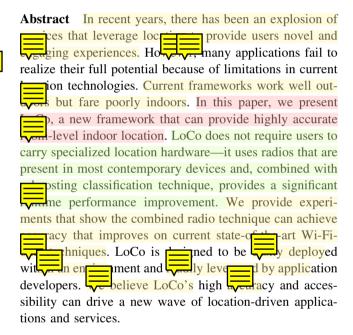
### ORIGINAL ARTICLE



# LoCo: boosting for indoor location classification combining Wi-Fi and BLE

Matthew Cooper<sup>1</sup> · Jacob Biehl<sup>1</sup> · Gerry Filby<sup>1</sup> · Sven Kratz<sup>1</sup>

Received: 9 January 2015/Accepted: 17 November 2015/Published online: 12 January 2016 © Springer-Verlag London 2016



Keywords Indoor location detection · Multi-radio indoor ioning · Location-aware application frameworks

Matthew Cooper cooper@fxpal.com

Jacob Biehl biehl@fxpal.com

Gerry Filby filby@fxpal.com

Sven Kratz kratz@fxpal.com

### 1 Introduction

Leveraging physical location to drive meaningful, context-specific user experiences has been specific user experiences and other satellite-based technologies provide pervasive outdoor location, reliable indoor location, reliable indoor location, remains out of reach despite years of research. A part adoption inhibitor to many proposed solutions is that they require users to wear or carry specialized devices for tracking, e.g., badges that recognize infrared beacons [37] or specialized radio devices.

Recently, there has been interest in leveraging existing structure and devices to determine indoor location. If it is included with smartphones such as Apple's iPhone and Google's Android platform. The typical approach to Wi-Fi localization is to use the client's received signal strength indication is to use the client's received signal strength indication of fixed Wi-Fi access points (APs) as a use of distance between the APs and the device. These inces, along with the locations of the APs, are used in a lateration calculation to predict the device's location. It to the inherent error of these calculations, most ing results have several meters of error.

alternative to multilateration is location classificationing RSSI scans collected by the mobile device. The approach trades precision for accuracy; the approach does not track a device's location as (x, y) coordinates, but provides a high confidence prediction of what ion the device is in using a predefined discrete set of locations.

this work, we propose new techniques that seek to practical to deploy and be leveraged by application opers. Our new framework, called LoCo, achieves ased room-level accuracy by combining Wi-Fi RSSI



FX Palo Alto Laboratory, Palo Alto, CA, USA

Is with signals from a new radio type, Bluetooth low y (BLE). While our approach utilizes a relatively new technology, it is important to note that many existing phones already contain BLE radios (e.g., Apple's iPhone 4S and newer, Samsung Galaxy SIII and S4) and off-theshelf BLE beacons are becoming increasingly available with steadily decreasing costs (e.g., *iBeacons*<sup>1</sup>, *ToDs*<sup>2</sup>, and *Estimotes*<sup>3</sup>).

The framework we present is designed to achieve roomlevel location accuracy that is meaningfully higher than jously developed techniques in order to enable new or location-based applications. State-of-the-art Wi-Fi rprinting methods (e.g., [6]) are documented to provide racy near 85 % for room-level localization. For applications such as equipment tracking or inventory management, this degree of accuracy is acceptable. However, for location applications and services that we envision, such as robustly detecting the presence of an individual in a room, this level of performance falls short. To improve accuracy, we extend Wi-Fi RSSI-based techniques with the use of beacons that are deployed explicitly for indoor locadentification. We show that integrating BLE enhances the accuracy of both our earlier work based on Wi-Fi [5] as well as a high-performing Wi-Fi RSSI-based baseline [6].

We present a series of experiments to validate this approach. First, we show that using BLE beacons alone allows for high-performance room-level location identification. Additionally, we show that combining signal strength measurements from the BLE beacons with those from Wi-Fi access points (APs) consistently improves performance over using either modality alone. In all configurations, the proposed approach dramatically reduces computational complexity, thus enabling location identifin to be performed more frequently and on resource trained devices (e.g., mobile smartphones). We perform additional analysis of the utility of the multiradio pach for location identification at sub-room level in an n plan" office space. Finally, we describe experiments that demonstrate that while multiradio location classificas superior, it also is resilient across a range of densities i-Fi access points.

### 2 Related work

We have structured the review of related work to cover previous contributions on indoor localization algorithms using only Wi-Fi, fusing Wi-Fi with other wireless technologies, and alternate sensing technologies.

<sup>&</sup>lt;sup>3</sup> http://www.estimote.com/.

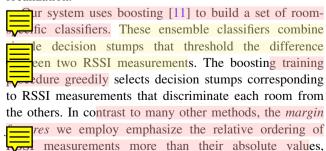


### 2.1 Wi-Fi-only localization

There is a great deal of prior work on using Wi-Fi RSSI measurements to locate devices and their users. Many methods aim for precise physical localization (i.e., multilateration) rather than quantizing position to a room or region. While room information may not be required for ization, the locations of wireless access points (APs) ally are needed.

u et al. [19] provide a comprehensive survey of Wi-Fi ion determination methods. They contrast triangulamethods that use geometric analysis to estimate location with scene analysis methods that match test measurements against templates of locations' radio frecy characteristics. More recent triangulation work has sed on approaches that model radio propagation statistically in combination with previously defined templates for the signal streeths observed at various locations [21, 40]. In [40], lowers are also clustered according to whether base station in range to accelerate location determination by temental triangulation". Other approaches include localization via genetic programming [7], Bayesian networks [20], and combined mapping of ions with tracking of devices as a semi-su ng problem [27]. There has been work that ration through of flight and/or angle of arrival (e.g., [8, 24, 39]). sophisticated, these approaches require modification and/or special software to run on the Wi-Fi base stations. More importantly, while gains have been made in addressing multipath issues in dense environments like office buildings (e.g., [32]), current techniques have not been shown to be robus real-world aptions.

whing methods represent the set of RSSI values at a given location by a known (absolute) set of values and minimize a distance measure tign a test data point to a proposition in location [6, 9, 18, 31]. It is prevised classification for machines and naive Bayes classifiers compared poorly with the Redpin and k-nearest-neighbors (kNN) approach in [18]. The work in [26] describes a method that combines generative modeling using EM to help address missing data and discriminative modeling with SVMs for localization.









<sup>1</sup> http://support.apple.com/en-us/HT202880.

<sup>&</sup>lt;sup>2</sup> http://www.todhq.com/.

both substantial complexity reductions relative to hing (i.e., kNN) methods, and superior accuracy.

We previously reported on the use of boosting for Willy classification in [5]. In this work, we introduce the beacons and present a much expanded evaluation with additional experimental results that explore the value and implications of a combined radio approach to room-level indoor location detection.

### 2.2 Wi-Fi and other wireless technologies

iguez et al. [30] study Wi-Fi, Bluetooth, and Zigbee ndoor localization. They use kNN on the raw RSSI mation for matching locations and observe an increase in location accuracy when multiple wireless sources are used. They mention problems with interference between wireless devices as more beacons are added. In the present, this problem is mitigated by using low transmission y on the Bluetooth beacons.

Incorporating additional location information from Bluetooth is proposed by [4]. This work uses Bluetooth pns as an ancillary technology to narrow down the space for Wi-Fi fingerprinting. They derive their results from simulations and modeling but did not conduct riments with real hardware and environments to verify results.

### 2.3 Alternate localization technologies

pative technologies are ultrasound, (active) radio frecy identification (RFID), and infrared (IR). The Active technologies are ultrasound sensors that use time-of-flight multilateration for localization. While ActiveBat achieves a high degree of accuracy, it relies on relatively heavy instrumentation (100 sensors for an area of 280 m²) of the localization space as well as the addition of ware to all tracked devices. The Cricket system [28] ws up RF transmissions with ultrasound bursts. The difference in speed between RF and sound can be used to calculate to a beacon.

The ton [16] system uses multiple RFID readers for location triangulation of active RFID tags. They reported relatively poor accuracy, providing location estimates in voxels with 3 m side lengths. Ni et al. [23] tags to better calibrate the base stations. Using kNN techniques, their system achieves a relatively precise (1–2 m) localization accuracy, but at the cost of requiring a

significant amount of fixed reference tags (one per m<sup>2</sup> co—1).

eBadge [37] was an early location system based on badges augmented with IR transmitters and IR detectors in the environment. The main limitation of using IR is that as a line-of-sight technology, badge occlusion poses problems. Also, the range is relatively limited (the authors state a range of about 6 m).

### 3 LoCo location framework

The Lyperamework consists of typeramemork is a pyment of Wi-Fi access points and Bluetooth low typeramemory (BLE acons, a client service running on a classification engine. We discuss each of these in detail below.

### 3.1 Constellation of Wi-Fi APs and BLE beacons

Lowerts signal strength information from two differences of wireless transmitters. As in past work, the Tramework legges the received signal strength indicator (RSSI) of the D beacons coming from in-range 802.11 Wi-Fi access points (APs). These values are easily accessible by most Wi-Fi frameworks, including most smartphone operating systems. Wi-Fi signals travel relatively far; a single AP may serve hundreds of square meters in an office building, providing data coverage for several users and offices. As a result, there can be significant distance, between the AP and the receiving device that introduce variable attenuation of the signal.

In contrast to the longer range Wi-Fi signal, we also porate small Bluetooth low energy (BLE) beacons. wanke "Classic Bluetooth" r a signal can travel a significant number of meters, devices can be configured such that the signal only travels a few meters. The beacons we use (Fig. 1) are based on the BlueGiga BLE112 SoC. These devices transmit and receive on the The ISM radio band, the same band used by 802.11 The transmit power of these beacons is an order of nitude lower than that of Wi-Fi. To addr nterferthe beacons divide the 2.4-Ghz band into the annels of 2 Mhz each from 2402 to 2480 Mhz. Three of these nels are dedicated to advertising packets and the Lining 37 to reliable data exchange in the connected mode.









The Bluetooth low energy beacons used as part of the LoCo

tising event, the advertising packet is sent on each of pree advertising channels sequentially. The advertising packet is are configured at frequencies largely unused by non Wi-Fi base stations. The time delay between advertising events is configurable to help manage battery consumption. In our deployment, configurable RF transmit and the advertising packets were setting, the device only consumes 27 mA of power to transmit and 0.5 μA when sleeping, thus allowing it to operate for over a year on a single CR2032 coin cell battery, or for several years when configured with a two "AA" cell battery pack.

experiments, we covered approximately superior of magnitude legislation of the simplicity of these devices, it is likely that their price will fall significantly as the economies of scale improve.

While we built out our beacons to provide some useful ionality (unique payloads with batter level, large ry capacity, etc.), they are relatively similar in functionality to Apple's iBeacon specification. With slight modification to our client software, a deployment of off-the-shelf OEM iBeacons could be used with our approach. As these beacons become more popular, we believe it is further evident that our proposed system could have wideranging applicability.

mon 3D printer (see Fig. 2). We coated the inside of the cases with a special graphite paint formulated to absorb 90 dB of RF energy. This allows us to control, in a way, the propagation of the BLE signals. Since the pns are deployed by attaching them to the ceiling (as in Fig. 3), the coating helps reduce signals

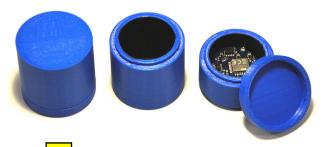


Fig. 2 Tweetes that we custom built using a 3D printer. The cases are made of ABS plastic, coated with a special graphite paint that is hed to absorb RF signals. Using this coating, the cases control, in crude way, the propagation of the BLE signals

traveling between rooms along the rails that hold the lling tiles in place. A variety of case heights is used, ranging from 10 mm for large spaces, to 45 mm for small, cubicle-sized locations. Experiments showed the distance the BLE antenna to the aperture created by the ated portion of the case gave us limited control of the desk/floor area a beacon's signal would cover. Figures 4 5 show the placement of the Wi-Fi APs and BLE used in our trial deployment and experiments. luegiga BLE112 firmware provides for control the power output of the Bluetooth radio signal ifier in the arbitrary range "1–15". In the initial configuration, we set the power to "1" to localize the radio transmission as much as possible. This was the setting for the beacons for the first experimental data set we collected described below. However, this configuration was npatible with the iBeacon use case which we rely on gger events in iOS clients. As a result, we increased the radio output to "5", thus increasing the range and visibility of the beacons as evidenced in the second data set we collected for our experiments.

### 3.2 Smartphone client service

latively simple service runs on users' smartphones to ct Wi-Fi and BLE signal measurements. We created nitial implementation of this scanning service using le's Android OS. We use the internal API to retrieve a f in-range Wi-Fi APs and the RSSI from those APs' D advertisements. A native Bluetooth low energy API not released until Android 4.3 (API level 18). To support devices not running 4.3, we also added support for e BLE APIs that have been n available by the us device manufacturers. For this enabled popular, yet older phones like the Samsung Galaxy SIII devices (model I-930 nning Android 4.1.2 to be comle with LoCo. We design the SIII as well as LG Nexus 4 G Nexus 5 devices in our experiments and evaluation below.







Fig. 3 A sketch depicting how the BLE beacons are deployed on ceilings in our environment

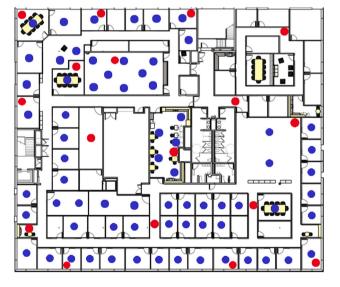
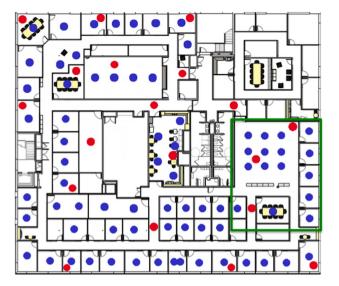


Fig. 4 Floor plan showing the placement of both the BLE beacons (blue dots) and Wi-Fi APs (red dots) used in our first data collection.

that we did not deploy beacons in closets, machine/server
, or restrooms (color figure online)

To make the LoCo framework available to application opers, the Android is implemented as an oid service. With few an 30 lines of code, an app



Floor please wing the placement of both the BLE beacons APs (red dots) used in our second experiment (larger deployment). The region enclosed by the green rectangle is for the smaller scale experiment focused on the open-plan area right side (color figure online)

developer can create a dependence on LoCo and make queries to the service for the device's location. Centralizing in a service not only makes LoCo easily available to ple apps, the model also ensures that only one eliminates duplicate scanning that can reduce battery eliminates. When an application subscribes to LoCo, it also specify how often the framework should deterits location.

## 3.3 sification engine

The classifier along with a trained definition library can be don to the smartphone allowing classification to be has low computational complexity, allowing it to run on smartphones without consuming a significant amount of structure to a cloud-based classification server. With this method, classification can be performed and stored on a server, allowing for external applications to leverage web services to query for a device's location.

stimate the device's room. While straightforward hing methods such as [6] provide high-level perforte, they typically require the storage of a search data application of specialized data structures with nonmetric distance measures can be difficult







Accelerated versions of kNN also scale poorly with re dimensionality, limiting expansion of the location system to areas with large numbers of APs and beacons.

erest; however, the method is applicable to any discrete quantization. Specifically, multiple rooms can be collinto a single location, and a single room can be subdivided into multiple locations for classification purposes; this ires only appropriate corresponding modifications to the procedure. This is explored in Sect. 4.2

tire ethods several aling attributes, including bicity, including vised classification method, the bulk of its computations in an off-line training procedure. While binary boosting algorithms have been extensively analyzed and multiple maximum series and employed procedure. While binary boosting algorithms have been extensively analyzed and multiple maximum series and extensively analyzed and employed procedure. While binary boosting algorithms have been extensively analyzed and employed employed entire training procedure. While binary boosting algorithms have been extensively analyzed and employed entire training procedure. While binary boosting algorithms have been extensively analyzed and employed entire training procedure. While binary boosting algorithms have been extensively analyzed and employed entire training procedure. While binary boosting algorithms have been extensively analyzed and employed entire training procedure. While binary boosting algorithms have been extensively analyzed and employed entire training procedure. While binary boosting algorithms have been extensively analyzed and employed entire training procedure. While binary boosting is a supplied to the procedure of the set of procedure and the procedure is a supplied to the procedure of the procedure in the procedure is a supplied to the procedure of the procedure in the procedure of the procedure in th

$$\operatorname{room}^*(S) = \underset{\operatorname{room}}{\operatorname{argmax}} F_{\operatorname{room}}(S) \tag{1}$$

ach room, we construct a binary classifier that outputs

pre representing the probability that the RSSI scan

r S was observed in that room:

$$F_{\text{room}}(S) = \sum_{m} \alpha_m h_m(S). \tag{2}$$

per-room classifier combines "weak learners",  $h_m$  ding to the scalar weights  $\alpha_m$ . The weak learners are ion stumps that compare a scalar feature to a threshold  $\theta_m$ :

$$h_m(S) = \begin{cases} 1 & X_m \ge \theta_m \\ 0 & \text{otherwise} \end{cases}$$
 (3)

ining, the holds  $\theta_m$  are tuned to minimize error. lefine the feature vector comprised of elements  $X_m$  that  $X_m$  from each RSSI vector S below.

ruct the per-room classifiers, we propose margin res. Given the observed RSSI vectors, we compute the f unique pairwise differences (margins) between the peacons, the peacons, the peacons, the peacons of the peacons.

ecifically, we transform the RSSI vector  $S \in \mathbb{R}^B$  into a in feature vector with elements:

$$= S(a_m) - S(b_m), \tag{4}$$

 $b_n, b_m \in \{1, \dots, B\}$ . Missing RSSI values for specific Ds or beacons in the training set are set to a nominal

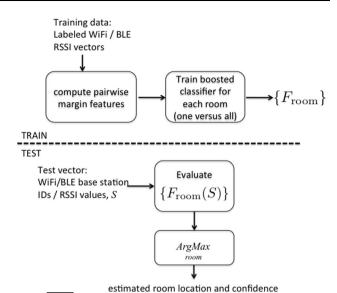


Fig. 6 Typhic describing the data flow using boosted classification margin features for room identification. Processing above the line (TRAIN) is performed on line, given a set of labeled training data. Processing below the describing (TEST) is performed for each test scan

 $R_{\text{min}}$  to indicate they were not visible to the mobile to the mobile to the fact that specific beacons or BSSIDs are visible at specific locations is incorporated into the features.

The mean features computed from the training scans form the input to classifier training. The training procedure ifies a location-specific set of weak learners of the in (3) that best discriminates that location from all s. The weak learners and their relative weights ( $\alpha_m$  in are learned in a greedy iterative procedure that optimizes error using a per-sample weighting over the training [14]. These per-sample weights are repeatedly adjusto emphasize misclassified samples. This process is repeated for each location to construct the set of classifiers  $\mathbf{m}$  in Fig. 6. The output of training  $F_{\text{room}}$  is the ification of its selected weak learners,  $h_m$  and their thts  $\alpha_m$  in (2). Each weak learner is determined by two ware IDs  $a_m, b_m$  (MAC addresses of specific Wi-Fi Ds and BLE beacons) and a threshold,  $\theta_m$ . We employ lticlass extension of the discrete adaboost classifier ed in [41] (SAMME—algorithm 2). ocation determination, each of the per-room clas-

s is applied to a given test RSSI scan vector  $S_{\text{test}}$ . We pute only the required set of RSSI differences of (4) were selected in classifier training. These differences are compared to the thresholds  $(\theta_m)$  determined during training and then combined linearly per (2). The final room estimate is determined by comparing the scores  $m(S_{\text{test}})$  and selecting the maximum per (1). If any BLE beacons or Wi-Fi BSSIDs that contribute to a per-







room classifier of present in  $S_{test}$ , the missing values are set to  $R_{min}$ . The following of the computation occurs in the off-line training of the computation occurs in the off-line training dure. The training complexity is governed by the size of training samples. The test complexity depends on the number of *selected* weak learners. Thus, unlike hing methods, classification test complexity is upled from the training set size.

### 4 Experiments

eport results of experiments designed to investigate:

- classification performance and time complexity,
- spatial granularity for defining locations,
- beacon and Wi-Fi access point density, and
- distance-based error analysis.

a single floor of our office. The Revalues were recorded by a variety of Android mobile devices that included Samsung Galaxy SIIIs, LG Nexus 4s, and LG We first assess the utility and limitions of the BLE mark our boosting method for room-level location classification against existing methods.

Because the BLE beacons are explicitly deployed for location sensing, we consider first a simple baseline that associates each location with the hardware ID of the beawith maximal RSSI in its corresponding scans in the ng set. At classification time, we simply determine the beacon with maximal RSSI in each test scan and then assign the location associated with that BLE as the estimate. This simple method provides the first baseline od for comparison and is denoted "MAX BLE" in each test scan and then assign the location associated with that BLE as the estimate. This simple method provides the first baseline of for comparison and is denoted "MAX BLE" in each test scan and then assign the location associated with that BLE as the estimate. This simple method provides the first baseline of for comparison and is denoted "MAX BLE" in each test scan and then assign the location associated with that BLE as the estimate. This simple method provides the first baseline of for comparison and is denoted "MAX BLE" in each test scan and then assign the location associated with that BLE as the estimate. This simple method provides the first baseline as and Figures below. We also evaluated versions of each test scan and then assign the location associated with that BLE as the estimate. This simple method provides the first baseline with the BLE as the estimate. This simple method provides the first baseline as and Figures below. We also evaluated versions of each test scan and then assign the location associated with that BLE as the estimate. This simple method provides the first baseline as and Figures below. We also evaluated versions of each test scan and then assign the location as a state-location as a state-

$$d_{RP}(X,Y) = \alpha \sum_{b} [(X(b) > 0) \land (Y(b) > 0)]$$

$$-\beta \sum_{b} [(X(b) > 0) \oplus (Y(b) > 0)]$$

$$+ \gamma C(X,Y).$$
(5)

in X and Y. The summation penalizes that appear in both X and Y. The cond summation penalizes

Ds/beacons that are visible in exactly one of X or Y. We egated euclidean distance as the correlation measure C, set  $\alpha = 1.0, \beta = 0.4, \gamma = 0.2$ . For classification, brute kNN search is executed for the given test RSSI vector. classification is determined by the majority vote among earest training samples. Redpin was developed for use with Wi-Fi RSSI data; however, we also report results for the method using the BLE RSSI data to provide a broader sense of eacons' utility. We also naively concatenate the Wi-Fi E RSSI data to establish a combined Redpin baseline. he boosting classification system described here, we classifiers using the BLE and Wife SSI data both idually and in combination. margin features from the RSSI data following (4), and run 45–90 iterations of clastic training off-line to learn each per-location classifier. esting, we apply the per-location classifiers to each test scan and \_\_\_\_n the highest scoring location as the estimate. Unles Herwise noted, we report accuracy and timing results averaged over fold cross-validation, and train all boosting classifiers 0 iterations. Wall timing results here are computed using a PC with a 2.8-GHz AMD processor.

### 4.1 Classification accuracy and efficiency

The first data set consists of 1181 scans collected in 55 ions in which RSSI values are observed from a total of LE beacons and 159 unique Wi-Fi BSSIDs. The large er of Wi-Fi devices is the result of two conditions. the 17 APs that we used are commercial Cisco APs service multiple networks and provide access Liple frequencies (mainly channels on lind 5 GHz). These APs produced 42 unique BSSIDs. 5 because our office shares a building with other companies and has close neighbors, the remaining unique BSSIDs are these organizations' APs. The data set contains an ge of 19.05 scans per location (SD = 5.21). Table 1 summarizes the experimental results for the data set. The right column shows the average time red to classify one test scan. The top row (MAX/BLE) s results using the predefined map of the BLE bea-This approach simply determines the maximum BLE RSSI and returns the corresponding location of the BLE as the estimate. It has lower accuracy, but demonstrates the value of a simple and efficient method using the BLE beacons. The next three rows show results for the Redpin method using BLE beacons alone (Redpin/BLE), Wi-Fi alone (Redpin/Wi-Fi), and the combined mode (Redpin/ BLE + Wi-Fi). The bottom three rows show corresponding results for location classification via boosting. Throughout, se boldface to indicate the result with the highest racy in the tables.



Table 1 Classification accuracy and (wall) timing results for the first data set with 55 locations and 1181 scans

Method	RSSI data	Accuracy	Time (s)
MAX	BLE	0.757	0.00072
Redpin [6]	BLE	0.952	1.354
	Wi-Fi	0.913	3.054
	BLE + Wi-Fi	0.930	4.335
Boosting	BLE	0.937	0.00492
	Wi-Fi	0.943	0.00562
	BLE + Wi-Fi	0.966	0.00431

Accuracy is computed using ninefold cross-validation. Test timing is averaged per RSSI scan in seconds. The boosting results here use 45 iterations of classifier training

Table 2 Classification accuracy and (wall) timing results for the second data set with 56 locations and 2294 scans

Method	RSSI data	Accuracy	Time (s)
MAX	BLE	0.629	0.00159
Redpin [6]	BLE	0.964	2.0835
	Wi-Fi	0.950	7.2236
	BLE + Wi-Fi	0.973	9.165
Boosting	BLE	0.965	0.00627
	Wi-Fi	0.956	0.00653
	BLE + Wi-Fi	0.991	0.00660

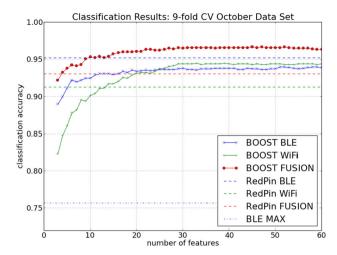
Accuracy is computed using ninefold cross-validation. Test timing is averaged per RSSI scan in seconds. The boosting results here use 90 iterations of classifier training

The single modality results show consistent levels of accuracy for both Redpin and boosting in the range 0.95. The Redpin BLE-only results outperform the BLE-only results (boosting/BLE) by a small margin, while boosting outperforms Redpin using Wi-Fi only. Also, BLE and Wi-Fi are able to provide complementary information that is exploited by the boosting classification method in the fusion (boosting/BLE + Wi-Fi) results to improve accuracy. The BLE + Wi-Fi boosting classifier performs the best among the evaluated methods with significantly lower computational overhead than Redpin systems.

number of BLE beacons, 74 and the number of Wi-Fi AP pn IDs, 254. In addition, data set uses the BLE ger BLE RSSI values. The mean BLE RSSI value ger BLE RSSI value at the higher power st data set is -95.012 (std = 4.41), while for these data set is -91.157 (std = 6.35)<sup>4</sup>. Between the collections of the two data sets used, our company

<sup>&</sup>lt;sup>4</sup> RSSI is measured on a logarithmic scale.





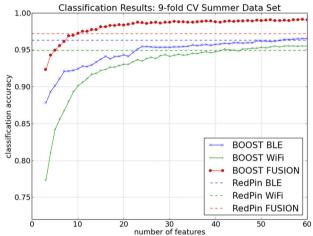


Fig. 7 The plots show the accuracy of different versions of both baselines and our system on the first (top) and second (bottom) experimental data sets. The boosting classification systems show ion with the number of training iterations (selected weak along the x-axis, while the baseline systems are horizontal lines in the plots

deployed additional Wi-Fi access points to support unrelated research activities. This data set consists of 2294 total RSSI scans collected using the same Android devices as e, with an average of 31 scans per location 5.73). The experimental results appear in Table 2 and include the same baseline and boosting location classification configurations as before.

Figure 7 graphically displays performance on the two data sets for the boosting location classifiers. As more RSSI margin features and their corresponding weak learners are integrated, performance improves. The boosting/BLE system asymptotes relatively quickly with approximately twenty weak learners in the left panel. This is because fewer BLE beacons are visible from each location due to their positioning, lower power, and shielding. The lower panel shows that the classifiers use more of the higher

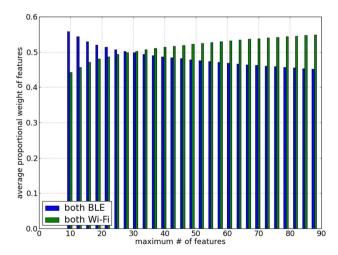
powered BLE beacons in the second data set. Notice that the increased range of the Breacons degrades the performance of the map-based //BLE method relative to the first data set. This is because on average 9.86 = 3.62) beacons are visible in each scan in the second set compared to only 2.95 (std = 1.32) beacons in the first data set.

In contrast, the boosting/Wi-Fi system shows more steady improvement with the addition of weak learners.

The reflects the fact that many more Wi-Fi BSSIDs may be useful for classification. The fusion also shows some gradual improvement and reaches accuracy with about thirty weak learners. The horizontal dashed lines show the performance for the Redpin and MAX baselines.

For the first data set, the Redpin/BLE system performs at a higher accuracy than the boosting/BLE classifier, while on the second data set they perform at the same level. For both data sets, the Redpin/BLE system shows excellent performance and is superior to the Wi-Fi-only variant. This trend is also evident in the boosting results, though with the inclusion of a large number of weak learners, the BLE-only and Wi-Fi-only performance converges to a similar level. These experiments show that BLE beacons offer a competitive alternative to Wi-Fi for room-level location identification.

both experiments, the boosting system that combines and Wi-Fi data shows the highest accuracy. For fusing the RSSI data, we compute margin features from each lity independently and combine them for classifier ng. Figure 8 visualizes the order of selection and



**Fig. 8** The plots show the relative weighting of BLE-based and Wi-Fi-based features in the combined boosting classifiers on the first (*top*) and second (*bottom*) experimental data sets. The weights are first normalized per room, and the proportions allocated to BLE and Wi-Fi are averaged per room and per fold to generate the plots

relative proportion in which the two modalities are used by posting classifiers on the second larger data set. ically, it shows the average normalized weights ( $\alpha_m$ in (2)) associated with BLE margin features and Wi-Fi in features. The horizontal axis represents the maxinumber of decision stumps allowed in each classifier. The weights are first summed per location and used to normalize the proportion allocated to each modality. These results are then averaged over all locations and cross-validation folds. The plots show that BLE-based features are greedily selected first, but because of the abundance of Wita and its larger range, more and more Wi-Fi features dded to the classifiers in subsequent iterations. This suggests that the BLE information is more discriminative than Wi-Fi as expected, since it is deployed explicitly for location classification. However, as training progresses, the relative density and longer range of Wi-Fi provide complementary information.

### 4.2 Spatial granularity

ditional experiments, we compared these methods for location classification at a finer spatial granularity. One portion of our office includes nine cubicles of size 2.5 by 2.5 m arranged in a three-by-three grid. The area is separated from a row of five conventional (walled) offices by a vay 1.5 m wide, and from a larger conference room by The area is enclosed in a green rectangle in Fig. 5. s area, we prmed location classification of the cubicles and ther rooms based on BLE, Wi-Fi, and their combination. As before, there is a BLE beacon located on the ceiling in each location, and various Wi-Fi BSSIDs visible both from our internal network and from neighboring organizations. We sampled a data set comprised of 445 scans as part of the larger second data set ed above. The scans include RSSI data from 17 BLE acons in the building and 163 unique Wi-Fi BSSIDs.

Experimental results for all previously evaluated baseline and boosting variants appear in Table 3. The results for

3 Classification accuracy results for the open space data set 5 locations and 445 scans

Method	RSSI data	Accuracy
MAX	BLE	0.504
Redpin [6]	BLE	0.818
	Wi-Fi	0.876
	BLE + Wi-Fi	0.907
Boosting	BLE	0.818
	Wi-Fi	0.925
	BLE + Wi-Fi	0.955

Accuracy is computed using sixfold cross-validation and 45 iterations of boosting classifier training



BLE alone in this setting show a decline from the earlier riments, suggesting that there are limitations to their al resolution for location classification. In particular, relying on the kimal observed BLE RSSI shows poor performance. age, each scan contains RSSI meanents from  $\frac{1}{1000}$  BLE beacons (std = 2.45). The ting classifiers that use Wi-Fi alone, or in combination with BLE show improved performance relative to Redpin in this setting. The boosting classification method with the complete set of measurements shows the best performance at 95 % accuracy for this higher granularity. Thus, the boosting classifiers again offer both an efficiency advantage and superior accuracy to the Redpin results using Wi-Fi only or together with BLE at this higher spatial granularity.

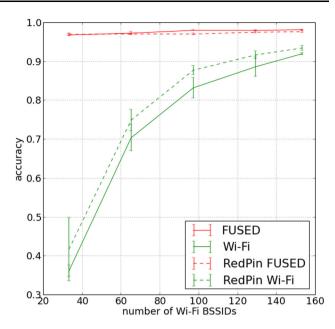
### 4.3 Wi-Fi access point density

A key to the relative efficiency of boosting in this appliin is the feature selection inherent in its classifier
ing. The increase in the number of visible hardware
IDs moving from the first to the second data set produced a
inficant complexity increase in the Redpin systems. On
ther hand, because the boosting system selects a small
t of margin features and corresponding hardware IDs
each room classifier, the complexity remained
stable and low for both data sets.

We performed additional experiments to investigate the impact of Wi-Fi density on the performance of these sysve made two changes to our experimental col. Fitwe used the second experimental data set, estricted the Wi-Fi information to accestints from rganization (152 of 254 total BSSI performed repeated experiments using standard cross-validation by randomly sampling a proportion of the access s and withholding all their associated RSSI measures during classifier training.<sup>5</sup> Here the error bars are buted over each fold and iteration. We graphically depict the results in Fig. 9. To create each point in the plots, we perform four iterations of random selection of access points to simulate lower Wi-Fi density. The plot s accuracy versus the number of Wi-Fi BSSIDs hable. Generally, the Redpin system outperforms boosting here by a small margin, but boosting is competitive over a wide range of Wi-Fi densities. The fusion results show that the boosting system does very slightly better than Redpin over the bulk of the experimental conditions, and both use BLE to provide high performance in relatively low-density Wi-Fi environments. As before, the

<sup>&</sup>lt;sup>5</sup> Multiple BSSIDs are associated with a single Wi-Fi access point. Here, we include and withhold BSSIDs by access point.





**Fig. 9** The plot shows the accuracy of Wi-Fi-only and combined classification using Redpin and boosting as the number of visible Wi-Fi hardware IDs varies (*x*-axis)

boosting system operates with substantially reduced computation.

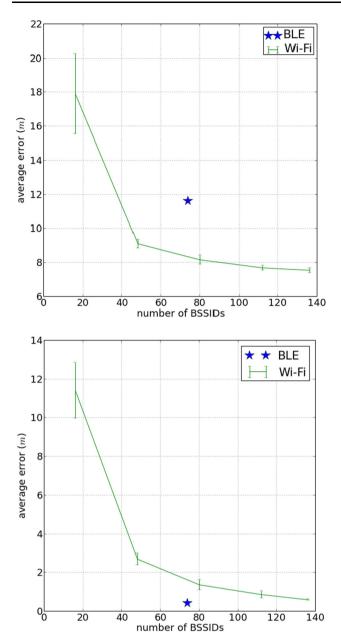
### 4.4 Distance-based error analysis

BLE and Wi-Fi with the boosting classification ach. For these experiments, we first associate a fic distance-based cost to each possible misclassifin, in contrast to the simple binary cost that defines the racy measure we have used to this point. For this, we pute distances between the centroids of each pair of locations in the second data set. Note that the analysis here the directly comparable to distance-based error for lateral systems because we still are using room-level ground truth rather than absolute position.

As mentioned earlier, because the BLE beacons are explicitly deployed for location classification, we expect that the mistakes BLE-based methods make will most often be other nearby locations relative to Wi-Fi. The results in Fig. 10 affirm this intuitive power a range of Wi-Fi densities simulated as before. Wi-Fi base stations are included, the variance shown by the error bars is higher. This is because the access points only effectively cover some subset of locations. As more and more base stations point comparable to and eventually lower than the BLE resulting the point in the plots.

Here, gain average the distance-based cost of all misclassifications averaged over the entire data set, again





**Fig. 10** The plots show the average misclassification distance in meters for Wi-Fi as the number of access points and their associated BSSIDs in the available RSSI data is varied. The single point represents the BLE performance. The *upper plot* shows the average error for misclassified scans only. The *lower panel* shows the average error results for all scans, including correctly classified results

using sixfold cross-validation and four iterations at each Wi-Fi AP density. We examine the distance-based error in two plots. In the top plot, we compute the average distance over only misclassified samples. In this case, the *number* of misclassifications does not influence the results, and BLE shows higher error (11.63 m) relative to Wi-Fi (approximately 9 m) at the same relative density. In the lower panel, we average the error over all scans, including correctly classified samples, and see that in this case BLE

shows a lower average error, because it makes fewer errors than Wi-Fi. As the Wi-Fi access point density increases, the distance-based error converges to a level (average error of 0.555 m) close to that of the BLE-only boosting system (average error of 0.405 m). The fusion system that uses BLE with Wi-Fi reduces this average error to 0.1172 meters, again demonstrating the complementary nature of these modalities for indoor location classification.

### 4.5 Experimental summary

nts validate our approach to location clasthe accuracy of the boosted classifiers sification. s a high level of performance using BLE and Wi-Fi individually and in combination. We show that the explicit deployment of BLE hardware for indoor location classification is effective and competitive with Wi-Fi-based methods using both Redpin and boosting. Distance-based rsis of the BLE-only classifier's mittees also shows a r average cost relative to Wi-Fi. S d, the boosting pach shows consistent performance gains by combin-LE with Wi-Fi information to demonstrate that BLE beacons can be deployed to complement existing Wi-Fi infrastructure for high-performance location classification. Additionally, analysis of the boosting classifiers and their composition and weighting of weak learners further affirms omplementary nature of the BLE and Wi-Fi RSSI urements. Finally, the error analysis above using nce validates the improved accuracy results offered by pining BLE and Wi-Fi. There are both fewer errors in terms of accuracy, and also lower cost errors in terms of distance, made by the combined classifier relative to any single modality result.

Importantly, the boosted classifiers off-load the bulk of omputational complexity to an off-line training proThe test complexity of the boosted classifiers is istently orders of magnitude faster than the Redpin is. While more effort could be made to optimize both the kNN and boosting code, these results demonstrate the significant computational savings provided by supervised infication relative to the matching approaches prevalent ior work.

### 5 Discussion

LoCo Framework, along with our analysis of its perance, demonstrates that it is possible to perform rate room-based location classification using practical, accessible methods. While our work is not the first to investigate leveraging multiple sources of radio signals to mine device location, we do believe we are the first to instrate the efficacy of a hybrid radio approach with



ting that achieves the level of accuracy and perforted demonstrated by our experiments. This is a significant step forward in providing a usable platform for creating location applications and services for settings such as the workplace.

We believe LoCo can have immediate utility in creating next generation location applications and services. A particular motivation for us is to drive improvements in office communication tools and technologies. For nce, providing real-time location information of coles in chat, email, and other communication applications is now easily implementable with LoCo. We have leveraged LoCo to enhance an existing office presence [38], and we are now planning a study to understand improved location information translates to meaningful changes in communication behavior. Even further, p could be leveraged to build important office safety ms. For example, a system that can inform building security persent if employees are present in the building after hours. arly, in the event of a fire or earthquake, the last known location of workers could be provided to first responders.

There are many additional applications for indoor localization. There have been various integrations of indoor localization with audio guides in exhibition spaces or museums [15, 34, 36]. Further applications include indoor location-based gaming [10, 33], location-aware shopping assistance [12, 42], location-based advertising [1], and indoor preserve awareness [2, 37, 38].

While are advantages introduced by our framework, it portant to note its limitations. A particularly parameter in process is simple but involves taking a device into room having it perform several scans that it submits to the process is simple but involves taking a device into room having it perform several scans that it submits to the feation engine with an attached ground truth label. It is not replaced, a new survey is likely necessary. The process are, however, several ways to address these issues. Open proach, as first proposed by [6], is to crowd source users in providing ground truth. For instance, when a user notices a location is not classified, or not classified correctly, he could simply tell the client application the correct

location and the particular scan could be sent to the classification engine as additional piece of data used in the training process. A proposed, which we are actively investigating, is to use a small robot to perform a survey. With this method, a new, up-to-date model can be created on a regular, ongoing basis with relatively low human overhead.

determine the precise location within a room a device is located. In large rooms, this would limit the utility of leveraging LoCo to classify activities of individuals or groups, or to actuate changes or events within the physical. Our belief is that achieving these goals will likely be through further location information with other sensor sources.

LoCo location information within-room methods of precise within-room techniques into the LoCo Framework.

### 6 Conclusion and future work

Location is playing an ever-increasing role in mobile computing. Many of the most popular applications and services used today exploit knowledge of the user's current location. For many of these applications, the accuracy of current technologies is adequate. However, we believe the next generation of mobile services will demand accurate, reliable location information in indoor environments. In this paper, we presented the LoCo framework. We discussed how it has been designed and developed to provide application and service developers access to indoor location information. We demonstrated an evaluation of the system and explained its performance and capabilities compared to existing state-of-the-art techniques. Further, we performed this evaluation using hardware and software that the comparison of the system and explained its performance and capabilities compared to existing state-of-the-art techniques. Further, we performed this evaluation using hardware and software that the comparison of the system and capabilities compared to existing state-of-the-art techniques.

we intend to deploy LoCo in different environments to understand its performance characteristics more broadly. This effort will also allow us to better investigate how well LoCo scales to larger configurations, for instance, within adensely urban environment. It is in a densely urban environment. It is in a densely urban environment. It is in a densely urban environment and novel applications that use LoCo's accurate location information to provide interesting and configurations and retail settings. It is in a densely urban environment and novel applications that use LoCo's accurate location information to provide interesting and configuration information to provide interesting and configurations that use LoCo's accurate location information to provide interesting and configurations that use LoCo available to others in the research community.



#### References

- Aalto L, Göthlin N, Korhonen J, Ojala T (2004) Bluetooth and WAP push based location-aware mobile advertising system. In: Proceedings of the 2nd international conference on mobile systems, applications, and services, ACM, pp 49–58
- An X, Wang J, Prasad RV, Niemegeers I (2006) OPT: online person tracking system for context-awareness in wireless personal network. In: Proceedings of the 2nd international workshop on Multi-hop ad hoc networks: from theory to reality, ACM, pp 47–54
- Andoni A, Indyk P (2006) Efficient algorithms for substring near neighbor problem. In: Proceedings of the seventeenth annual ACM-SIAM symposium on discrete algorithm, SODA '06, ACM, New York, NY, USA, pp 1203–1212. doi:10.1145/ 1109557.1109690
- Aparicio S, Perez J, Bernardos A, Casar J (2008) A fusion method based on Bluetooth and WLAN technologies for indoor location.
   In: IEEE international conference on multisensor fusion and integration for intelligent systems, 2008. MFI 2008, pp 487–491
- Biehl JT, Cooper M, Filby G, Kratz S (2014) Loco: a ready-to-deploy framework for efficient room localization using Wi-Fi. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing, UbiComp '14, ACM, New York, NY, USA, pp 183–187. doi:10.1145/2632048.2636083
- Bolliger P (2008) Redpin-adaptive, zero-configuration indoor localization through user collaboration. In: Proceedings of the first ACM international workshop on mobile entity localization and tracking in GPS-less environments, MELT '08, ACM, New York, NY, USA, pp 55–60. doi:10.1145/1410012.1410025
- Chintalapudi K, Padmanabha Iyer A, Padmanabhan VN (2010) Indoor localization without the pain. In: Proceedings of the sixteenth annual international conference on mobile computing and networking, MobiCom '10, ACM, New York, NY, USA, pp 173–184. doi:10.1145/1859995.1860016
- Cong L, Zhuang W (2002) Hybrid tdoa/aoa mobile user location for wideband cdma cellular systems. IEEE Trans Wirel Commun 1(3):439–447. doi:10.1109/TWC.2002.800542
- Dempster A, Li B, Quader I (2008) Errors in deterministic wireless fingerprinting systems for localization. In: ISWPC 2008, pp 111–115
- Ferdinand P, Müller S, Ritschel T, Wechselberger U (2005) The eduventure—a new approach of digital game based learning combining virtual and mobile augmented reality games episodes.
   In: Pre-conference workshop? Game based Learning? of DeLFI 2005 and GMW 2005 conference, vol. 13. Rostock
- 11. d Y, Schapire R (1997) A decision-theoretic generalization line learning and an application to boosting. J Comput Syst Sci 55(1):119–139. doi:10.1006/jcss.1997.1504
- Gross HM, Böhme HJ, Schröter C, Müller S, König A, Martin C, Merten M, Bley A (2008) Shopbot: progress in developing an interactive mobile shopping assistant for everyday use. In: IEEE international conference on systems, man and cybernetics, 2008. SMC 2008. pp 3471–3478
- Harter A, Hopper A, Steggles P, Ward A, Webster P (2002) The anatomy of a context-aware application. Wirel Netw 8(2/ 3):187–197
- 14. T, Tibshirani R, Friedman J (2001) The elements of sta-
- Heller F (2008) Corona: realizing an interactive experience in visually untouchable rooms using continuous virtual audio spaces. Master's thesis, RWTH Aachen University
- Hightower J, Want R, Borriello G (2000) Spoton: An indoor 3D location sensing technology based on RF signal strength. UW

- CSE 00-02-02, University of Washington, Department of Computer Science and Engineering, Seattle, WA 1
- Li P (2010) Robust logitboost and adaptive base class (abc) logitboost. In: UAI 2010, proceedings of the twenty-sixth conference on uncertainty in artificial intelligence, pp 302–311. arXiv:1203.3491
- Lin H, Zhang Y, Griss M, Landa I (2009) Enhanced indoor locationing in a congested Wi-Fi environment. Tech. Rep. MRC-TR-2009-04, Carnegie Mellon Silicon Valley
- Liu H, Darabi H, Banerjee P, Liu J (2007) Survey of wireless indoor positioning techniques and systems. IEEE Trans Syst Man Cybern Part C Appl Rev 37(6):1067–1080
- Madigan D, Elnahrawy E, Martin RP, Ju WH, Krishnan P, Krishnakumar AS (2005) Bayesian indoor positioning systems.
   In: Proceedings of the 24th joint conference of the IEEE Computer and Communication Societies (INFOCOM 2005)
- Martin E, Vinyals O, Friedland G, Bajcsy R (2010) Precise indoor localization using smart phones. In: Proceedings of the international conference on Multimedia, MM '10, ACM, New York, NY, USA, pp 787–790. doi:10.1145/1873951.1874078
- Mukherjee I, Schapire R (2010) A theory of multiclass boosting.
   In: Lafferty J, Williams CKI, Shawe-Taylor J, Zemel R, Culotta A (eds) Advances in Neural Information Processing Systems, vol. 23., pp 1714–1722
- Ni LM, Liu Y, Lau YC, Patil AP (2004) Landmarc: indoor location sensing using active RFID. Wirel Netw 10(6):701–710
- Niculescu D, Nath B (2003) Ad hoc positioning system (APS) using AOA. In: INFOCOM 2003. Twenty-second annual joint conference of the IEEE computer and communications. IEEE Societies, vol. 3., pp 1734–1743. doi:10.1109/INFCOM.2003.1209196
- Olguín DO, Waber BN, Kim T, Mohan A, Ara K, Pentland A (2009) Sensible organizations: technology and methodology for automatically measuring organizational behavior. Trans. Syst. Man Cybern. Part B 39(1):43–55. doi:10.1109/TSMCB.2008. 2006638
- Ouyang RW, Wong AKS, Lea CT, Chiang M (2011) Indoor location estimation with reduced calibration exploiting unlabeled data via hybrid generative/discriminative learning. In: IEEE transactions on mobile computing, vol. 99. (PrePrints). doi:10. 1109/TMC.2011.193
- Pan JJ, Pan SJ, Yin J, Ni LM, Yang Q (2012) Tracking mobile users in wireless networks via semi-supervised colocalization. In: IEEE transactions on pattern analysis and machine intelligence, pp 587–600
- Priyantha NB, Chakraborty A, Balakrishnan H (2000) The cricket location-support system. In: Proceedings of the 6th annual international conference on Mobile computing and networking, ACM, pp 32–43
- Rishabh I, Kimber D, Adcock J (2012) Indoor localization using controlled ambient sounds. In: International conference on indoor positioning and indoor navigation (IPIN), pp 1–10. doi:10.1109/ IPIN.2012.6418905
- 30. Rodrigues ML, Vieira LFM, Campos MF (2011) Fingerprinting-based radio localization in indoor environments using multiple wireless technologies. In: IEEE 22nd international symposium on hal indoor and mobile radio communications (PIMRC), 203–1207
- Roxin A, Gaber J, Wack M, Nait-Sidi-Moh A (2007) Survey of wireless geolocation techniques. In: IEEE Globecom
- Sen S, Lee J, Kim KH, Congdon P (2013) Avoiding multipath to revive inbuilding wifi localization. In: Proceeding of the 11th annual international conference on mobile systems, applications, and services, MobiSys '13, ACM, New York, NY, USA, pp 249–262. doi:10.1145/2462456.2464463



- 33. Sukthankar G (2002) The dynadoom visualization agent: A handheld interface for live action gaming. In: Workshop on ubiquitous agents on embedded, wearable, and mobile devices (conference on intelligent agents and multiagent systems), Bologna, Italy
- Terrenghi L, Zimmermann A (2004) Tailored audio augmented environments for museums. In: Proceedings of the 9th international conference on Intelligent user interfaces, ACM, pp 334–336
- 35. Tod beacons (2014) http://www.todhq.com/
- 36. Wakkary R, Newby K, Hatala M, Evernden D, Droumeva M (2004) Interactive audio content: an approach to audio content for a dynamic museum experience through augmented audio reality and adaptive information retrieval. In: Museums and the web conference
- Want R, Hopper A, Falcão V, Gibbons J (1992) The active badge location system. ACM Trans Inf Syst 10(1):91–102
- 38. Wiese J, Biehl JT, Turner T, van Melle W, Girgensohn A (2011) Beyond 'yesterday's tomorrow': towards the design of awareness technologies for the contemporary worker. In: Proceedings of the

- 13th international conference on human computer interaction with mobile devices and services, MobileHCI '11, ACM, New York, NY, USA, pp 455–464. doi:10.1145/2037373.2037441
- Xiong J, Jamieson K (2012) Towards fine-grained radio-based indoor location. In: Proceedings of the twelfth workshop on mobile computing systems and applications, HotMobile '12, ACM, New York, NY, USA, pp 13:1–13:6. doi:10.1145/ 2162081.2162100
- Youssef M, Agrawala A (2005) The horus whan location determination system. In: Proceedings of the 3rd international conference on Mobile systems, applications, and services (MobiSys '05), pp 205–218
- Zhu J, Zou H, Rosset S, Hastie T (2009) Multi-class adaboost. Stat Interface 2:349–360
- 42. Zhu W, Owen CB, Li H, Lee JH (2004) Personalized in-store e-commerce with the promopad: an augmented reality shopping assistant. Electron J E-commer Tools Appl 1(3):1–19

