Discovering Semantically Meaningful Places from Pervasive RF-Beacons

Donnie H. Kim, Jeffrey Hightower[†], Ramesh Govindan[‡], Deborah Estrin
UCLA CSD, Intel Labs Seattle[†], USC CSD[‡]
{dhjkim, destrin}@cs.ucla.edu, jeffrey.r.hightower@intel.com[†], ramesh@usc.edu[‡]

ABSTRACT
Detecting visits to semantically meaningful places is important for many emerging mobile applications. We present PlaceSense, a place discovery algorithm suitable for mobile devices that exploits pervasive RF-beacons. By relying on separate mechanisms to detect entrance to and departure from a place and buffering overlapping data for subsequent visits, it is more robust than the state-of-the-art, especially in detecting short visits, places where people are mobile, or where inconsistent beacons are prevalent due to interference. We experimentally evaluate PlaceSense's effectiveness in discovering semantically meaningful places, and compare with other approaches that use coordinates or RF-beacon fingerprints. Our results demonstrate that PlaceSense correctly discovers 92% (compared to between 28% and 65% for previous work) of the visited places and accurately detects their entrance and departure times from both real-life and scripted data sets.

Author Keywords
Place Learning, Location, Beacon Traces

ACM Classification Keywords
H.5.2 Information interfaces and presentation: Miscellaneous

General Terms
Algorithms, Experimentation, Human Factors

INTRODUCTION
Recent advances in location technology and mobile devices have opened the door for many interesting mobile applications. An increasing number of mobile devices are now capable of locating themselves based on different technologies including satellite, mobile telephony, and 802.11 (Wi-Fi). These technologies offer different opportunities and limitations; the Global Positioning System (GPS) provides worldwide coverage except in buildings and underground, while technologies based on Wi-Fi and cellular signals can potentially provide relatively coarse location estimates anywhere wireless internet and voice services are available [11]. Several commercial products [18, 17] have shown that a mixture of GPS and RF-beacon-based location can allow a device to compute its position ubiquitously and with high availability. The raw coordinates provided by these location systems enable location-aware applications such as navigation and emergency response that require absolute locations for only a short period of time.

Many emerging mobile applications, however, can benefit from *places*, which are colloquially labeled representations of locations such as "Home", "My Office", or "Joe's plumbing store", instead of a series of raw coordinates. Places can directly support applications ranging from simple location-aware reminders to personalized mobile searches based on place preference [13]. Automated place discovery can help studies of human spatial and temporal behavior, which have historically depended on laborious manual recording or direct observation [4]. More generally, place information can help device intelligent algorithms for applications that capture and share a user's context [3]. Such applications collect streams of location, image, acoustic, or text data to continuously understand and record people's activity and mobility patterns, report information about their environment (e.g. traffic, pollution levels), or exchange whereabouts among friends and family [15, 5, 16]. A place discovery technique can help schedule data collection only when it matters and can effectively summarize the collected data.

Place learning algorithms attempt to find a locale that is important to an individual user and carries a semantic meaning. In this paper, an important locale is defined as a place where the user spends a substantial amount of time and/or visits frequently. Typically, the input to a place learning algorithm is a sequence of time-series sensor data (e.g. GPS coordinates, CDMA/GSM cell towers, Wi-Fi Access Point MAC addresses, etc.) and its output is a sequence of tuples (date, enter_time, leave_time, place name). A number of interesting place learning algorithms have been proposed both based on coordinates provided by location systems (GPS or Place Lab [11]) or on raw RF-beacon (Wi-Fi Access Point or cell tower) fingerprints.

In this paper, we present the PlaceSense algorithm which is an evolutionary step in this line of research. PlaceSense collects Wi-Fi or cell tower radio fingerprints by scanning the environment, detects place entrance and departure using multiple successive scans. It cannot, of course, automatically assign semantically meaningful names to places, but

UbiComp 2009, Sep 30 – Oct 3, 2009, Orlando, Florida, USA.

accurately identifies place entry and exit to enhance recognizing when places are visited. It improves upon prior work in two ways: (1) it is more robust since it uses separate mechanisms for entrance and departure, and (2) is more responsive since it uses history information to rapidly detect subsequent visits. Newly-seen beacons trigger entrance determination, but the algorithm robustly avoids beacon instability (caused by weak beacons in hallways, for example). On the other hand, departure is determined by the disappearance of *representative beacons* seen in a place, but this disappearance is carefully made to avoid false positives.

In addition to describing PlaceSense, we discuss possible errors in finding places, describe new place discovery evaluation metrics, and demonstrate PlaceSense's effectiveness with a comparative evaluation to two published place algorithms based on coordinates and RF-beacon fingerprints. To evaluate our algorithm, we gathered radio traces (GPS, Wi-Fi, and GSM cellular) from three volunteers following scripted visits (for accurate ground-truth) to multiple places and as they went about their normal routines for four weeks. Each volunteer collected radio traces and kept a written diary of places they visited. Using these two sets of data, we compare PlaceSense's effectiveness in discovering visited places and the accuracy of the detected entrance and departure time with previous work, and demonstrate that it outperforms the other methods in real-life applications. Intuitively, PlaceSense performs better than algorithms that use geographic coordinates since semantically meaningful places are often indoors where current location systems suffer in continuously providing accurate positions. Furthermore, it outperforms RF-beacon fingerprint based algorithms by being more robust to inconsistent beacons, and more responsive in detecting short visits.

RELATED WORK

Place learning algorithms can be divided into two classes based on the characteristics of the source location data: geometry and fingerprint. We discuss each class and describe in more detail the techniques we use to compare PlaceSense.

Geometry-based Place Learning

Geometric algorithms produce coordinates, circles, or polygons to describe places the algorithm believes are significant to the users. These algorithms take a history of locations (e.g., from GPS devices) and find locations where the person stays for significant periods of time. The algorithms vary based on the type of sensor data and the specific clustering algorithm they use. Coordinate-based systems include GPS and RF-emission based coordinate inferring systems [11, 18, 17], which typically provide location as a pair of latitude and longitude.

Identifying densely clustered regions from the geometric coordinate trace is basically a clustering problem. However, standard clustering algorithms can include transient or erroneous coordinates, making the clusters unnecessarily large and imprecise [8]. More fundamentally, GPS-based place learning systems cannot accurately identify indoor places. Early efforts in place learning with GPS used loss of signal to infer the location of important indoor places. Marmasse et al. [14] identify a place as a region, bounded by a certain fixed radius around a point, within which GPS disappears and then reappears as when a user enters and leaves a building. This approach is sufficient to identify indoor places that are smaller than a certain size (e.g. a home), but cannot identify distinct places within larger indoor places (e.g. multi-floor buildings), and is prone to generating false positives caused by the many possible outdoor GPS shadows (a recent study [11] and our own experiments show that GPS coverage is available only 5-30 % of the time on average for a device carried by users during a typical day). A similar but improved approach was proposed by Ashbrook et al. [2]. Sets of important coordinates are identified as those at which the GPS signal reappears after an absence of 10 minutes or longer. These sets are then clustered into "significant locations" using a variant of the k-means clustering algorithm. Toyama et al. presented a variation of this work that employs multiple radius parameters to detect meaningful locations at different granularities [20]. These methods overcome some of the place-size limitations and most of the false positives in Marmasse's approach, but the use of GPS signal loss to infer places still leaves these techniques unable to infer important outdoor places and multiple places within a single building. Other recent GPS-based methods have been proposed, but have similar problems with indoor places [22, 12].

Kang et al.

Kang et al. [8] proposed an approach based on distance and time-related heuristics similar to the idea proposed by Hariharan et al. [6] that does not depend on GPS signal losses. Their approach allows using continuous RF-emission based coordinate systems as location sources. Both find a new place when the distance of the new coordinates from the previous place is beyond a distance threshold and when the new locations span a significant time threshold. However, unlike Hariharan et al., who compute the distance between all pairs of coordinates after every new location measurement, Kang et al. incrementally compare the distance between the mean of the current cluster and the new measurement against the distance threshold. Unlike other clustering algorithms that require offline clustering of complete location traces, their time-based clustering algorithm incrementally extracts stays without expensive computation. However, this approach still does not resolve the inherent problems of GPS or RF-emission based coordinate systems which requires the intermediate step of acquiring geographical coordinates for every beacon scan and discovering places closer than the localization error of the systems.

Fingerprint-based Place Learning

Fingerprint algorithms detect stable radio environments that indicate a stay but provide no absolute location information for a place. These algorithms define the fingerprint of a place as a vector of currently visible cell towers or Wi-Fi access points, and use it to recognize when the device returns to a place. However, place learning is different from fingerprint-based localization as it attempts to discover a collected representation of locations. Algorithms can be categorized by the constraints on the fingerprints they use: either currently connected beacons or every neighboring beacon. Some of the algorithms are constrained by vendor APIs that only reveal the cell tower to which a phone is connected. Eliminating this restriction can favor better performance, but it is unclear when this restriction will be lifted on some platforms.

Laasonen et al. proposed a place learning algorithm based on currently connected cell towers [10]. Cliques are found by clustering cell towers. These cliques become places if their duration is longer than a threshold, but implementing the clustering on the phone is computationally expensive [19]. Froehich et al. identify places by triggering human intervention when a new cell-tower is connected [4]. Tangentially, related to our work, Krumm et al. measured the variance of the signal strength of the strongest Wi-Fi access point for smoothing transitions between the inferred states of "still" and "moving" [9]. Finally, Ahmad et al. proposed a fair election algorithm that finds the best representative beacons for various length of stays that a recognition method can use, but they do not address the problem of discovering the entrance to and departure from a place [1].

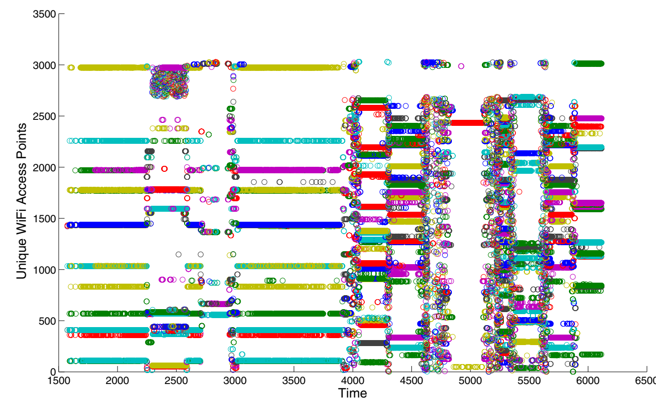
BeaconPrint

Most closely related to our work is that of Hightower et al. who discovered places by using multiple scan windows to distinguish beacons seen infrequently while a device is in the same place from new beacons seen as the device is physically departing [7]. entrance and departure are found when new beacons are continuously not found (or found) for more than c_{max} scan windows. A new beacon is one that was not seen in the previous w time interval. However, as it relies on new beacons and disregard familiar beacons, BeaconPrint often fails to find places where users are continuously mobile (e.g., marketplaces), or places where severely inconsistent beacons are found due to interference. Using multiple scan windows may also delay and hinder finding subsequent brief visits especially when the travel time between places is short. Finally, the BeaconPrint recognition phase simply compares the fingerprint seen by the device to a list of place fingerprints learned by the system. A histogram uses each unique beacon as a bucket and counts the time it was detected during the visit. A place fingerprint which shares a set of beacons with the highest weight in the tested fingerprint's histogram is selected as a match.

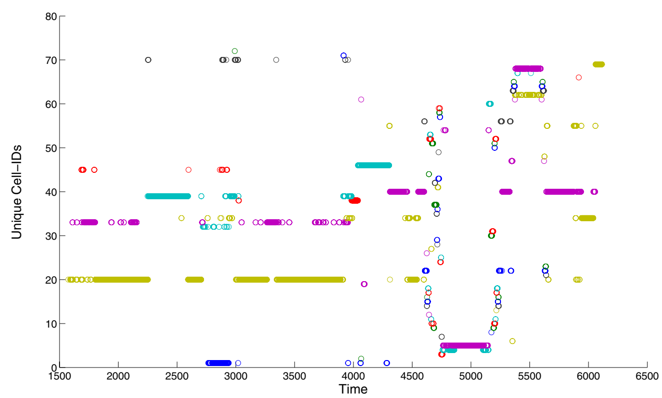
In our evaluation section, we evaluate PlaceSense against the most prominent algorithm from each of the two categories of place learning algorithms: Kang et al.'s geometric algorithm, and BeaconPrint.

THE PLACESENSE ALGORITHM

The PlaceSense algorithm is designed to discover places by continuously monitoring the radio beacons in the environment around a mobile device. It uses radio beacons from Wi-Fi access points (APs) or cell towers as its signals, which are pervasive and can be detected by most mobile devices. These beacons contain unique identifiers (the AP's MAC address or the tower's Cell-ID) for both communication set-



(a) Every neighboring 802.11 (Wi-Fi) Access Point



(b) Currently connected GSM cell towers

Figure 1. RF-Beacons found by a single person following normal routines. Each dot in the scatter graph is a beacon found from a RF-scan. The unique identifier of each beacon is on the vertical axis and the scan number is on the horizontal axis. Sampling rate is 0.1Hz. Both traces are from a single day.

up and hand-off. Mobile devices can periodically scan for these nearby beacons without connecting to or communicating with them. The IDs of these beacons are visible even if the network is private. A timestamped log of these beacon scans is the input to PlaceSense's discovery phase.

Discovering a place

Discovering a place involves two steps: detecting a stable radio environment that indicates an entrance to a place, and recognizing when it changes signaling a departure. A stable radio environment can be detected when consecutive radio scans contain familiar beacons or do not include new beacons for a specified amount of time. A beacon is considered new if none of the previous scans contained it; otherwise it is considered familiar. Thus, a set of beacons that were found in previous scans determines whether a beacon is considered new or familiar.

The biggest challenge facing such an algorithm is dealing with intermittent beacons. Such beacons may be seen briefly, for example, when crossing a hallway, when interference results in beacon losses, or when transiently traversing the edge of the transmission range of an AP. Simple techniques, such as discarding low response rate beacons, cannot, for example, correctly deal with lossy beacons.

In PlaceSense, we overcome this challenge with robust beacon inference. Like other approaches, a scan window, as opposed to a single scan, is used to tolerate noise and beacon losses. A window size, w, of a scan window defines the smallest time unit in which the algorithm will determine an entrance or departure of a place. For example, if one minute is chosen as the window size, a group of non-overlapping successive 1-minute scans are used by our algorithms (not sliding windows). Our algorithms use this group of scans to infer *fingerprints*, defined by a list of beacons and their response rate.

Entrance to a place.

Continuously seen stable scan windows imply a potential entrance to a place. A scan window is stable if it does not contain any new beacons. Stable depth, s_{max}, specifies at least how many stable scan windows must be seen to indicate an entrance to a place. Stable depth combined with window size ($s_{max} \times w$) specify how long the person must stay somewhere for it to be considered a place. To start an examination of a potential stay at a place, the current scan window is saved and compared against the following scan window. A stable value counter s is incremented when a scan window contains beacons that are a subset of the beacons found in the preceding scan windows. However, any scan window containing a new beacon terminates the examination and initiates another examination by setting s to 0 and clearing the history. A place entry is declared when s reaches s_{max}.

PlaceSense more conservatively detects entries than BeaconPrint which decrements s when new beacons are found, and aborts when s drops below 0. This approach is useful for avoiding infrequent beacons unnecessarily dividing a single place into multiple places. However, when discovering an entrance to a place, it can include unnecessary beacons in the fingerprint that is used to determine new beacons, so may be less robust when identifying a previously visited place. In contrast, PlaceSense's early termination quickly discards unnecessary beacons and only includes beacons that matter.

Departure from a place.

Detecting new beacons or missing familiar beacons indicate that the radio environment is changing and implies the mobile device is leaving the place. To suppress the influence of infrequent beacons, only selected beacons are considered when detecting changes. After entering a place, PlaceSense selects a set of representative beacons based on their response rate. A representative threshold r_{rep} is defined and every beacon with a response rate higher than the threshold is included in the representative set. If none of the beacons meets this condition, a beacon that has the highest response rate is used as a single representative beacon. Departure from a place is detected when all of the representative beacons disappear and new beacons are found in a scan window. Beacons are considered new if they are not seen during the stay. By relying on representative beacons, PlaceSense is more robust to infrequent beacons and discovering places where a device is highly mobile.

The response rate is a function of multipath, signal fading, MAC layer characteristics, and interference. LaMarca et al. [11] showed that when a device is stationary, the percent of scans which sees a particular beacon can be more effective in predicting the distance to that beacon than the signal strength values reported by the wireless network interfaces of both Wi-Fi cards and GSM phones. In PlaceSense, we measure the response rate as the ratio of detection count and total number of scans for each beacon:

$$\mathcal{R}_{k,x} = \frac{\sum_{i=1}^{n_k} b_{x,i}}{n_k},\ b_{x,i} = \begin{cases} 1 & \text{if beacon } x \text{ found in } i\text{th scan} \\ 0 & \text{otherwise} \end{cases}$$

where $\mathcal{R}_{k,x}$ is the response rate of beacon x at place k and n_k is total scan count since the place was entered. To update the response rate of every detected beacon during a stay, the detection count of each beacon is maintained by the fingerprint.

To avoid a single scan window determining a departure, a tolerance value t is used which can range from 0 to t_{max}. Tolerance depth, t_{max}, specifies at least how long scans must be unstable to indicate a leave from a place. Choosing t_{max} too high causes distinct places to be merged and long delays on decisions, while choosing it too low results in erroneous fragmentation. Tolerance value t decrements with every additional scan window that satisfies the above condition while incrementing when more than one representative beacon is found. If t reaches 0, the fingerprint is recorded and a stay is terminated.

Finally, PlaceSense uses a novel buffering strategy to rapidly detect place entry after quick transitions. Without this strategy, if the departure decision takes at least T minutes (defined by t_{max}) but a new place is entered within that time, entry determination will be delayed until the former place is terminated. PlaceSense buffers overlapping data and starts entry determination in parallel, as soon as the t value is below t_{max}. As soon as the previous place terminates, the buffered statistics are used to examine the next place.

Our two main innovations, a robust departure determination and a responsive short place transit determination, are responsible for PlaceSense's superior performance, as we show in the next section.

Adjusting to beacons in use

We can adjust PlaceSense to work with different types of RF-beacons. RF-beacons can be categorized by the constraints on the beacon fingerprints used: currently connected beacons and every neighboring beacon. When a fingerprint of every beacon is available, such as Wi-Fi scans depicted in Fig. 1(a), PlaceSense, as described above, can be used directly. However, when only currently connected beacons are available (e.g. cell tower-ID), as shown in Fig. 1(b), an adjustment can be made to the definition of representative beacons to improve performance. A threshold used to select representative beacons is no longer effective but the decision on which beacons are included in the set can make a difference. We tested two options. The first option, which we call PlaceSenseGSM, regards cell-IDs found during the speculation phase as representative beacons. However, this can be inaccurate at times and divide a place incorrectly if a handoff occurs during a long stay. The other version we tested is designed to reduce unnecessary Wi-Fi scans for discovering places. By additionally including new beacons to the representative set when representative beacons are found again before the value of t decreases below 0, PlaceSenseGSMCoarse detects larger-scale places and may trigger WiFi scans for finer resolution.

EXPERIMENTS

Evaluating the performance of a place discovery technique is not an easy task. Unlike the localization problem where the evaluation metric is often the distance between a real coordinate and an estimated coordinate, a place is typically not a single point nor has a universal spatial shape or size.

We use a novel evaluation methodology, and base our definition of a place on how semantically meaningful places are referred to by people. Thus, rather than attempt to find a geometric definition of a place, we used human participants to log any place they visited and stayed for more than five minutes. In our experiment participants were not provided with a specific definition of a place. Despite this, we show that there exists a definition of a stable radio environment (different from that in [7]) that well approximates human-labeled meaningful places in many cases. We now describe this methodology.

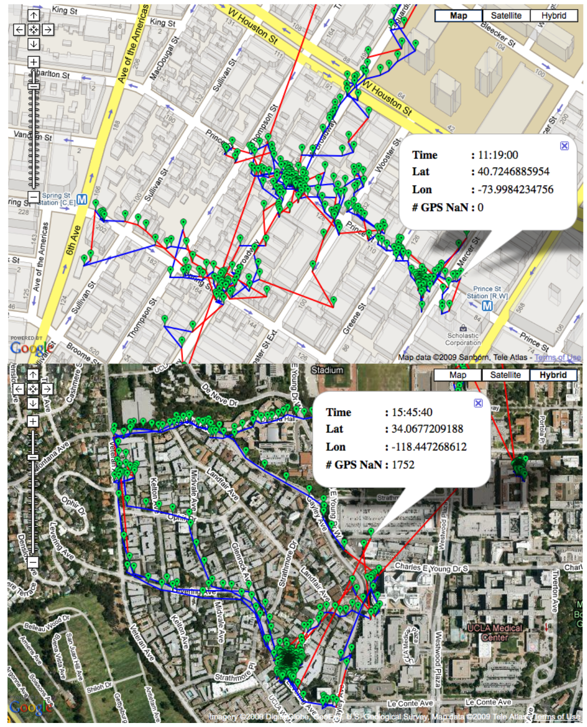


Figure 2. A webpage illustrating the GPS coordinates on a map allowing data collectors to correct any erroneous log entries. Each info balloon represents a GPS coordinate. Users can click on the balloons to get additional info: time, latitude, longitude, and the number of NaN (Not A Number) value returned due to GPS failures. Red lines indicate GPS failures in between data points.

Data Collection

We collected location trace logs using a Nokia N95 mobile phone, equipped with integrated GPS and built-in Wi-Fi. The phones were loaded with software configured to collect GPS, Wi-Fi, and GSM traces every 10 seconds (sampling at 0.1 Hz provided sufficient data resolution without a detrimental impact on battery life). All the nearby Wi-Fi beacons were logged, while only currently connected GSM cell towers were recorded. Traces were uploaded to a server for further analysis after data collection to reduce power requirements. Phones lasted for 4 to 5 hours and required a recharge twice a day during a day-long data collection. External battery packs were provided for long travels.

For initial evaluation, we conducted a scripted tour of 30 different places in 12 buildings and 4 outdoor plazas. Each data collector individually selected 10 places they go to often on the UCLA campus which included various building rooms, library floors, stores, gyms, patios, and food courts. Several places were within a single building and some places overlapped between participants. Three distinct visit durations (8, 10, and 15 minutes) were distributed to data collectors (10 visits per each). Distance between places varied from 1 to 10 minutes by a normal walk. Data collectors were asked to stay at a place for a predefined amount of time and entrance and departure times were recorded whenever they entered or left the room. For outdoor places (e.g. outdoor tables), time was recorded when the data collector started or ended being stationary. For further validation, we collected 4 week-long location trace logs from each of three data collectors as they went about their normal lives. They mostly stayed within the local city limits, while a couple of traces were also collected in three different cities during a trip. All results are presented together as no significant difference was observed.

To collect ground-truth, each data collector was asked to keep a diary of the name and time they entered and left every place they stayed more than 5 minutes during the experiment. Home, office, lunch places near work, coffee shops, bus stops, and class rooms were frequently visited, while various restaurants, stores, markets, clinics, etc. were less frequently visited. At the conclusion of each data collection, a webpage illustrating the GPS coordinates on a map with timestamps was provided to allow the data collectors to correct any erroneous log entries (Figure 2). These diaries and maps provided the ground-truth information about the coordinates of the actual places the data collector went as well as the actual times they arrived and left those places. However, as GPS data was not available in most of the indoor as well as many outdoor locations, there were limitations on achieving accurate time information. While participants initially kept an accurate diary, the time accuracy deteriorated within the first few days. Many time entries were inaccurate and short visits were often not recorded. Each data collector found *interesting* places when reviewing the results.

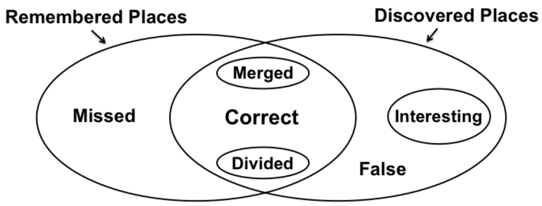


Figure 3. ***Remembered places*** **(recorded by people) and** ***discovered places*** **(found by place discovery techniques). More** ***correct*** **and** ***interesting*** **places and fewer other erroneous places indicate better performance.**

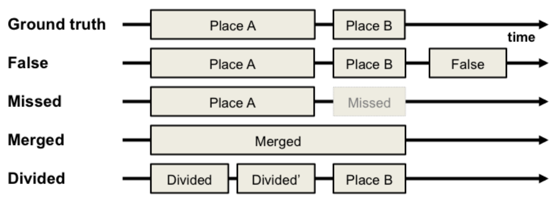


Figure 4. Four types of erroneous place discovery. ***False*****: erroneous third place is found,** ***missed*****: place B is missing,** ***merged*****: two distinct place A and B are merged as a single place,** ***divided*****: place A is divided into two distinct places.**

Evaluation Metrics

To quantify the effectiveness of a place discovery technique, we define a set of meaningful and erroneous places following Zhou's work with two additional types: *merged* and *divided* [21]. As shown in Figure 3, places recorded in a user's diary are called *remembered places* and places discovered by a place discovery technique are called *discovered places*. *Remembered places* that are not discovered are called *missed*, while places that are both remembered and discovered are further categorized as *correct*, *merged*, and *divided*. As illustrated in Figure 4, if two different places reported by the user are discovered as a single place, the places are called *merged*. Likewise, if a single place reported by the user is discovered as two or more places, it is labelled as *divided*. Others that were both remembered and discovered are classified as *correct*. Of the remaining discovered places, places are categorized as *interesting* if the user claims it as a legitimate stay while reviewing the results, otherwise it is classified as *false*. More *correct* and *interesting* places indicate better performance, while the distribution of erroneous places allows us to understand the strength and weakness of each technique. We further define *precision* and *recall* as follows:

$$Precision = \frac{\#\,Correct + \#\,Interesting}{\#\,Discovered},\quad Recall = \frac{\#\,Correct}{\#\,Remembered}$$

The accuracy of a discovered place is further evaluated by time offsets of the entrance or departure time of a place. Offset is measured as the difference between the time determined by a place discovery technique and the time manually recorded by data collectors. Accurate time information of a visit to a place may not be critical to users, but accurate segmentation of radio signals indeed improves the performance of place learning and recognition techniques. However, remembering to log the exact time whenever visiting a place is challenging for the data collectors during their daily lives. Thus, we first conducted a scripted tour of 30 different places on campus varying the stay from 8 to 15 minutes accompanied by accurate time records in order to evaluate the timeliness of each technique.

Implementation

In addition to PlaceSense, two previously proposed algorithms, BeaconPrint [7] and Kang et al. [8], were implemented for comparison. Improvements to the basic PlaceSense algorithm were implemented incrementally to evaluate the effect of each change. For our implementation of BeaconPrint, every beacon discovered during both the stay and the time window w (used for checking stable scans at the beginning) were included in the fingerprint. We used the parameters for window size $w = 2$ minutes (120 seconds) and confidence depth $c_{max} = 3$, as in BeaconPrint. Kang et al. designed a time-based clustering algorithm to overcome problems found by previous algorithms that depended on GPS failures in indoor locations. This approach takes as input a stream of timestamped geographic coordinates derived from any location system. For our experiments, we use their suggested parameters for time $t = 300$ seconds and distance $d = 300$ meters. GPS coordinates were used as input. GPS failures were regarded as a stay within the distance threshold. For every version of PlaceSense, a one minute window size was used which contained six samples (i.e. sample rate = 10 seconds). Stable depth s_{max} and tolerance depth t_{max} were both set as 3, similar to the certainty value suggested by BeaconPrint. Time series of Wi-Fi or GSM traces were used as input.

Results

To investigate the representative threshold r_{rep}, we examine the performance of PlaceSense with different thresholds. Then, we illustrate the enhancements achieved by our incremental improvements: tolerance depth and buffering beacon scans for identifying the next potential place. Data sets from a scripted tour are used as they provide more accurate ground-truth over the data collected from collectors following their normal lives. The performance of our final version is compared against other algorithms including Kang et al. [8], and BeaconPrint [7], and our own PlaceSenseGSM. Finally, we evaluate how well our approach works on real-life data by comparing PlaceSense against BeaconPrint and Kang et al. using data collected from normal lives.

r_{rep}	Correct	Merged	Divided	Missed	False	Recall
0.5	20	10	0	0	0	0.67
0.6	20	10	0	0	0	0.67
0.7	22	8	0	0	0	0.73
0.8	22	6	2	0	0	0.73
0.9	24	4	2	0	0	0.80
1.0	12	0	3	15	0	0.40

Table 1. Number of erroneous places discovered by basic PlaceSense. Larger representative threshold value r_{rep} results in fewer merged places, but more divided places.

r_{rep}	Correct	Merged	Divided	Missed	False	Recall
0.5	18	12	0	0	0	0.60
0.6	18	12	0	0	0	0.60
0.7	18	12	0	0	0	0.60
0.8	24	6	0	0	0	0.80
0.9	26	4	0	0	0	0.87
1.0	22	0	0	8	0	0.73

Table 2. Erroneous places discovered by PlaceSense using tolerance depth (t_{max}). Introducing t_{max} parameter reduces erroneously divided places while allowing larger r_{rep} threshold value to reduce the number of merged places.

As shown in Table 1, smaller r_{rep} results in less divided places and more merged places as it becomes conservative in determining a departure. Larger r_{rep} decreases the population of representative beacons and further increases the chances of missing all of the representative ones. All of the five merged places (when threshold is $0.5 - 0.7$) are from the case when the data collector traveled from one floor to another in a single building. Smaller threshold values allow more beacons with lower response rates (which are possible on different floors) as representative beacons. Larger values resolve some of these cases but also introduce erroneously divided places by sensitively determining departures. An extreme threshold value of 1.0, which requires discovering the beacon in each scan window during the stay, severely degrades performance by setting the bar for becoming a representative beacon too high. Thus, larger threshold values (except 1.0) are preferable for mitigating merging effects but require a different approach to reduce erroneous departures.

The basic version determines a departure when no representative beacons are discovered in a single window. Any single window missing every representative beacon aggressively determines that the device is leaving a place. Instead, the tolerance parameter requires at least t_{max} windows to not detect the representative beacons, reducing erroneously divided places while allowing larger threshold values to reduce merged places. Tolerance depth t_{max}, the maximum tolerance value, is set as three windows which we used to determine a stable scan when entering a place. By avoiding a single window, the number of erroneously divided places is reduced. Table 2 shows that the tolerance parameter eliminates the two divided places when using threshold 0.8 and 0.9. Smaller threshold values, on the other hand, resulted in more merged places as the tolerance value reduced the opportunities to terminate a stay.

Table 3 compares the performance of different place discovery techniques on traces from a scripted tour. BeaconPrint missed places where Wi-Fi signals are significantly inconsistent. Short trips, such as moving from one floor to another floor in the same building, often did not provide enough new beacons to terminate a place and start a new one. In contrast, PlaceSense is robust against noisy beacons as long as more than one beacon is consistently discovered during a stay. PlaceSense was also more sensitive when changing venues unless more than one beacon was found strongly in both places. PlaceSenseGSM, which only uses the currently connect cell tower, performed reasonably well, even better than Kang et al. [8], but showed limitation in discriminating between closely located places. Some places had many handovers during a short period of time, resulting in being identified as a missed place. Most of the places missed by Kang et al. were due to GPS failing to lock enough satellites in time. All of the merged places resulted from failing to discriminate between different floors. While using Wi-Fi-based coordinate systems may reduce the number of missed places, merged places are unlikely to be reduced dramatically as the coordinates are too coarse to differentiate the floors.

Algorithm	Correct	Merged	Divided	Missed	False	Recall
PlaceSense	26	4	0	0	0	0.87
BeaconPrint	18	10	0	2	0	0.60
Kang et al.	9	14	0	7	0	0.30
PlaceSenseGSM	11	16	0	8	0	0.40

Table 3. Number of erroneous places discovered by Kang et al., BeaconPrint, PlaceSense (Wi-Fi), and PlaceSenseGSM from a 30-places-scripted-tour data set.

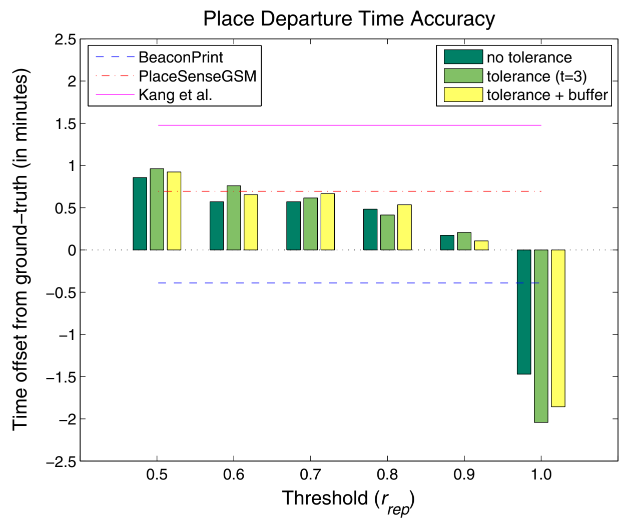


Figure 5. The time offset of departure time in Kang et al., BeaconPrint, PlaceSenseGSM, and PlaceSense. For PlaceSense, larger r_{rep} value decrease the time it takes before terminating a stay.

We further evaluated the accuracy of each discovered place by its entrance and departure times. To measure the boundary accuracy of places found by place discovery algorithms, we measured the time offset of entrance and departure times of each place. We excluded missed places in this evaluation and used only the beginning and end that matched with the ground-truth in the cases where the place was divided or merged. We discuss the departure time first as our results exhibit effects of delayed leaving time on the subsequently visited place's entrance time. Many of the places we visited during the scripted tour were within three minutes walking distance and could effect each other's time boundaries. As shown in Figure 5, larger threshold values decreased the time it takes before terminating a stay. Fewer beacons become representative when larger threshold values were used and increase the opportunity to lose all of them earlier. Setting the threshold extremely large, such as 1.0, dramatically in-

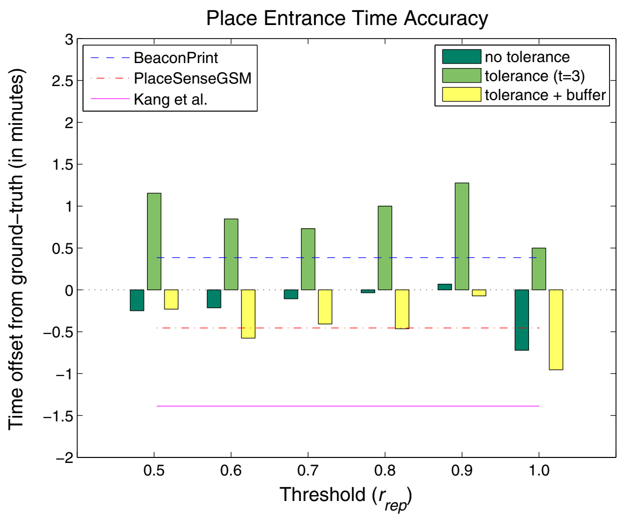


Figure 6. The time offset of entrance time in Kang et al., BeaconPrint, PlaceSenseGSM, and PlaceSense. For PlaceSense, large hidden delays in departure time when r_{rep} is small prevails the decreasing effect of lowering r_{rep}. Buffering reduces the hidden delay effect.

creased the incidence of the early termination of a stay. We did not observe a noticeable delay in departure time when the tolerance value is used, since departure time was recorded when the tolerance value starts dropping below t_{max}. Consequently, the departure time illustrated in Figure 5 did not illustrate the delay in making the decision. However, this hidden delay may impair the accuracy of the entrance time to the next place visited, when the travel time between the two places are less than $t_{max} \times w$.

Similar to departure time, entrance time, as shown in Figure 6, is delayed when larger values of r_{rep} are used. However, in the case of using tolerance values, the hidden delay postpones discovering the subsequent visit and delays the entrance time. Large hidden delays in departure time, when the threshold value is $0.5 - 0.6$ (Figure 5), dominates the decreasing effect of lower threshold (Figure 6). However as a larger threshold value reduces the delay on departure time, the threshold effect on the entrance time begins to dominate and increases the delay. Consequently, we have introduced a buffer that starts speculating about the next potential place while terminating the current one. By overlapping the departure and entrance period, effects of the hidden delay are significantly reduced. But buffering did not change the number of correctly discovered places or the distribution of erroneously discovered places.

Average offsets of entrance and departure time of BeaconPrint, Kang et al., and PlaceSenseGSM are depicted in Figure 6 and Figure 5. BeaconPrint, in general, is more sensitive to changes than the other place discovery algorithms. This leads to comparatively earlier departure time and delayed entrance time. On the other hand, GSM and GPS based approaches are less sensitive to changes as GPS and GSM provide lower location resolution. With a high enough threshold value (except 1.0) accompanied with our incremental improvements, PlaceSense outperforms other place

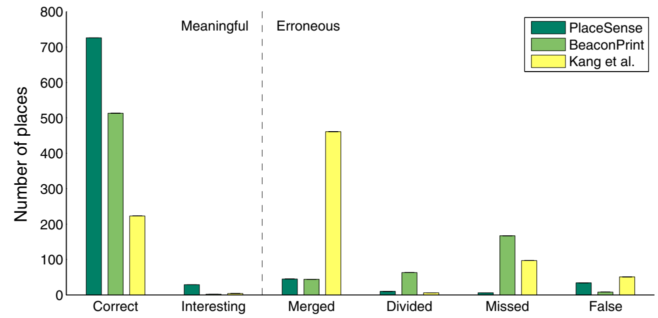


Figure 7. Number of places found from real-life traces by PlaceSense, BeaconPrint, and Kang et al. PlaceSense reduces the number of missed places while also increasing the number of interesting and false places.

	Aron			Bryan			Chris			All		
	PS	BP	KA	PS	BP	KA	PS	BP	KA	PS	BP	KA
Cor.	233	156	81	251	182	63	242	175	79	726	513	223
Int.	10	1	2	6	1	0	13	0	2	29	2	4
Mer.	6	14	138	23	15	185	16	15	138	45	44	461
Div.	2	21	1	2	30	3	6	12	2	10	63	6
Mis.	0	50	21	3	52	28	3	65	48	6	167	97
Fal.	6	2	10	14	2	20	14	4	21	34	8	51
Recall	0.97	0.65	0.34	0.90	0.65	0.23	0.91	0.66	0.30	0.92	0.65	0.28
Precision	0.95	0.81	0.36	0.87	0.80	0.23	0.88	0.85	0.33	0.89	0.82	0.30

Table 4. The distribution of discovered places by different users (PS: PlaceSense, BP: BeaconPrint, KA: Kang et al.)

discovery techniques both in the number of place errors as well as time boundary accuracy.

Following our initial evaluation, we further validated our algorithm by running it on multi-day traces. Three data collectors collected these traces for four weeks each, following their normal lives. Our data collectors were assigned the pseudonyms Aron, Bryan, and Chris. Traces contained various routines from ordinary work and home routines to a multi-day trip to other cities. The results of PlaceSense on these real-life traces generated with a representative threshold $r_{rep} = 0.9$ and tolerance depth $t_{max} = 3$ (optimal configuration obtained from our initial evaluation) are shown in Figure 7 and is compared against BeaconPrint and Kang et al. using their suggested parameters. However, we did not evaluate the time boundary accuracy as the error range of time records provided by the data collectors were often more than five minutes.

By focusing on representative beacons and buffering data for subsequent visits, PlaceSense reduces the number of missed and divided places while also increasing the number of interesting and false places compared to BeaconPrint (Fig 7). Kang et al. based on GPS resulted in significantly more merged places as many proximate places located in nearby buildings are identified as a single place. No significant difference among the three data collectors are found (Table 4). PlaceSense has the highest overall recall and precision of all three data collectors. The recall numbers are also consistent with our results from the scripted tour (Table 3). Aron drives to work and visited various buildings during work hours. 63 different places are visited, and 36 places including various

	5 -10 min			10 - 30 min			30 - 2 hrs			2 - ∞ hrs		
	PS	BP	KA	PS	BP	KA	PS	BP	KA	PS	BP	KA
Cor.	158	45	47	180	138	61	213	187	54	175	143	61
Int.	26	2	1	3	0	3	0	0	0	0	0	0
Mer.	12	15	97	13	12	102	17	12	158	3	5	104
Div.	0	0	0	3	7	1	5	27	1	2	29	4
Mis.	4	114	30	1	40	33	1	10	23	0	3	11
Fal.	33	6	38	1	2	12	0	0	1	0	0	0
Recall	0.91	0.26	0.27	0.91	0.70	0.31	0.90	0.79	0.23	0.97	0.79	0.34
Precision	0.80	0.69	0.26	0.92	0.87	0.36	0.91	0.83	0.25	0.97	0.81	0.36

Table 5. The distribution of discovered places by their visit durations (PS: PlaceSense, BP: BeaconPrint, KA: Kang et al.)

restaurants and shops are visited once. Regular places other than work and home include a gym, grocery stores, gas stations, and lunch places near work. Bryan takes buses to work and frequently visits several indoor rooms in a single building at work: offices, server rooms, and meeting rooms. An outdoor patio for lunch, coffee shops, bus stops, and a couple of outdoor recreation places were often visited. Among the 50 different places he visits, 26 places are only visited once during data collection. Chris, an undergraduate student, walks to school and visits 108 different places in four weeks. In the first two and a half weeks, many classrooms, friend's residences, and lab rooms are regularly visited. For the remaining days, Chris visits multiple cites during spring break, which led to more single visits than the other data collectors (81 places).

We further investigate the distribution of discovered places by their visit duration in Table 5. PlaceSense shows strength in discovering brief visits as well as other long-term places where the radio environment is unstable with many infrequent beacons. Places that BeaconPrint missed, but PlaceSense discovered, include short visits to a convenience store, gas station, post office, and various indoor rooms as well as a long stay in meeting rooms and seminar rooms with many unstable beacons. Visits to various stores where people generally roam around within a larger restricted area were also often missed. False places found by both algorithms include unrecognizable short stays that are not recorded in the data collector's diary. PlaceSense additionally finds a slow walk through the hallway or open area as a place when a strong beacon is found continuously during the walk. A ride on a Wi-Fi enabled cab is also discovered as a place. However, many of these cases are repeated and can be recognized so that they can be filtered out. Interesting places mostly are unrecorded brief visits to various place such as bus stop, gas station, copy room, parking lot, etc., that are recoverable during diary and map reviews.

Finally, to investigate the overall improvement of PlaceSense in recognizing revisited places by enhancing place discovery, we compare against BeaconPrint using the same recognition algorithm it uses. We focus on how well places that were actually visited are recognized. The first visits to a place discovered correctly by both algorithms are given to the algorithms for learning. Any subsequent visits are used to evaluate recognition accuracy. Our real-life data set included 143 places that have a valid learning visit, but only 63 places are visited more than once. Additionally, there are 62

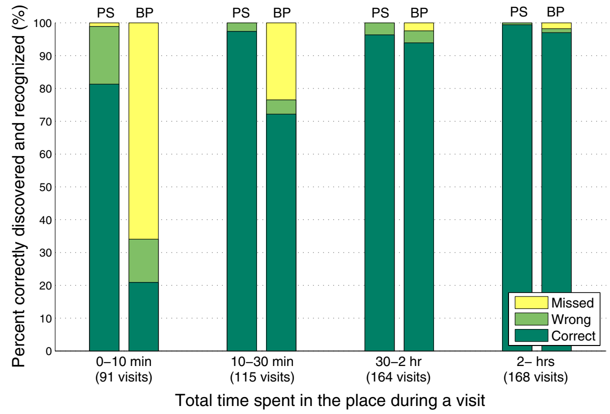


Figure 8. Percent correctly discovered and recognized from real-life traces by time spent at the place (PS: PlaceSense, BP: BeaconPrint). PlaceSense significantly improves accuracy in recognizing short visits.

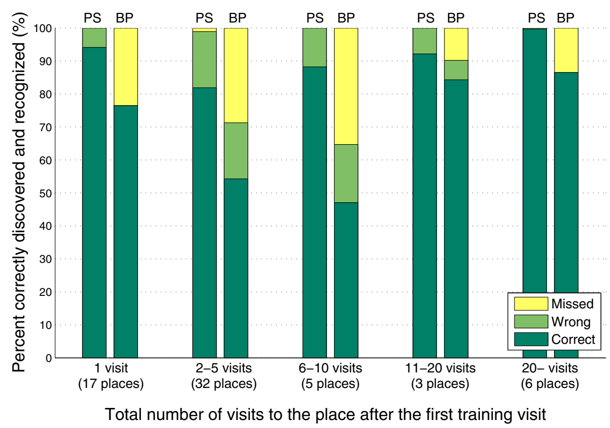


Figure 9. Percent correctly discovered and recognized from real-life traces by the number of visits to the place (PS: PlaceSense, BP: BeaconPrint). Frequently visited places are often briefly visited.

places (49 were visited once) that only PlaceSense has a correct learning visit, and 6 places (5 are visited once) that only BeaconPrint has. 10 places have no valid learning visit for both algorithms. Kang et al. is excluded as its performance in correctly discovering places is significantly lower than the others. Places with no learning data are excluded. The percentage an algorithm correctly identifies a place is the ratio of the total number of places the algorithms correctly predicts to the total number of places the data collectors actually visited. Each error by the algorithms is further broken down to *missed* and *wrong* places. Figure 8 shows PlaceSense's notable strength in recognizing short visits. For both algorithms, similar percentage of places were incorrectly recognized as nearby places that shares similar beacons. Figure 9 implicates that frequently visited places are often visited for less than 30 minutes.

CONCLUSION AND FUTURE WORK

Our results show that PlaceSense provides a significant improvement in the ability to discover and recognize places. Precision and recall with PlaceSense are 89% and 92% versus the previous state-of-the-art BeaconPrint approach at 82% and 65% precision and recall. Because it uses response rate to select representative beacons and suppresses the influence of infrequent beacons, PlaceSenses accuracy gains are particularly noticeable in challenging radio environments where beacons are inconsistent and coarse. PlaceSense also detects place entrance and departure times with over twice the precision of previous approaches thanks to judicious use of buffering and timing. It has the ability to overlap the departure fingerprint of one place with the arrival fingerprint of the subsequent place. Lastly, PlaceSense is accurate at discovering places visited for short durations (less than 30 minutes) or places where the device remains mobile. Accuracy in short-duration and transient places is a significant contribution because these types of places are valuable to emerging applications like life-logging and social location sharing.

We believe PlaceSense has solved some of the major open technical problems and corner cases with beacon-based place detection. Although there is always value in incremental research to refine and validate the place learning algorithms themselves, we think most future research should focus in two new areas. First, we promote a research agenda that moves up the application stack. Just as coordinate-based location applications have exploded recently with services such as Google Maps Mobile and Apple iPhone map applications, we believe that place will be a critical feature in the next wave of killer mobile applications. People are going to want to-do lists, instant messaging, photo journals, and other everyday applications to have awareness that goes beyond simple numerical latitude and longitude coordinates. PlaceSense can run on todays commodity mobile devices because it uses existing Wi-Fi and GSM radios as its sensors and has only modest computational requirements, which means researchers now have both the devices and place detection capabilities to deploy, instrument, and evaluate everyday applications that use place data. The outcome of these field studies will hopefully be a much more nuanced insight into how people actually perceive places in specific application contexts. The results could feed back into the place learning algorithms themselves to extend them with the ability to be tuned for particular uses or situations. The second future direction we see is research to combine place learning with other sorts of contextual information. We believe the two most important contexts to complement Where am I? are Who am I with? and What am I doing? Combining location with activity and social interaction offers the exciting goal of rich, natural, personalized interactions with mobile devices and instrumented environments. Integrating these contexts involves many challenges in joint modeling, machine learning, sensing, and interaction but could pay off with a whole new level of capability for high-value applications like health monitoring, gaming, and social communication.

ACKNOWLEDGEMENTS

This work is supported by NSF Grant CNS-0453809 and UC MICRO Grant 07-048, co-sponsored by Nokia and Cisco. Hardware is provided by Nokia Research Labs under the Sensor Planet initiative.

REFERENCES

1. U. Ahmad, B. J. d'Auriol, Y.-K. Lee, and S. Lee. The election algorithm for semantically meaningful location-awareness. In *MUM '07*, pages 55–63, 2007.
2. D. Ashbrook and T. Starner. Using gps to learn significant locations and predict movement across multiple users. *Personal Ubiquitous Comput.*, 7(5):275–286, 2003.
3. A. Campbell, S. Eisenman, N. Lane, E. Miluzzo, R. Peterson, H. Lu, X. Zheng, M. Musolesi, K. Fodor, and G.-S. Ahn. The rise of people-centric sensing. *Internet Computing, IEEE*, 12(4):12–21, July-Aug. 2008.
4. J. Froehlich, M. Y. Chen, I. E. Smith, and F. Potter. Voting with your feet: An investigative study of the relationship between place visit behavior and preference. In *Ubicomp '06*, pages 333–350, 2006.
5. S. Gaonkar, J. Li, R. R. Choudhury, L. Cox, and A. Schmidt. Micro-blog: sharing and querying content through mobile phones and social participation. In *MobiSys '08*, pages 174–186. ACM, 2008.
6. R. Hariharan and K. Toyama. Project lachesis: Parsing and modeling location histories. In M. J. Egenhofer, C. Freksa, and H. J. Miller, editors, *GIScience*, volume 3234, pages 106–124. Springer, 2004.
7. J. Hightower, S. Consolvo, A. LaMarca, I. E. Smith, and J. Hughes. Learning and recognizing the places we go. In *Ubicomp '05*, pages 159–176, 2005.
8. J. H. Kang, W. Welbourne, B. Stewart, and G. Borriello. Extracting places from traces of locations. In *WMASH '04*, pages 110–118, New York, NY, USA, 2004. ACM.
9. J. Krumm and E. Horvitz. Locadio: inferring motion and location from wi-fi signal strengths. *MOBIQUITOUS '04*, pages 4–13, 2004.
10. K. Laasonen, M. Raento, and H. Toivonen. Adaptive on-device location recognition. In *Pervasive '04*, pages 287–304, 2004.
11. A. LaMarca, Y. Chawathe, S. Consolvo, J. Hightower, I. E. Smith, J. Scott, T. Sohn, J. Howard, J. Hughes, F. Potter, J. Tabert, P. Powledge, G. Borriello, and B. N. Schilit. Place lab: Device positioning using radio beacons in the wild. In *Pervasive '05*, pages 116–133, 2005.
12. L. Liao, D. Fox, and H. Kautz. Extracting places and activities from gps traces using hierarchical conditional random fields. *Int. J. Rob. Res.*, 26(1):119–134, 2007.
13. P. J. Ludford, D. Frankowski, K. Reily, K. Wilms, and L. Terveen. Because i carry my cell phone anyway: functional location-based reminder applications. In *CHI '06*, pages 889–898. ACM, 2006.
14. N. Marmasse and C. Schmandt. Location-aware information delivery with commotion. In *HUC '00*, pages 157–171. Springer-Verlag, 2000.
15. E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell. Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application. In *SenSys '08*, pages 337–350. ACM, 2008.
16. M. Mun, S. Reddy, K. Shilton, N. Yau, P. Boda, J. Burke, D. Estrin, M. Hansen, E. Howard, and R. West. Peir, the personal environmental impact report, as a platform for participatory sensing systems research. In *MobiSys '09*. ACM, 2009.
17. Navizon. http://www.navizon.com.
18. SkyhookWireless. http://www.skyhookwireless.com/.
19. N. Song. Discovering user context with mobile devices: location and time. Thesis, Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science., 2006.
20. N. Toyama, T. Ota, F. Kato, Y. Toyota, T. Hattori, and T. Hagino. Exploiting multiple radii to learn significant locations. In *LoCA '05*, pages 157–168, 2005.
21. C. Zhou, D. Frankowski, P. Ludford, S. Shekhar, and L. Terveen. Discovering personally meaningful places: An interactive clustering approach. *ACM Trans. Inf. Syst.*, 25(3):12, 2007.
22. C. Zhou, P. J. Ludford, D. Frankowski, and L. G. Terveen. How do people's concepts of place relate to physical locations? In *INTERACT '05*, pages 886–898, 2005.