain as to the practicality and usefulness of the technology for pervasive computing applications. Though many applications have been proposed for passive RFID [12, 20, 4, 5, 30, 32, 38] to our knowledge, very few studies have been conducted on the behavior and performance of EPC Class-1 Generation-2 UHF RFID in a realistic, long-term pervasive computing setting. Most previous studies were limited to controlled laboratory environments [7, 19, 37]. As such, the community is left with a number of questions that laboratory studies cannot answer, such as how well do results from the laboratory map to an everyday-life setting? How much data does

a community of users in a pervasive RFID system generate and what does the data look like? How well the tags can be read when mounted on everyday objects by users? By experts? Are users eager to adopt this new technology? How do they manage their private location data? How do usage trends evolve over time? Are there any trends or correlations in the data that might inform system and application design? What techniques could be used to cope with uncertainty in the data resulting from sensor failure or limitations in sensor coverage? Are these techniques effective?

To begin answering the above questions, we conducted a four-week study of a building-wide RFID deployment with 47 readers (160 antennas), 67 participants, and 324 tags. During the study, participants had access to a variety of applications including a friend finder, an object and tag manager, a data browser and a Facebook application. We also complemented the study with microbenchmark experiments conducted in a laboratory and in a building-wide deployment. In this paper, we present results in connection with four key aspects of an RFID deployment:

(1) *System and application utilization*: We measure system utilization in terms of the amount and character of data generated in addition to patterns of application use. We also identify trends in these quantities over time. As part of these measurements, we record how often users interact with their RFID data or delete sensitive data. Overall, we find that an RFID infrastructure produces small data volumes (approximately 0.8 MB of data per participant over 4 weeks on average) and that this data can nicely be compressed (*e.g.* 20KB per person on average). We find more than an order of magnitude difference in the amount of RFID data generated by different users and objects and that at least four factors are responsible for this trend: differences in user mobility, differences in how people carry their tags, differences in materiality of tagged objects and differences in the degree to which people use the applications (heavy application users tend to produce the most RFID data). Interestingly, we also found that few people deleted any data and no one withdrew from the study.

(2) *System performance*: We measure the load distribution across readers and over time. As in previous mobility studies (*e.g.*, [3]), we find that some locations are hot spots. In the case of an RFID deployment, however, such trends can amount to an order of magnitude difference in the quantity of data produced at different locations. More importantly, we estimate the reliability of the system: *i.e.*, how often readers detect different types of tags in their vicinity when tags are managed and used by non-experts. We find that the detection rates are relatively low overall and highly variable. We find significant differences in detection rate among types of tags, types of objects, and individuals. Some individuals and objects are detected upward of 80% of the time. The median detection rates, however, are below 40%.

(3) *Inferring higher-level information from RFID data*: Building all but the simplest applications directly on raw RFID data is difficult because tags often go undetected and readers cannot cover all areas of interest. To address this challenge, recent work proposes to build and use probabilistic model-based views [10, 9, 39, 50]. We investigate the challenges involved in the practical application of such a technique in a large scale RFID deployment. We use standard methods to build probabilistic views over the collected RFID data which transform the raw RFID data into a

smooth distribution of the tags' possible location over time. We find that this technique is promising with room-level location recalls up to 60% in spite of the significant inaccuracy and uncertainty in the underlying RFID streams. Additional care and tuning, however, are needed to achieve higher recall rates. Finally, these techniques impose significant overhead: the amount of data increases by a significant factor (a factor of 13.5 in our experiments). This overhead, however, can be dramatically reduced by ignoring the lowest-probability data. Applying such techniques at scale in an RFID-based pervasive computing environment is thus challenging.

(4) *Miscellaneous Challenges*: During the study and in preparation for the study, we encountered a variety of challenges. For example, deployment of antennas in a busy, shared public space; management and maintenance of a system with hundreds of distributed components; and ensuring user privacy when experimenting with a technology that often considered to be insecure and invasive in nature. We report on the lessons that we learned.

In summary, pervasive, passive RFID deployments enable interesting applications that users are willing to adopt. Exploiting RFID infrastructures in pervasive computing environments, however, raises significant technical challenges primarily due to the low detection rates of tags, especially when these tags are managed directly by users.

The rest of this paper is organized as follows. In Section 2, we first present the RFID hardware, deployment, and performance benchmarks. We describe the RFID Ecosystem applications in Section 3. We describe our study and present its results in Sections 4 and 5 respectively. We discuss additional challenges that we encountered in Section 6, present related work in Section 7, before concluding in Section 8.

2. THE RFID ECOSYSTEM

Our work is centrally focused on experimentation with a building-wide RFID deployment called the RFID Ecosystem. The goal of this deployment is to simulate the conditions of an RFID-saturated future environment at scale for long periods of time, and in a carefully controlled setting. We discuss the details of the RFID Ecosystem in this section. In particular, we describe the RFID equipment used as well as the deployment layout, parameters, software, and alternatives we decided against. We also present results from laboratory experiments and microbenchmarks designed to characterize the behavior of our system in terms of data generated and success rates when operating under various conditions.

2.1 RFID Readers and Tags

As explained in the introduction, the two fundamental components of any RFID deployment are the RFID readers, which are installed throughout an environment, and the RFID tags, which may be attached to people or objects. In the RFID Ecosystem, we use equipment that implements the popular EPC Class-1 Generation-2 (EPC Gen 2) standard [2]. EPC Gen 2 RFID operates in the 860MHz-960MHz frequency range using the passive-backscatter, interrogator-talks-first approach and has an operating range that is typically measured in meters.

We use Speedway readers from Impinj [22]. These readers are equipped with Ethernet and run Linux with proprietary firmware that exposes tag inventory results to user-authored code. To each reader, we attach between 1 and 4 Cushcraft

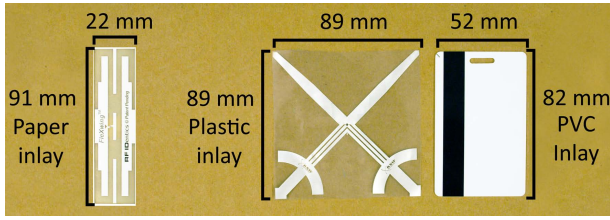


Figure 2: The three tag types used in our study: the FleXwing, the Excalibur, and the RFIDentics card.

S9028PC antennas. The S9028PC measures 25.4cm x 25.4 cm x 3.8cm, has a gain of 9 dBic, is circularly polarized (left or right), has a VSWR of 1.5:1 and a beamwidth of 60° by specification. We set the readers in *dense reader mode* (Miller-4 tag coding) [2]. This mode is designed to reduce RF interference from other nearby readers in our deployment and has been shown to provide a longer read range than other modes [7]. Finally, we set the reader transmit power to the maximum allowed 30 dBm so as to allow for maximum operating range. In combination with the Speedway reader we use three types of tags (see Figure 2): the RFIDentics FleXwing [41] tag, the KSW Microtec Excalibur tag [26] and an RFIDentics tag embedded in a PVC magstripe card [41]. All three tags use the Impinj Monza chip and store a manufacturer-set 96-bit Electronic Product Code (EPC). In this paper, we refer to the EPC as the tag identifier or tag ID since we tag personal objects and people rather than products.

To characterize the relative performance of our three types of tags, we conducted a laboratory experiment in which we measured the *read-rate* for each tag at various positions and orientations relative to a reader antenna. As Hodges et al. [19] point out, read-rate is a simple and popular way to assess an RFID system’s operating range that takes into account all factors affecting range in an environment. We formally define read-rate as:

DEFINITION 2.1. *Read-rate: The rate at which a reader can successfully read a tag.*

We began the experiment by placing a single reader and antenna against one wall of a 6m x 8.5m room containing tables, desktops and wooden shelves (with metal braces) at its sides. We then moved each type of tag through a series of 245 points inside a half-sphere with a 5m radius in front of the reader antenna. At each point the tag was oriented parallel to and then perpendicular to the face of the reader antenna. We measured read-rate over a 10 second period for each position. A horizontal ($\phi = 90^\circ$) slice of the half-sphere is shown for each tag type in Figure 3; darker shade indicates higher read rate. Figures 3(a),(b) and (c) show the results for parallel tag orientation. Notice first that in general, the read rates are highest in the *major detection field* [17], an arc directly in front of the antenna, and decrease as the tag moves out of this arc, through the antenna’s *minor detection field*, to the periphery of the full operating range. The Excalibur has the best read-rate throughout the full operating range while the FleXwing and PVC card offer good rates over a slightly smaller range. While the Excalibur and FleXwing tags were quite robust to change in orientation, the PVC card shows significant performance degradation when oriented perpendicular to the reader antenna (see Figure 3(d)). These results suggest that the Excalibur tag

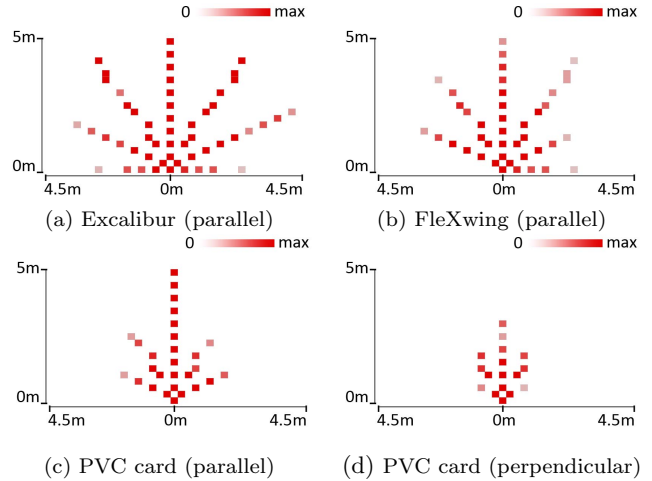


Figure 3: The read rates for the three types of tags at points inside a horizontal slice of a half-sphere centered in front of the reader antenna.

should offer the best performance in the variety of situations, followed by the FleXwing and then the PVC card. More rigorous laboratory studies on passive UHF RFID performance can be found in related work [7, 19, 37]. The goal of this experiment was simply to understand the basic performance properties of the tags used in our study.

2.2 Building-Scale RFID Deployment

We deployed 47 readers and 160 reader antennas throughout the 7 floors of our 8,000 square meter Computer Science and Engineering building with antennas located as shown in Figure 4. Two adaptations of standard industry configurations [15] were used to install antennas: *halfway portal* (HP) and *gateway choke point* (CP). In the HP configuration, an antenna is mounted above a hallway pointing downward to read tags passing through that section of hallway. The CP configuration points antennas with opposite polarization inward (towards each other) from both sides of an entrance to a particular part of the building (e.g., atrium, laboratory). HPs afford a cheaper installation with wider coverage using less hardware while CPs are designed to provide higher read-rates at important transition points. Full deployment details are summarized in Table 1. These deployment parameters were shaped not only by inevitable budgetary and regulatory considerations, but by the aesthetic and ethical concerns of our community as well. As such, we also had to hide CP installations inside custom-built cherry wood boxes and avoid deploying antennas in offices altogether so as to prevent RF interference with telephones and other personal equipment. The latter constraint, which limits reader topology, is shared by important domains like hospitals, where minimizing interference from RFID equipment is a crucial requirement [45]. We present one approach to coping with the problem of limited reader topology in Section 5.4.

All readers in our deployment run custom software (700 lines of C code) that processes new RFID data before streaming it over wired or wireless Ethernet to a central server. This software continuously polls the firmware for new tag reads and packs any data into one *tag read event* (TRE) per tag per antenna per second, a tuple with the schema: (tag ID, antenna ID, start time, number of reads). For example, if within a second a tag *a* is read 32 times by an-

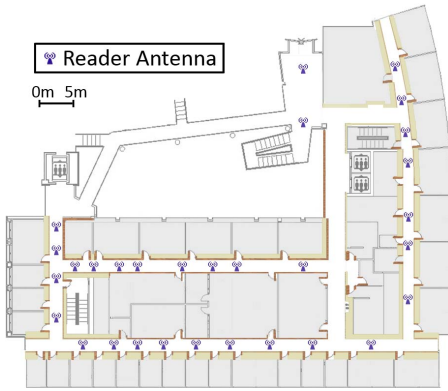


Figure 4: Layout of a representative floor.

Count	Mounting	Config	Elevation
132	Under cable tray	HP	2.4 m
5	Under lounge ceiling	HP	3 m
5	Above stairwell	HP	3 m
10	Beside building entry	CP	45 cm
2	Beside conference rooms	CP	45 cm
6	Under lab cable trays	HP	3 m

Table 1: RFID antenna deployment parameters.

tenna 1 starting δ_1 milliseconds after time τ and 20 times by antenna 2 δ_2 milliseconds after time τ , then our software will generate and send two TREs to the server: $(A, 1, \tau + \delta_1, 32)$ and $(A, 2, \tau + \delta_2, 20)$. Each reader also runs ntp to synchronize its clock with the rest of the system.

The central server stores all TRE data in a SQL Server database [44] that also contains metadata on the deployment including the latitude and longitude of each antenna and a symbolic antenna name (e.g., “front entrance”, “4th floor stairwell”). Custom Java code (56 classes) orchestrates data transmission between the readers and the database using JDBC and Apache’s MINA library for efficient, secure networking. All applications run on top of the RFID Ecosystem are implemented using the Cascadia system libraries [50]. However, with the exception of the probabilistic data in Section 5.4, all of the data reported and analyzed in this paper come directly from the RFID Ecosystem and are not processed by Cascadia. Finally, a set of administrative and diagnostic web applications are hosted with Apache and Tomcat on a separate web server.

2.2.1 Deployment Benchmarks

Unlike the idealistic single antenna laboratory case, RFID performance in a large deployment is more complex and depends on a variety of other factors like velocity and changing orientation of tags, ambient RF interference and interference from materials composing tagged objects or people. To better understand performance in our deployment amidst these effects, we extracted some baseline measures using three microbenchmarks. In the first microbenchmark (HP), a researcher walked through a sequence of 145 HPs carrying 15 tags attached to various objects. In the second (CP), the same researcher walked through just 2 CPs carrying the same tagged objects. A third microbenchmark (MT) investigates the *multi-tagging* technique proposed by Ravindranath et al. [38], in which multiple tags are attached to an object at different orientations in an effort to increase the probability that the object is detected. In the MT microbenchmark,

Object	Material	Tag	HP	CP	MT
3-ring binder	paper	EX	0.16	0.8	0.6
DVD	metal	FW	0.09	0.2	0.12
Hardcover	paper	EX	0.18	0.7	0.5
Magazine	paper	EX	0.54	0.6	–
Mobile phone	metal	FW	0.04	0.25	0.1
Paperback	paper	EX	0.39	0.55	0.7
Power cable	metal	FW	0.27	0.4	0.3
Water bottle	plastic	FW	0.61	0.8	–
Backpack	cloth	EX	0.58	0.9	–
Badge	plastic	EX	0.77	1.0	–
Belt	cloth	PVC	0.02	0.15	0.02
Hat	cloth	EX	0.82	1.0	–
Jacket	cloth	EX	0.28	0.3	0.8
Shirt	cloth	EX	0.07	0.4	0.1
Wallet	cloth	FW	0.05	0.0	0.06

Table 2: Tagged objects (the first 8 were carried inside the backpack), type of tag used (EX is Excalibur, FW is Flexwing, PVC is PVC card) and average detection rate for each microbenchmark.

one additional tag was added to each object having poor results in the HP microbenchmark, after which the researcher walked through the same sequence of 145 HPs. For each microbenchmark, we calculate detection-rate, which we define as follows:

DEFINITION 2.2. *Detection-rate*: The probability that a tagged object passing in the vicinity of an HP or CP is detected at least once.

We compute the detection-rate for each object as the fraction of HPs (or CPs) passed which detected at least one of the object’s tags. The first three columns of Table 2 describe the tagged objects as well as the type of tag used.

2.2.2 Results

Across 4 trials of the HP microbenchmark, 10 trials of CP and 4 trials of MT, we collected 155 minutes of trace data. Table 2 shows the average detection-rate across all trials and benchmarks. The results are consistent with the laboratory experiment in showing that type of tag affects performance, with Excaliburs providing an average detection rate that is .28 greater than Flexwings and .47 greater than PVC cards across all objects and trials. The already well-known tendency [15, 49] for a tag’s readability to depend strongly on the materiality of the object it is attached to shows up here as well, with metal objects being more frequently missed than paper, cloth or plastic objects.

The way in which objects were carried and tags oriented also had a noticeable impact. Tags on objects kept close to the body (e.g., wallet, shirt, belt) are not read as well. Moreover, for all microbenchmarks there was variation in measured detection rate across trials, especially for objects inside the backpack which had detection rates with standard deviation greater than .1. This was likely due to shifts in the relative position of objects inside the backpack during and between trials (indeed, the backpack was unpacked and repacked between the second and third HP trial and between microbenchmarks). The results of the MT microbenchmark show that detection rate roughly doubled for objects not severely limited by the materiality problem or the way in which the object was carried. Dramatic increases in detection rate for MT were most likely due to a change in object orientation inside the backpack (e.g., for the 3-ring binder) or an additional tag being added in a much more readable

orientation (*e.g.*, in the jacket’s breast pocket as well as the side pocket). These results show promise for the multi-tag technique. However, we do not study the multi-tag technique further in this work as it complicates the study of end-user tag management as discussed in Section 4.2.

It is also clear from Table 2 that antenna configuration played a major role in determining whether or not a tag was detected, with CP providing an average detection rate across all tags and trials that is .21 better than the average HP rate. This is probably due to the fact that communication between Speedway readers and tags is *forward-link-limited* [21], meaning that detection depends on a tag’s ability to harvest adequate power from the reader signal. In our case, the additional antennas in the CP configuration increase the probability that a tag is in an antenna’s major detection field, making it more likely that the tag can harvest enough power to respond. Nevertheless, we chose to keep an HP-heavy deployment layout because the CP configuration was prohibitively intrusive and expensive to deploy widely throughout the building.

3. APPLICATIONS AND TOOLS

As an incentive for building occupants to use our system, we created a set of simple applications and tools for pervasive computing. All applications and tools featured secure web-based interfaces and were implemented using the Google Web Toolkit [16]. As such, interfaces were implemented in JavaScript with AJAX calls to retrieve dynamic content from an Apache Tomcat Servlet that used the Cascadia system [49] for queries on people, objects and events. We present a brief overview of the applications and tools below and discuss them in detail elsewhere [51].

Location-Aware Applications. Three applications provide location-aware services that facilitate an awareness of the location of one’s friends and one’s personal objects. Two of the applications, Rfidder and a Facebook application, report near-real-time information on the location of a user’s friends. Rfidder, shown in Figure 5(a), provides a stream of location updates for one’s friends in the form of a microblog with entered-room and exited-room events. The Facebook plugin displays a user’s friend list along with the most recently recorded location for each friend. A third application, the Tag Manager (see Figure 5(b)), allows users to view and edit metadata concerning their tagged objects and also displays the last recorded location for each object.

Tools for Managing Privacy and Personal Data. A set of tools facilitates application deployment and end-user management of personal privacy. The first tool is a Data Browser (see Figure 5(c)) which displays a table containing a dump of RFID data collected about the user. The Data Browser can be used to survey what data has been collected as well as to delete any data at the granularity of a TRE, an hour, or a day. An Access Control Interface allows users to create and manage context aware access control rules (*e.g.*, users in group “friends” can see my location during business hours when I am in the lounge) that are used to protect privacy in the Rfidder and Facebook applications. A third tool called the Place Manager assists users in the definition and labeling of meaningful places [18] within the building. This Place Definition tool presents a Google Map mash-up with an overlay of the building floorplan including antenna locations. Users can define new places by selecting a set of antennas corresponding to the place of interest. Thus, a user

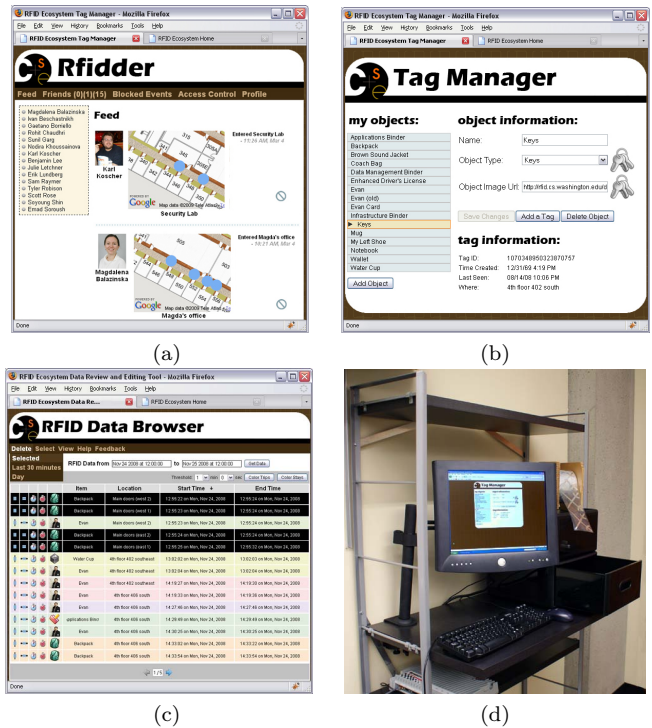


Figure 5: (a) Rfidder receives a stream of friend-location events. (b) Data Browser for review and optional deletion of a user’s data. (c) Tag Manager for end-user management of tags and data; integrates with (d) The RFID Kiosk: PC and RFID reader with one antenna installed above the plastic drawer to the right of the display.

could define a place called “My Office” as the set of antennas on either side of his office.

All applications and tools are accessible via links from a central RFID Ecosystem Homepage that also displays a dynamically generated plot showing how much data has been collected about the user over the past week. In order to better understand how users interact with RFID data, we logged all operations made through a tool or application which resulted in a database transaction (*e.g.*, add friend, query friend’s location, delete range of data).

4. FIELD STUDY SETUP

While laboratory and microbenchmark studies provide valuable insight into the performance of RFID equipment, they do not answer important questions regarding community use of a large-scale passive RFID system over a long period of time. For example, the character of the data generated and consumed by a community remains unknown as does the feasibility of end-user tag and data management and the performance of tags in a long-term everyday setting. To begin to answer such questions, we conducted a four week field study from November 2 to November 28, 2008. In the study, participants carried tags on themselves and on their objects. The participants also had access to the applications and tools described in Section 3 and periodically answered short surveys, but otherwise went about their everyday lives. Table 3 summarizes the high-level parameters of our study. We present the study methodology in detail throughout the remainder of this section.

Entity	Measure
Participants	67
Object types	19
Total tags	324
Duration (weeks)	4

Table 3: Parameters of our four-week user study.

4.1 Recruiting Participants

We recruited study participants from the community of faculty, staff, undergraduate and graduate students that occupy our building. During a 2-week recruiting period, we enrolled 67 participants (30 graduate students, 33 undergraduates, 2 faculty, 2 staff; 46 male, 21 female) who carried a total of 324 tags on 19 different types of objects. At the time of consent, all participants were educated on RFID, pervasive computing and the purpose of the study. Prospective participants were also presented with a demonstration of the applications and tools they would have access to if they chose to participate. As an additional incentive to enroll in our study, we offered participants \$30 for up to four months of participation with a chance to earn up to \$20 more for completing a series of short, web-based surveys (see Section 4.3). All participants were informed that they could withdraw from the study at any time and that they could also delete any subset of their data at any time using the Data Browser tool.

4.2 Tag Management

Upon consenting, participants were asked to carry one RFID tag as a “personal badge” and as many additional tags as they liked for use with personal objects. In order to determine whether end-users could effectively mount and manage their own tags, we employed a two-phase strategy for tag management. In the first phase, participants chose, mounted and managed their own tags directly; in the second phase we assisted participants in remounting their tags in an optimal fashion. We describe each phase in detail below.

On entrance to the study, participants chose tags, created a digital link between each tag and the object to be tagged, and then attached the tags to the objects. To facilitate this process, we built an *RFID Kiosk* in our laboratory. The Kiosk (see Figure 5(d)) was built into a computer desk and includes a supply of RFID tags, a PC with an RFID reader and a single antenna configured for short-range reading. Participants accessed the Tag Manager tool through the PC, which allowed them to create metadata for a particular object as well as to poll the Kiosk’s reader for a tag to associate with that metadata. We did not encourage that multiple tags be associated with a single object because this would complicate the process of optimizing a tag’s mounting on a given object. Later on, participants could access the Tag Manager from any web browser if they decided to edit their object metadata; however, creation or alteration of tag-object associations was only possible at the RFID Kiosk.

Participants were given guidelines as to how RFID tags should be attached to objects and to themselves for the best performance (*e.g.*, “a tag will be read infrequently or not at all if it is too near to your body or to a metallic object”). Throughout the first 3 weeks of study, participants were urged to follow the guidelines and reminded by email that they could directly see how well their tags were being read by checking the dynamic plot on the RFID Ecosystem Homepage or by reviewing the collected data with the Data Browser. In the last week of the study, we invited

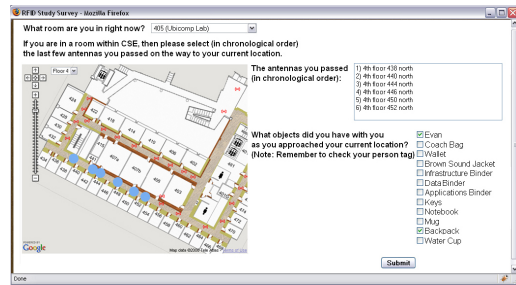


Figure 6: A web-based survey uses a Google map mashup to collect participants’ current location, a list of carried tags, and the sequence of antennas recently passed.

participants with less frequently detected tags back to our laboratory and helped them find a more optimal mounting technique. The results of this remounting experiment are covered in Section 5.3.3.

In addition to user-centered management of tags, we also wanted to investigate the durability of each type of tag. To this end, we distributed laminated and non-laminated versions of the Excalibur and FleXwing tags and made note of all incidents in which participants reported a tag as broken.

4.3 Collecting Ground Truth

To evaluate the our system’s ability to track people and objects, we needed some form of ground truth on the locations of tags at various times throughout the study. We collected two forms of ground truth information: survey results and card-key logs. Every weekday during work hours, a custom-built survey system prompted participants via email at random times to visit a survey web application at the soonest possible time. The web-based surveys asked where a participant currently was, what tagged objects he had with him and which antennas he passed on the way to his current location (if any); surveys were designed to take less than 20 seconds to complete. The timestamped survey responses allowed us to compute a sample-based representation of: 1) when and how often a participant was in the building or in a particular room, 2) when and how often a participant carried a particular RFID tag, and 3) approximately when a participant passed a particular antenna. We also administered a slightly longer exit survey to participants at the end of the study which asked a variety of questions on participants’ experience including tag wearing, application usage and privacy. In addition to surveys responses, we collected card-key logs for 4 members of our research group. This allowed us to pinpoint group meeting times for meetings that occurred in card-key access controlled rooms. We could then try to detect these meetings using a probabilistic view of the collected RFID data (see Section 5.4).

4.4 Cleaning RFID Data

It is well known that streams of passive RFID data are noisy and inconsistent [23, 31], with many false negatives as a result of RF interference, limited read range, tag orientation and other intermittent environmental phenomena. As such, we apply some basic data cleaning and compression algorithms to the collected data before analysis. In particular, we use the adaptive window-based cleaning algorithm from Jeffery et al. [23] to smooth over statistically insignificant gaps in the RFID data, obtaining a smoothed stream

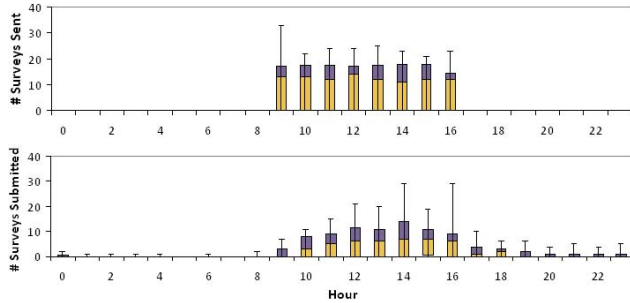


Figure 7: Quartiles for the distribution of the number of survey prompts (Surveys Sent) and responses (Surveys Submitted) by hour of the day over all workdays.

of TREs for each antenna. We then compress the smoothed data into STAYs. A STAY is a tuple that represents the length of time a tag stays at a particular antenna and has the following schema:

(tag ID, antenna ID, start time, end time, number of reads)

Smoothed TREs that occur consecutively are grouped into a single STAY record: *i.e.*, a tag must leave the read-range of an antenna for at least 1 second in order for a new STAY record to be generated. We found that this threshold works well in our environment but it is configurable.

5. FIELD STUDY RESULTS

In this section, we present the results of our study in terms of survey responses, infrastructure and applications utilization, system performance, and inference regarding detailed participant location and higher-level user events.

5.1 Survey Responses

Over the course of the study, we administered 2226 surveys and recorded 1358 responses. Though survey prompts were emailed during business hours, responses were submitted at all hours of the day. Figure 7 shows the distributions. On average, survey responses took 18 seconds to complete, this does not include the time taken to log in with username and password. Throughout the course of the study there were 7 incidents in which a participant self-reported an error in a survey response. In all cases these errors consisted in reporting that a particular tag was with the participant when in fact it was not. The effect of such errors on our results is that the reported detection-rate for that tag will be artificially lowered (see Section 5.3).

5.2 System and Application Utilization

The first key question that our study strives to answer is that of system utilization: How much data does a community of users generate and what does the data look like? What are the system utilization trends over time? Are users eager to adopt this new technology and are privacy concerns a significant barrier to adoption? We answer these questions in this section.

5.2.1 Data Rates

Table 4 summarizes the total and hourly data rates generated throughout the field study. Overall, the RFID infrastructure generates a small amount of data and that data

	Raw	Compressed
total data	1.5M TREs (52.95 MB)	38K STAYs (1.3 MB)
average rate	2295 TREs/hour	57 STAYs/hour
max rate	8408 TREs/hour	601 STAYs/hour
min rate	0 TREs/hour	0 STAYs/hour

Table 4: Total RFID data generated during the study and average, minimum, and maximum hourly data rates.

can be nicely compressed. On average, the system produced only 0.8MB of TRE data and a little over 20KB of STAY data per participant during the entire month. These data rates are much less than those generated by other sensor deployments (*e.g.*, audio sensors may sample at 16000Hz, acceleration sensors at 550Hz [6]). As such, RFID data may be potentially easier to manage. However, as we discuss in Section 5.4, post-processing the raw RFID data to infer more detailed location information causes a significant data blow up (easily one order of magnitude).

Figure 8 shows the overall system utilization in more detail. Figure 8(a) shows the quartiles for number of STAYs per hour across all working days and all days off (*e.g.*, weekends and holidays). Figure 8(b) shows the number of distinct personal badges read per hour. As expected, the data generated by an RFID infrastructure approximately matches the level of activity in the building. The figures show clear diurnal patterns with activity occurring later in the afternoon on days off. The number of distinct tags and the total data produced during a day also parallel each other.

Figure 9 shows the system utilization trends by day throughout the study in terms of total number of STAYs and total number of distinct tags read. These data rates also closely match the normal building utilization. The first day corresponds to November 2nd. We can see the patterns for the four weeks of the study with dips on weekends and on the November 11th, 27th, and 28th holidays. It is notable that the total number of STAYs again closely mirrors the total number of distinct tags read, suggesting that each tag read may contribute equally to the total number of STAYs generated. However, as we see in Section 5.2.2, many tags are seldom read at all throughout the course of the study, so we can conclude that the smaller number of tags which *are* read each day are responsible for the majority of data generated.

Overall, the activity trends measured by our RFID infrastructure are consistent with earlier mobility studies using WiFi [3, 25].

5.2.2 Data on People and Objects

To further characterize the data generated by a community of users we looked at the data for individual participants and for various object types. Among the 67 participants, each carried 1 personal badge tag and an average of 3 additional tags. The most tags carried by any participant was 21, while several participants carried only their personal badge.

Figure 10(a) shows a cumulative distribution function (CDF) of the number of STAYs generated by each personal badge, by each object, and by each participant (*e.g.*, the sum of STAYs for that participant’s tags). The plot shows that badges were read more frequently than objects, but that the most frequently read objects were detected as often as the most frequently read badges. The plot for participants’

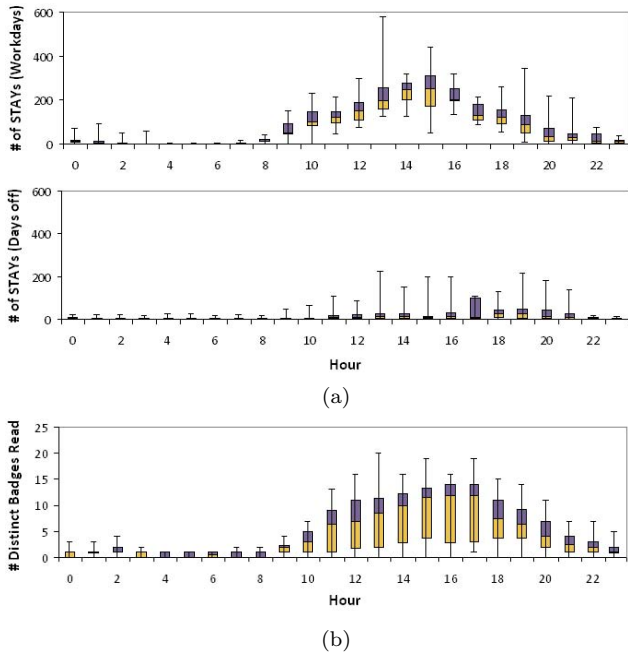


Figure 8: Quartiles for STAYs per hour over (a) all working days and all days off. The quartiles for (b) number of distinct badges read per hour over all days.

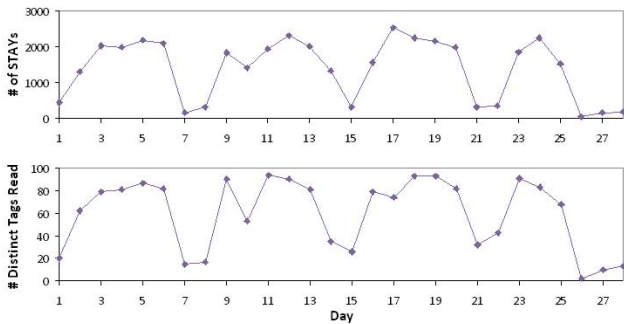


Figure 9: For each day of the study: total STAYs recorded and number of distinct tags read.

shows that the majority (65%) of participants produced less than 1,000 STAYs, while 10% of participants produced upwards of 3,000 STAYs. This means that the most active participants were detected moving through the building approximately 3 times as much as the median participant and 30 times as much as the least active participants. The plots for badges and objects show the same trend, although a little less pronounced. There are several reasons for these differences in level of activity between different people and their objects which we discuss below.

Figure 10(b) shows a CDF of the average and max length of STAYs generated by badges and objects. As one would expect, we see that participants tend to have slightly shorter average STAY times than their objects. This is because participants often leave objects behind in a location while they move. We also see that the top 10% of both objects and badges in terms of maximum STAY lengths have approximately the same maximum STAY length of 2.5 to 3 days (about the length of a long weekend). Deeper analysis of

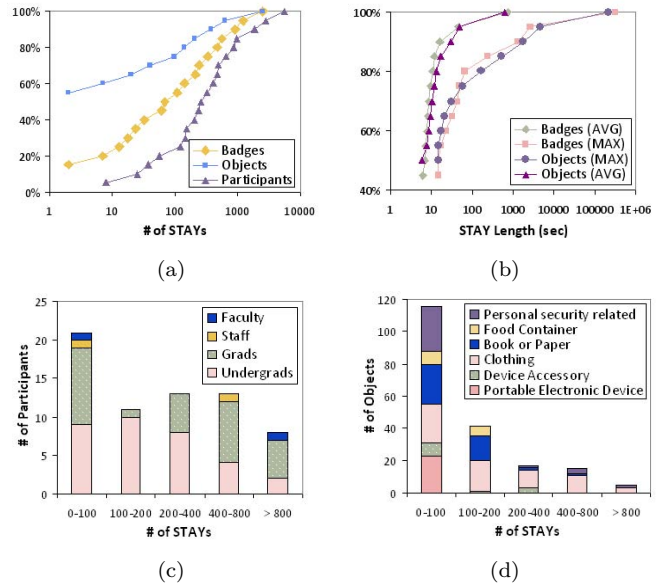


Figure 10: CDFs showing (a) STAYs generated by badges, objects and participants, and (b) Average and max STAY lengths for badges and objects. PDFs showing the number of STAYs generated by (a) different types of people and (b) different types of objects. Note the x-axis has a log scale for plots (a) and (b).

the data in connection with location survey data showed that these STAYs were all cases where a participant left one or more tags out on his or her desk in range of an antenna for several days at a time. Such events also created the most raw TRE data overall but are easily compressed into STAYs.

The probability density functions (PDFs) in Figure 10(c) and (d) show how the type of participant or object affects the amount of data generated. Faculty, staff, and graduate students tend to fall in one of two categories: they either produce significant amounts of data (more than 400 STAYs) or they produce almost no data (less than 100 STAYs). In contrast, undergraduate students tend to produce either almost no data or a small amount of data (200 to 400 STAYs). These groups tend to separate into large or small data producers because many participants simply wore their badges less often. In the exit survey, participants cited several reasons for this: 41 said they forgot to carry their tag on a weekly basis, 8 said they felt socially awkward carrying a tag, and 8 said they stopped wearing their tag because it didn't work well enough. Another reason for the split in data producers is that some participants just spent less time in the building. However, as we show below, there is a definite correlation between those who produced the most data and those who most actively used the applications. It is also likely that graduate students, faculty and staff produce more data than undergraduates because they occupy the upper floors of the building where there are many more antennas.

For objects, we see again a clear split between large and small data producers in all categories. These differences can partly be explained by how often participants move the objects. However, a second factor also affects the data generated by objects: the type of material. As the PDF shows, most portable electronic devices are metallic and so they are read quite infrequently even though we expect that participants carry them often. Meanwhile, non-metallic objects

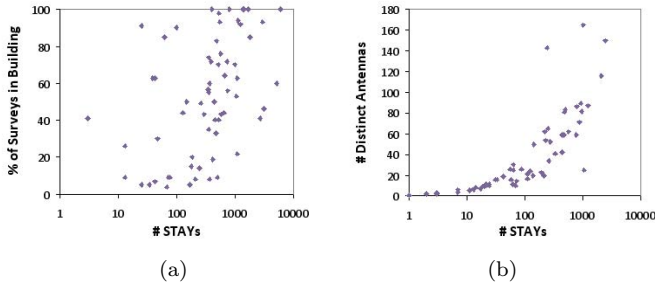


Figure 11: The impact of user mobility on total data collected. (a) The percent of “in building” survey responses by a participant generating a particular number of STAYs. (b) The number distinct antennas visited by a participant generating a particular number of STAYs. Note the log scale on the x-axes.

like clothing, books and papers are read most often. Thus, even with a variety of mounting strategies used across participants, the object material plays a very strong role in a tag’s readability.

In summary, there are at least four factors that can affect the amount of RFID data produced by a person or object: mobility, application utilization, how and when the tags are carried, and the material the tag is attached to. These results confirm our microbenchmarks in Sections 2.1 and 2.2.1 as well as the results of the next section. We now study how mobility and application usage correlate with system utilization.

5.2.3 Person and Object Mobility

Figure 11 shows scatterplots that depict the number of STAYs generated by each participant against either (a) the percent of survey responses in which that participant indicated he or she was in the building or (b) the number of distinct antennas visited by that participant. As expected, there is a clear correlation (with coefficient .38) between the approximate amount of time a participant spends in the building and the number of STAYs he generates. We also find that participants who occupy a larger portion of the building, as indicated by a larger distinct number of antennas visited, are likely to generate more STAYs (we computed a correlation coefficient of .71). We also examined the relationship between STAYs generated and average length of STAYs to find that there is neither a strong positive nor strong negative correlation (the coefficient is .16). The latter result is somewhat surprising and seems to indicate that independent of how much time some tags spend sitting in trackable areas, they are not actually more or less mobile.

5.2.4 Application Usage

Figure 12(a) shows application usage measured in operations (as described in Section 3). The plot shows that application usage decreased over time. This may be explained in part by the fact that students and faculty had an increased workload near the end of the academic quarter during the second half of the study. In addition, the exit survey revealed that 18 participants who stopped using applications felt that it was too much effort to sign-in to a web page, while 15 others didn’t have enough friends participating to keep Rfidder interesting. Although most participants explained that Rfidder and object tracking through the Tag Manager

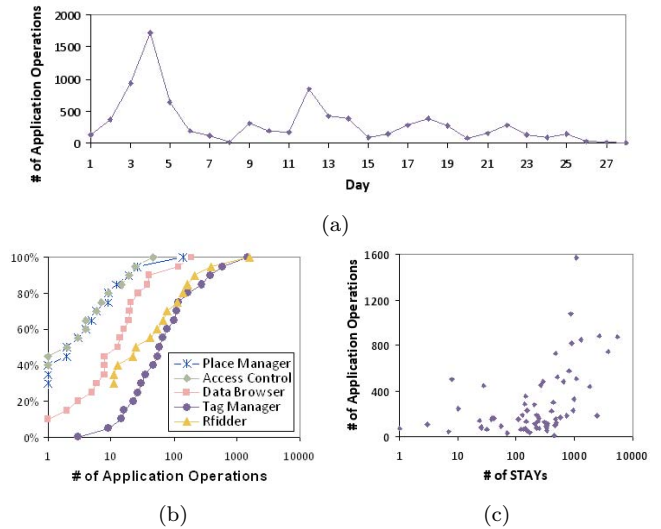


Figure 12: Application utilization: (a) over the course of the study, (b) as a CDF over all participants, broken down by application, and (c) per participant against data generated by that participant. Note the x-axis has a log-scale for (b) and (c).

were appealing only as novelties (only 13 found these applications useful for coordinating activities), these were still the most popular applications as shown in Figure 12(b). Another trend which emerged is that the amount of data generated by a participant is correlated with that participant’s level of engagement with the applications. Figure 12(c) plots application usage in operations against amount of data generated by that participant. The most active application users definitely generate the most data, although the overall correlation coefficient is only .37 for these two variables.

It is interesting to note that participants deleted very little of their data. Out of 67 participants, only 3 made any deletions. Moreover, the 4 deletions that were made correspond to only single STAYs (as opposed to an entire hour, day or trip through the building). Participants explained these operations as “a test to check that the system worked” in the exit survey. In addition, only 7 participants said they removed their tags for privacy reasons at least once throughout the study. This seems to indicate that for the first month of this study in a controlled, familiar setting, participants had little concern over the potential for a privacy breach. It may also be that participants did not understand the potential for privacy breach - even though the risks were carefully explained during the consent process. A more in-depth analysis of participants’ reactions to the applications and tools is published elsewhere [51].

5.3 System Performance

In addition to basic system utilization, we are also interested in studying the performance of an RFID infrastructure: What is the load distribution across different readers? How well are tags on various types of objects detected in practice? Are tags mounted by users detected as well as those mounted by experts? Are tags durable enough to last for long periods of time in an average user’s possession? We address these questions in this section.

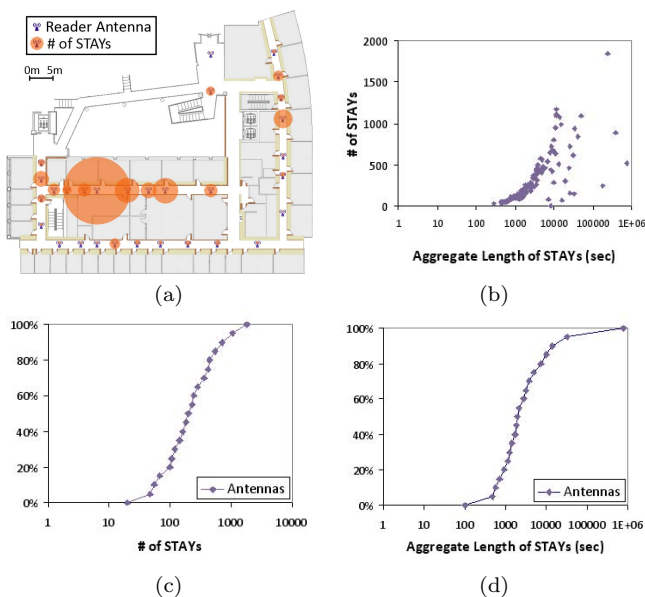


Figure 13: (a) Map of RFID deployment highlighting antennas where the most STAYs occur. (b) Plot of aggregate length of all STAYs against total number STAYs for each antenna. The CDF over all antennas of (c) STAYs registered and (d) the aggregate length of all STAYs. Note the x-axis has a log-scale for plots (b), (c), and (d).

5.3.1 Load Distribution

Figure 13(a) shows a map of the 3rd floor of the Allen Center that highlights regions where the most STAYs occur. Consistent with previous location-tracking studies [3], certain locations and thus antennas are “hot spots” that generate significantly more data than others; the two hot spots in the map are outside the student store and outside a lab where several of the study participants spend time.

Looking more closely, we see trends in Figures 13(c) and (d), the CDFs over all antennas of total STAYs occurring and total aggregate length of all STAYs. Both curves exhibit clear logarithmic trends. The first CDF shows that there are approximately two orders of magnitude difference between the number of STAY records at the least and most visited antennas. The second shows that there is an incredible four orders of magnitude difference between aggregate length of STAYs. These plots illustrate two different types of hot spot: one is a location that mobile tags pass frequently, and another is a location where tags dwell for very long periods of time. The data in Figure 13(b) is correlated with coefficient .29 and shows that while a single antenna can have both of these characteristics, many also have a much longer aggregate STAY length relative to the number of STAYs.

5.3.2 Detection Rate

We use the timestamped survey responses to compute approximate read probabilities for tags used in the study. Recall that the 1398 survey responses each include the participant’s current location, antennas passed on the way to that location (if any) and the list of tags carried while passing those antennas.

An important challenge when processing the survey data is that participants were sometimes working for a long time at their desk when they got the survey. In those cases, the survey responses correspond to data collected much earlier

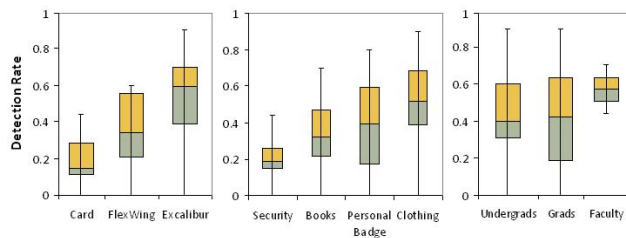


Figure 14: Quartiles for detection rate of (a) different types of tags (b) different types of objects (c) different types of people.

than the survey timestamp. At other times, participants passed a series of antennas without being detected in which case there is no data corresponding to the survey response. For this reason we chose the following approach to compute the detection rates from the survey data which underestimates the detection rates.

We compute approximate detection rates by counting antenna “hits” and “misses” for each tag. To do so we take each survey response and search backwards from the timestamp to find the first STAY for each carried tag. If a tag’s first STAY did not begin within 1 hour of the timestamp then we add N misses for that tag where N is the number of antennas passed. We basically assume that if no data is found within this time window, the participant was not detected at all during the reported trip. If a tag’s first STAY did occur within 1 hour of the timestamp then for each antenna reported in the survey, we search for a STAY occurring at that antenna at most 15 minutes prior to the last STAY record. We add one hit for each such STAY found and one miss when no such STAY is found. The 15 minute window serves to limit the search for hits to the most recent trip by the participant. It is possible, though not common in our deployment, that a participant passes multiple times under the same antenna within such a short period of time. We compute an approximate detection rate as hits divided by the number of hits and misses. It should be noted that not all survey data is usable because some responses indicated that participants were not currently in the computer science building or that no tags were currently carried. We also threw out survey data for tags that had a total number of antenna hits and misses less than 10 across all survey responses. Overall, we use 956 surveys to compute detection rates. We present this approximate detection rate data below.

Figure 14 shows the detection rate aggregated across tag types, object types, and class of participants. As expected given our micro-benchmarks (see Section 2.1), the Excalibur tags achieve the best performance among the three types of tags. These tags are also have the largest antenna of the three models. The figure shows significant difference between types of objects. Clothing articles achieve the best performance while security related objects including keys, wallets and purses have the worst performance. Metallic objects do not appear on the graph as there were not enough hits in survey data for these objects. The differences between objects are due to a combination of object material and tag mounting. We further explore the effect of tag mounting next. Finally, badges exhibit much larger variance than tags attached to objects indicating that different people carry their badges very differently and these choices

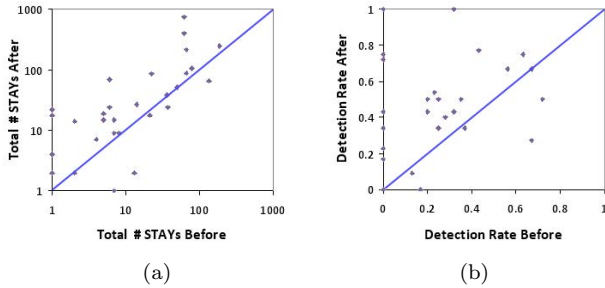


Figure 15: Plots showing the (a) total STAYs collected and (b) detection rate for each remounted tag before and after remounting.

	Broken	Lost	Returned
EX	2	1	0
EX (laminated)	0	0	0
FW	3	1	0
FW (laminated)	1	0	0
PVC	0	1	2

Table 5: Summarizing results for tag durability.

significantly affect the detection performance. Interestingly, different groups of participants do not systematically get better read probabilities than others (the faculty group was simply smaller than the other ones).

5.3.3 Tag Mounting

To explore the effect of tag mounting by users as opposed to mounting by experts we let participants mount and manage their own tags for the first 3 weeks of the study. In the last week of the study we requested that participants carrying tags with very low detection rates come to our lab to try an alternate tag mounting strategy for a week. Figure 15 shows the total STAYs collected and detection rate for each remounted tag before and after it was remounted. The plots show that although some tags remained consistently difficult to read, the readability improved for the majority of remounted tags. This suggests that in an RFID-based pervasive computing system it is not enough to give basic tag mounting guidelines. Either experts must assist the tag mounting process or users must be sufficiently incentivized (possibly through compelling applications) to find an effective tag mounting strategy.

Regarding the durability of tags and the ability to endure use by a variety of participants, we observed a number of events which are summarized in Table 5. A total of 6 tags were broken, most of which were unlaminated tags that had been physically torn or creased. One laminated FleXwing was broken after being mangled inside a backpack. Three tags (one of each type) were lost, and 3 PVC cards were returned or thrown away during the study because the participant felt they didn’t work well enough.

5.4 Inferring Information from RFID Data

As the above results show, raw RFID data can contain significant uncertainty about tag locations: RFID readers may cover only a subset of the areas of interest (*e.g.*, no offices in our deployment) and detection rates can vary significantly based on object material and mounting strategy. A commonly used technique to facilitate application development given such uncertain raw data is to build and expose

a probabilistic model in which sensor locations and errors are abstracted away. The application logic then consists of queries on top of this probabilistic view [10, 9].

As in previous work on event queries over RFID data [28, 29, 39, 50], we build the probabilistic view using a Hidden Markov Model (HMM) [36] and apply (*Bayesian filtering and smoothing* [14]) using a particle filter based inference technique [11]. The output of this process is a stream in which each timestep contains a *distribution* over possible locations for a tag: *e.g.*, Alice was in her office with probability 0.2, in her neighbor’s office with probability 0.1, and in the corridor with probability 0.7.¹ We materialize the resulting probabilistic streams to disk.

We create a probabilistic view for all tags in the study over the 4 weeks. First, we tried simply to run the particle filter non-stop over the four weeks. This approach, however, lead to a significant amount of low-quality data due to long periods of tag immobility: in the absence of RFID data, the view is highly uncertain of a tag’s location, producing unusable results. In a second approach, which we report on here, we only ran the particle filter over periods of time in which TREs occurred. We call these periods *slots*.

We evaluate the quality of the probabilistic view by comparing the survey results to the materialized view for the appropriate tag. We use two comparison techniques T1 and T2. In T1, for each survey, we look at the tag’s probabilistic view in a time window centered on (*i.e.*, preceding and following) the survey response time. If the probability that the tag is in the correct location (according to the survey) is greater than a threshold, we consider the view to be correct. In T2, we only consider the last time slot before or containing the survey response time. T2 assumes that a tag does not move in between time slots.

For each technique, Figure 16 shows (1) the fraction of surveys where the view correctly identified a tag’s location, (2) the fraction of surveys when the view was wrong, and (3) the fraction of surveys for which no RFID data for the correct floor was registered within the given window (these surveys are indicated as “not applicable”). For clarity, Table 6 shows the corresponding detection rate values. The plots for probability threshold 0.1 correspond to instances where the view is confident that a tag is located in the given room (probabilities are low in general due to location ambiguity in cases where a tag could be in one of three adjacent rooms or in the corridor; spreading the probability values across several locations). The plots for threshold of 0 show instances where the view only localizes the tag to the correct vicinity (*e.g.*, corridor or floor).

For T1, on average, the probabilistic view places the tag in the correct vicinity most of the time, but in the correct room less than half the time. Looking only at the most recent slot before a survey (T2) instead of the time around a survey (T1) yields worse results simply because fewer slots are considered. For example, if the system does not notice when a tag enters a room but correctly identifies that the tag left, T1 will report the correct location but T2 will not. Overall, *continuous*, fine-grained tracking in an RFID deployment is thus difficult. There are times when the system fails to capture the correct tag movements even after smoothing the sensor data.

¹The probabilistic view includes correlation information between consecutive timesteps, but we ignore these correlations in this paper.

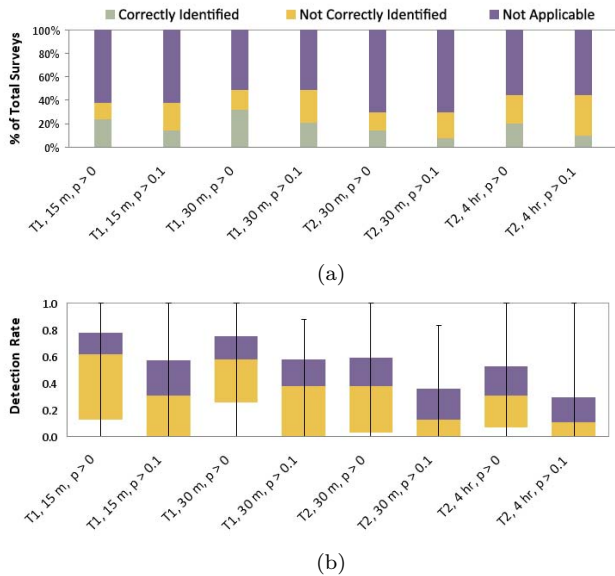


Figure 16: Detection rate results for probabilistic view methods in terms of (a) Surveys correctly identified, and (b) quartiles.

Technique	Win	Threshold	Detection rate
T1	15 m	$p > 0.0$	0.62
T1	15 m	$p > 0.1$	0.37
T1	30 m	$p > 0.0$	0.65
T1	30 m	$p > 0.1$	0.43
T2	30 m	$p > 0.0$	0.48
T2	30 m	$p > 0.1$	0.24
T2	240 m	$p > 0.0$	0.45
T2	240 m	$p > 0.1$	0.21

Table 6: Average detection rate.

The overall low detection rates are partly due to the fact that performance varies considerably among tags depending on the properties of the tag and tagged object (recall Figure 14). Figure 16(b) shows the quartiles for detection rate of different types of tags. For the best 25% of tags, the probabilistic view places the tag in the correct vicinity upwards of 80% of the time and in the correct room upwards of 60% of the time. These results are thus encouraging: a probabilistic view can smooth away significant errors and ambiguity. This approach, however, clearly fails with large error rates, suggesting that a practical deployment must take further measures to ensure tags have sufficiently high read rates.

In a second experiment, we studied the utility of the probabilistic view for inferring simple, higher-level events. Specifically, we defined a simple meeting event: two people meet if they appear in the same physical location within at most five minutes of each other. We then used the card-key logs for the 4 researchers to pinpoint 10 meetings. Of these 10, 6 were correctly identified using the probabilistic views, thereby echoing the results of the room-level location experiment.

Overall, using probabilistic views over RFID data in large-scale, continuous tracking scenarios is thus necessary and promising but it raises important challenges in practice. As we showed above, probabilistic views help only when the data is already of sufficiently high quality. For alerting applications, the process must also be sufficiently fast to keep

Prob. threshold	0	0.05	0.1	0.2	0.3	0.6
Data ratio	13.54	2.51	1.69	1.05	0.74	0.25

Table 7: Probabilistic view data blow-up. The ratio is computed as the number of records in the materialized probabilistic view over the number of TREs when counting only data above a given probability threshold.

up with the input data (although techniques to achieve this goal exist [27, 34]). Finally, the per-timestep location distributions in probabilistic views can cause a data blow-up. Table 7 shows the data blow-up for the month long view over all participants. The entire view uses more than 13 times the space of the raw data! Eliminating data with low probability (less than 0.05), however, quickly reduces the overhead to just a small factor. A study of the impact of such compression techniques on performance is an interesting area of future work.

6. ADDITIONAL LESSONS

The process of deploying a large-scale RFID system in a busy, shared public building was difficult and time consuming. We also learned a great deal about effective deployment practices and found several very helpful insights and strategies for working with key problems.

6.1 Antenna Deployment Strategy

Deploying antennas in many locations was difficult due to health and safety concerns and regulations, building management and facilities issues, and even aesthetic issues. After experimenting with many alternatives for hiding the HP antennas (*e.g.*, placing them above ceiling tiles or inside plenum-rated enclosures, painting them to match the wall), we found the best strategy was to “hide the antennas in plain sight”. The antennas are actually quite non-descript and are often assumed to be yet another piece of a modern building’s complex technical infrastructure. Many building occupants didn’t notice the antennas until we announced their presence months after they were installed.

6.2 Web-Based Diagnostic Tools

We implemented a set of web-based diagnostic tools to help in the early stages of system deployment and debugging. These tools included a Google map mashup of readers, antennas, and live tag readings; system monitors that send SMS updates and alerts to researchers regarding data collection and system status; and data visualization tools that automatically plot aggregate data in a web interface. These tools greatly facilitated our work and expedited both the deployment and debugging processes.

6.3 Building-in Privacy Assurances

From the beginning, we designed our study to allow users a large amount of direct control over their own data and over what information is disclosed about them. We also took special care to make sure the user’s privacy control and data management interfaces were understandable and usable. This effort proved very helpful when we were recruiting users for the field study. Many participants were apprehensive about the study at the time of consent and immediately became much more comfortable and interested on seeing the privacy controls available to them.

7. RELATED WORK

Passive RFID has been used in a variety of pervasive computing experiments over the last decade. An early paper on the subject by Want et al. [48] spurred a variety of projects that use RFID to detect interactions with objects [12, 13, 20]. This direction has culminated in a substantial amount of work on RFID-based human activity recognition [31, 35, 42]. Other work from the pervasive and mobile computing community has investigated the design and implementation of fine-grained location systems that use RFID [30, 32, 33]. Our work does not currently include detection of interaction with objects or precise tag localization. Instead we focus on systems, techniques and applications that use coarse-grained (*e.g.*, room-level) localization. A variety of recent work has also focused on RFID in this capacity [4, 5, 24]. Most similar to our vision is MSR India’s SixthSense system for RFID-based Enterprise Intelligence [38]. However, recent work on SixthSense has focused on automatic inference of object type and ownership information while we explore more explicit, user-centered mechanisms for tagging and management of objects in an RFID system. To our knowledge, we also study RFID using a much larger scale deployment than any past pervasive or mobile computing research project.

Other work from the pervasive computing and networking communities has evaluated the performance of passive UHF RFID in carefully controlled laboratory studies. Ramakrishnan and Deavours [37] present a set of benchmarks for evaluating RFID performance in a variety of environmental conditions (including near metal and water). Hodges et al. [19] present a sophisticated, robotically automated technique for assessing the operating range of an RFID system. Buettner and Wetherall [7] measured EPC Gen 2 performance with a software defined radio and found opportunities for improvement in that standard’s physical and MAC layers. All these studies contribute a great deal to our understanding of RFID performance and its influencing factors. However, they are all constrained to a laboratory setting and thus omit many of the conditions affecting pervasive RFID deployments such as ambient RF interference, tag mounting, movement and user behavior patterns. In a previous study [49] we looked at the performance of EPC Gen 1 RFID technology in a shorter, smaller scale pervasive deployment but did not recruit non-researcher participants nor did we collect as much ground truth information.

A variety of other projects have studied large-scale technology deployments like ours. Early projects like ParcTab [47] and Sentient Computing [1] were quite similar in that they offered indoor location aware services to a community of users who each carried a tracking device. Our work differs chiefly in that these projects used active, battery-powered tracking devices and were most appropriate for ubiquitous tracking of people and not objects. More recent work has measured sensor performance and mobility in both home and public spaces. Logan et al. [31] evaluated the effectiveness of various sensors in a long-term human activity recognition study and found that RFID was among the worst performing sensors. However, in this case the researchers were using a different type of RFID at short range to detect interaction with objects which differs from our goals. Measurement studies on wireless networks have found similar patterns in daily human mobility as evident in network utilization [3, 25]. Wireless network data is different in character than RFID data though because there are

more data sources (*e.g.* objects as well as people, and many more readers than access points) and because each RFID data source is less reliable than an access point.

Numerous consultancies, corporations and other stakeholders have produced case studies on the performance of real-world RFID deployments. Much of this work is focused on the supply chain and contactless payment domains which pose problems that are only peripherally related to our work. However, there has also been increased interest in passive UHF RFID for tagging and tracking patients, personnel and equipment in hospitals [43, 40, 45]. Though work on RFID in hospitals typically emphasizes the requirements for cost effective deployment and business integration, it also raises many challenges we share: an indoor environment, tags on highly mobile people and objects, and increased chance of RF interference. As such, the hospital domain is an interesting area to watch and one to which some of our results may transfer.

8. CONCLUSION

In this paper, we presented the results from a four week study of a building-wide EPC Gen 2 UHF RFID deployment with tens of participants and hundreds of tags. We analyzed the system from the perspective of the data generated and consumed by the community which uses it as well as the performance and reliability of tags carried on various participants, objects and tags. We also experimented with building and using a probabilistic model-based view to enhance the utility of large-scale RFID data sets.

We found that our RFID deployment produces a very manageable amount of data overall, but with orders of magnitude difference among various participants and objects. Among other factors, the amount of data produced by participants was influenced by their application usage. Compelling applications may thus drive the future utility of such deployments.

We found that the tag detection rates were relatively low and with a large variance across tags. Many end-users also needed expert help in properly mounting their tags to objects. We also found that probabilistic model-based views seem a reasonable approach to smooth errors and gaps in the data generated by such deployments. However, we find that building and exploiting such views is challenging: care must be put into building these views in order to get good results and these methods impose significant overhead. A materialized model-based view can easily increase the data size by more than a factor of 13 (though compression can be achieved by throwing away the lowest probability data).

Overall, passive RFID deployments present a promising sensor infrastructure for a variety of location-oriented applications, but exploiting such deployments in a large-scale pervasive computing environment presents considerable technical challenges.

9. ACKNOWLEDGEMENTS

We like to thank Garret Cole, Sam Raymer, and Julia Schwarz for their excellent work in support of the RFID Ecosystem project and its applications. This work was partially supported by NSF Grants IIS-0713123, CNS-0454425 and CRI-0454394 as well as by the UW College of Engineering. We would also like to thank the anonymous reviewers for their helpful comments on early versions of this paper.

10. REFERENCES

- [1] Addelee, M. et al. Implementing a Sentient Computing System. *IEEE Computer*, 34(8):50–56, 2001.
- [2] EPCglobal UHF Class 1 Gen 2. <http://www.epcglobalinc.org/standards/uhfc1g2>, 2004.
- [3] Balazinska, M. and Castro, P. Characterizing Mobility and Network Usage in a Corporate Wireless Local-Area Network. In *MobiSys 2003*, pages 303–316, May 2003.
- [4] J. E. Bardram. Applications of context-aware computing in hospital work: examples and design principles. In *SAC '04*, pages 1574–1579, 2004.
- [5] Borriello, G. et. al. Reminding about Tagged Objects using Passive RFIDs. In *Proc. of the 6th Ubicomp Conf.*, pages 36–53, Sept. 2006.
- [6] Brunette, W. et al. Some sensor network elements for ubiquitous computing. In *IPSN '05*, 2005.
- [7] Buettner, M. and Wetherall, M. An Empirical Study of UHF RFID Performance. In *MobiCom 08*, pages 223–234, Sept. 2008.
- [8] Computerworld. Procter & Gamble: Wal-Mart RFID effort effective. <http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=284160>, Feb. 2007.
- [9] Deshpande, A. and Madden, S. MauveDB: Supporting Model-based User Views in Database Systems. In *Proc. of the 2006 SIGMOD Conf.*, pages 73–84, 2006.
- [10] Deshpande, A. et al. Using Probabilistic Models for Data Management in Acquisitional Environments. In *Proc. of the 2nd CIDR Conf.*, pages 317–328, Jan. 2005.
- [11] Doucet, A. et al., editor. *Sequential Monte Carlo methods in practice*. 2001.
- [12] Fishkin, K. et al. I Sense a Disturbance in the Force: Unobtrusive Detection of Interactions with RFID-tagged Objects. In *UbiComp 2004*, pages 268–282, Oct. 2004.
- [13] C. Floerkemeier and F. Mattern. Smart Playing Cards - Enhancing the Gaming Experience with RFID. In *PerGames 2006*, pages 79–88, May 2006.
- [14] Gelman, A. et al. *Bayesian Data Analysis*. Chapman and Hall/CRC, 2003.
- [15] Glover, B. and Bhatt, H. *RFID Essentials*. O'Reilly, 2006.
- [16] Google Web Toolkit - Google Code. <http://code.google.com/webtoolkit/>, 2008.
- [17] Hahnel, D et al. Mapping and Localization with RFID Technology. In *ICRA*, pages 1015–1020, 2004.
- [18] Hightower, J. et al. Learning and Recognizing the Places We Go. In *UbiComp 2005*, volume 3660, pages 159–176, 2005.
- [19] Hodges, S. et al. Assessing and Optimizing the Range of UHF RFID to Enable Real-World Pervasive Computing Applications. In *Proc. of the 5th Pervasive Conf.*, pages 280–297, 2007.
- [20] S. Hsi and H. Fait. RFID enhances visitors' museum experience at the Exploratorium. *Commun. ACM*, 48(9):60–65, 2005.
- [21] Impinj. RFID Communication and Interference. <http://www.impinj.com/WorkArea/>, 2007.
- [22] Speedway Product Brief. http://www.impinj.com/files/MR_SP_PB_00001_SpeedwayProductBrief.pdf.
- [23] Jeffery, S. R. et al. Adaptive Cleaning for RFID Data Streams. In *Proc. of the 32nd VLDB Conf.*, pages 163–174, Sept. 2006.
- [24] Kittredge, L. and Borriello, G. An RFID-based Game to Encourage Social Interaction. *PerGames 2006*, May 2006.
- [25] D. Kotz and K. Essien. Analysis of a campus-wide wireless network. *Wirel. Netw.*, 11(1-2):115–133, 2005.
- [26] KSW Microtec. <http://www.ksw-microtec.de/>.
- [27] Kwok C. et. al. Real-time Particle Filters. *Proc. of the IEEE*, 2(92):469–484, 2004.
- [28] Letchner, J. et al. Challenges for Event Queries over Markovian Streams. *IEEE Internet Computing*, 12(6):30–36, 2008.
- [29] Letchner, J. et al. Access Methods for Markovian Streams. In *Proc. of the 25th ICDE Conf.*, Apr. 2009.
- [30] Liu, X. et. al. Ferret: RFID Localization for Pervasive Multimedia. In *Proc. of the 8th Ubicomp Conf.*, pages 422–440, Sept. 2006.
- [31] Logan, B. et al. A Long-Term Evaluation of Sensing Modalities for Activity Recognition. In *UbiComp 2007*, pages 483–500, Sept. 2007.
- [32] Nemmaluri, A., et al. Sherlock: automatically locating objects for humans. In *MobiSys 2008*, pages 187–198, 2008.
- [33] Ni, L. M. et al. LANDMARC: Indoor Location Sensing using Active RFID. *Wirel. Netw.*, 10(6):701–710, 2004.
- [34] Nie, Y. et al. Probabilistic Inference over RFID Streams in Mobile Environments. In *Proc. of the 25th ICDE Conf.*, pages 1096–1107, Apr. 2009.
- [35] Philipose, M. et al. Inferring Activities from Interactions with Objects. *IEEE Pervasive Computing*, 3(4):50–57, 2004.
- [36] Rabiner, L. R. A tutorial on hidden Markov models and selected applications in speech recognition. pages 267–296, 1990.
- [37] Ramakrishnan, K. M. and Deavours, D. D. Performance Benchmarks for Passive UHF RFID Tags. In *Proc. of the 13th GI/ITG Conf. on Measurement, Modeling, and Evaluation of Computer and Communication Systems*, pages 137–154, Mar. 2006.
- [38] Ravindranath, L. et al. SixthSense: RFID-based Enterprise Intelligence. In *MobiSys 2008*, pages 253–266, 2008.
- [39] Ré, C. et al. Event Queries on Correlated Probabilistic Streams. In *Proc. of the 2008 SIGMOD Conf.*, pages 715–728, June 2008.
- [40] RFID Update. Reva Launches RTLS Solution Based on Passive RFID. <http://www.rfidupdate.com/articles/index.php?id=1774>, Apr. 2009.
- [41] RFIDentics Products. <http://www.rfidentics.com/products.html>.
- [42] Smith, J. R. et al. RFID-Based Techniques for Human-Activity Detection. *Communications of the ACM*, 48(9):39–44, Sept. 2005.
- [43] Spyglass Consulting. Trends in RFID 2008. http://www.spyglass-consulting.com/spyglass_whitepaper_RFID.html, July 2008.
- [44] SQL Server 2008 Overview. <http://www.microsoft.com/SQL>.
- [45] van der Togt, R. et al. Electromagnetic Interference From Radio Frequency Identification Inducing Potentially Hazardous Incidents in Critical Care Medical Equipment. *Journal of the American Medical Association*, 299(24), June 2008.
- [46] Wang, F. and Liu, P. Temporal management of RFID data. In *Proc. of the 31st VLDB Conf.*, pages 1128–1139, Sept. 2005.
- [47] Want, R. et. al. An Overview of the ParcTab Ubiquitous Computing Experiment. *IEEE Personal Communications*, 2(6):28–33, Dec 1995.
- [48] Want, R. et. al. Bridging Physical and Virtual Worlds with Electronic Tags. In *CHI*, pages 370–377, 1999.
- [49] Welbourne, E. et. al. Challenges for Pervasive RFID-based Infrastructures. In *PERTEC 2007*, pages 388–394, Mar. 2007.
- [50] Welbourne, E. et. al. Cascadia: A System for Specifying, Detecting, and Managing RFID Events. In *MobiSys 2008*, pages 281–294, June 2008.
- [51] Welbourne, E. et. al. Building the Internet of Things Using RFID: The RFID Ecosystem Experience. *IEEE Internet Computing*, May 2009.