

Indoor self-localization via bluetooth low energy beacons

Azkario Rizky Pratama^{1,2} · Widyawan² · Alexander Lazovik¹ · Marco Aiello¹

Abstract Indoor localization is concerned with mapping sensory data to physical locations inside buildings. Location of a user or a mobile device is an essential part of the context, and is therefore very useful for pervasive computing applications. Many proposals exist for solving the localization problem, typically based on image or radio signal processing, though the problem is still generally considered to be open, especially when costs and privacy constraints play an important role. In this paper, we propose a solution based on the emerging Bluetooth Low Energy (BLE) standard and off-the-shelf hardware. Such approach proves to satisfy economic constraints, while challenging in terms of accurate location. To translate beacon signals into locations, we consider several approaches, i.e., cosine similarity, nearest neighbourhood classification, and the nearest beacon. Our experiments indicate a vector based approach as the most suited one. In fact, we show its effectiveness in an actual office deployment consisting of five indoor areas: three multiuser offices, a social corner, and a hallway. We achieve 90% and 80% for accuracy and F-measure, respectively.

Keywords: Ubiquitous computing · Smart buildings · Indoor localization · Bluetooth · Low energy beaconing

1 Introduction

Location information is a central component of the context and in turn of most modern pervasive system applications. Typical scenario for its uses are emergency response,

Azkario Rizky Pratama:
a.r.pratama@rug.nl

Widyawan:
widyawan@ugm.ac.id

Alexander Lazovik:
a.lazovik@rug.nl

✉ Marco Aiello:
a.lazovik@rug.nl

¹ Distributed Systems Group, Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, The Netherlands

² Department of Electrical Engineering and Information Technology, Universitas Gadjah Mada, Indonesia

user navigation, environment optimization, and smart buildings. Depending on the application, the requirements for localization can be quite diverse. Precision in the localization can be traded off with costs of the sensors required for determining location (e.g., for energy saving in smart buildings), while this is not the case for emergency response systems that need reliable context information.

Our main research interest is in the area of smart building with particular emphasis on energy saving [1,2]. In particular, we are interested on approaches that are affordable, least invasive as possible, and sufficiently reliable [3].

In typical current office buildings, Wi-Fi access points (APs) are standard to provide wireless connection to occupants, Received Signal Strength (RSS) from each AP can be used further as an indication of where people are located. While Wi-Fi has been successfully used in buildings over the past two decades, new standards are emerging that can help address the localization problem. Bluetooth Low Energy (BLE) standard, first announced as iBeacon by Apple in 2013, is a low range, low energy communication protocol that extends the widely diffuse bluetooth standard. In BLE, a tiny packet is broadcasted periodically from each node to the surrounding environment, and mobile phones receive packets containing unique IDs and signal transmission strengths. Receiving phones need not authenticate nor be identified, thus helping maintaining privacy. Compared to Wi-Fi, the BLE beacon hardware is of smaller size and much lower energy requirements, thus facilitating its deployment and adoption. Fixed emitting stations, such as Wi-Fi and BLE, act as beacons useful for localization.

1.1 Landmark-based localization

Landmark-based localization utilizes unique features of the physical environment for the purpose of identifying the navigated space. The features can be natural or human artefacts; the features can be passive or actively signalling their unique identifier. Radio based beacons send radio signals on a regular basis with their unique ID. The received signal strength (RSS) is then used by the receiver to infer the location.

The most adopted method to utilize radio based beacons is that of constructing dense fingerprint maps of RSS to have comprehensive picture of distribution of signals in relation to the map of the location. Such an approach is unfortunately not scalable and portable, as it requires major time investments in creating the fingerprint maps.

1.2 Approaches

Given the requirement of having a low-cost, scalable and portable solution to the indoor localization problem, we consider several approaches based on BLE. In particular, we compare three – one, based on the strongest signal strength; second, based on magnitude distance between points; and third, based on angular distance among signal strengths vectors.

1.2.1 Nearest beacon

The simplest to beacon based localization is to consider the strongest signal as the indication of location, thus as the Nearest Beacon (NB). Assuming there are n -beacons deployed in the environment. Given an observation o formed by the readings of all known beacons $o = \{b_1, b_2, \dots, b_n\}$ where $b_i \in \mathbb{R}^n$ is a RSS, then the location is the location assigned to $\max(o)$. This approach is based on the fact that, in open space environments, the RSS has a direct correlation with the distance.

1.2.2 K-Nearest neighbor

The Nearest Neighbor (NN) algorithm considers the Euclidean distance in a feature space and classifies according to the minimal distance of the observation with known feature points. In our case, assuming that there are c -room classes, for each class, we arbitrarily collect k -instances as reference data, where each instance is a set of beacon readings $r_i^c = [b^1_i, b^2_i, \dots, b^n_i]$ each one of which is associated with a location.

Then, given observation o , we compute the Euclidean distance for all elements of k and take (1).

$$\min(\text{distance}(o, r)) = \sqrt{\sum_{i=1}^n (o_i - r_i)^2} \quad (1)$$

1.2.3 Cosine similarity

Considering the feature space formed by all possible beacon values, one can consider reading as a vector. By considering the angle between the read vector and vectors representing locations, one can classify the location of the user.

Let $v = [v_1, v_2, \dots, v_m]$ be a set of vectors representing certain locations and each location can be represented by one or more vectors. Given observation o , we can compute the cosine distance using (2).

$$\theta_{o, v_i} = \cos^{-1}\left(\frac{o \cdot v_i}{|o||v_i|}\right) \quad (2)$$

From experimentation, one then defines a threshold value γ used to decide location. This is used to define a step function of location membership based on the cosine distance between the read vector and the references in the following way:

$$\cos \theta(v_i, v_j) = \begin{cases} \text{united, if } (\gamma > \cos \theta \geq 1) \\ \text{spreaded, if } (0 \geq \cos \theta \geq \gamma) \end{cases} \quad (3)$$

1.3 Contribution and paper organization

We demonstrate a scalable and portable indoor localization solution that exploits readily available resources and features of the emerging BLE standard. To increase

its portability, we do not specify predetermined rules for reference data collection. Thus, we accept a lower accuracy in the classification, to avoid long and tedious predefined mapping of the deployment spaces. We address the possible shortcomings of such an approach by investigating an approach to classification based on the cosine distance.

The paper is structured as follows. In Section 2, we present an overview and the design of the proposed system as well as its concrete implementation. In Section 3 and Section 4, we elaborate on the case study set-ups; and provide results and discussions, respectively. Related work is presented in Section 5 and conclusions are presented in Section 6.

2 Design and implementation

Using an office space actually in use, we deploy a set of BLE beacons and design a system for the localization with the intent of testing several approaches to self-localization.

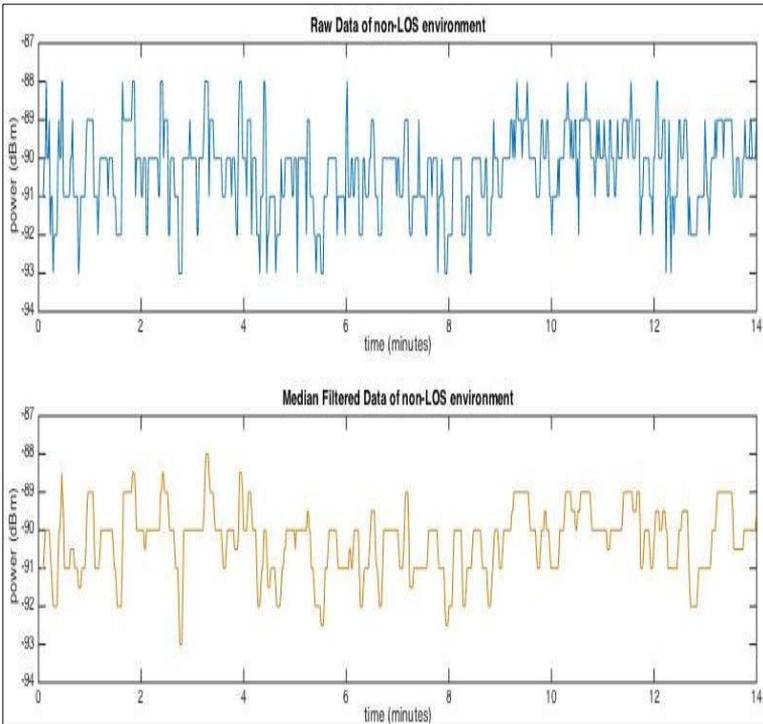


Fig. 1 (a). Raw data; (b). Median filtered data with window size of 6 samples

2.1 Design and assumptions

We set working desks as *observation points* for room offices, that is, locations where the user is likely to reside for a longer period of time. Each class can have one, more than one, or even no observation point. With respect to classes that have no

observation points (e.g., a hallway or a social corner), we define an *observation area*, where reference data is collected while the occupant is moving around.

The underlying assumption is that a person can be represented by one device that is always brought by the individual wherever s/he goes. The device runs a dedicated application and the Bluetooth interface must be on.

2.2 Bernoulliborg implementation

We test our system and several approaches to localization in our own office building, the Bernoulliborg, in the Zernike campus of the University of Groningen. We select our offices on the 5th floor, as shown in Figure 2. There are three office spaces (i.e.: R1, R2, R3), a Social Corner (SC), and Hallway (HW) interconnected among the spaces. To collect raw data, we write an android-based application. The mobile phone measures beacon's signal strength according to the iBeacon protocol [4]. These data are transmitted over Wi-Fi and stored in a centralized database.

We begin by copiously collecting BLE signals in the frequency 1 Hz from a single beacon node in the environment for 14 minutes length. It consists of about 800 samples. Such data set is sufficient to illustrate the signal fluctuation. We find that even though a device is positioned at a fixed position, the RSS fluctuates often (see Figure 1(a)). To have a more stable value, we apply low pass filtering with a specific window size to eliminate outliers from the raw signal, as shown in Figure 1(b). The window size is determined on the basis of the expected time a person will need to transition to a new room. As the rooms are about 7x4 meters, it is reasonable to assume that people need 20-30 seconds to move to another space. Given the 0.25Hz sampling rate, a window size of 6 instances, corresponding to approximately 24 seconds, is a reasonable size.

In addition to using the RSS, we consider the steps count as a feature that detects occupant's movements.

Table 1 Defined parameter

Parameter	Value
Transmitted power	20dBm
Beacon density	2 beacons per-30m
Broadcast	950ms
Listening	2500ms
Waiting	1500ms

We access the value that represents step occurrence in a specific time, value that is typically available in modern mobile OS.

3 Experiment

We perform an experimental evaluation in the Bernoulliborg offices to determine the feasibility and accuracy of localization at the room level based on BLE.

3.1 Setup

As proximity based localization system, Bluetooth Low Energy offers some configurable parameters that need adjusting according to the specific utilization. Our goal in determining the parameter values is to preserve energy of both beacon nodes and user's mobile phone, while gaining a sufficient accuracy. Our choice of parameters for the experiment is shown in Table 1.

Broadcasting power is set to $-20dBm$ (instead of the default value of $-12dBm$). This is to extend beacons battery life expectancy to more than 3 years. With such power configuration, each beacon theoretically will be able to reach devices within a 3.5 meter range. This shorter coverage entails the deployment of more beacons per unit of space. In our experiment we used 12 beacons to cover 3 working offices, a SC, and HW, totally approximately $170 m^2$. Each beacon has the same $950ms$ broadcasting interval, while a mobile phone will listen for beacon messages for $2500ms$ and wait for next batch for $1500ms$. We then collect reference data in 10 observation points and an observation area for each HW and SC, shown as red dots in Figure 2.

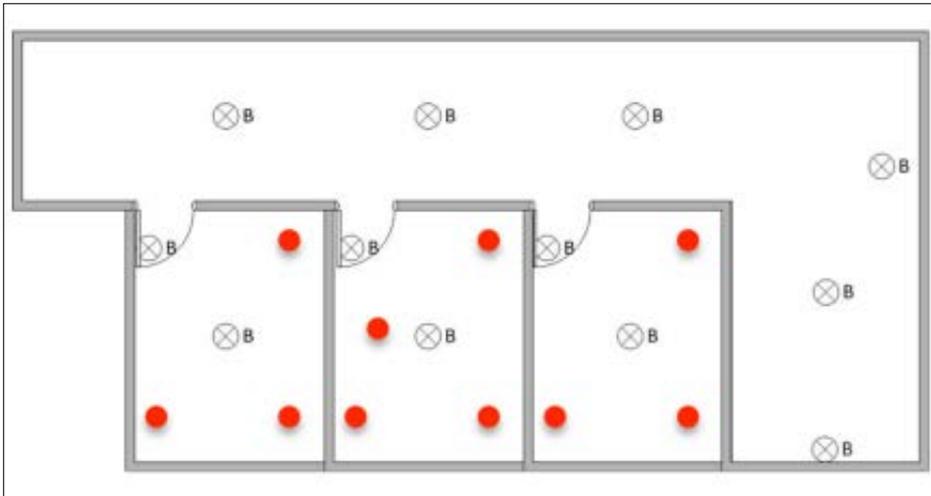


Fig. 2 Beacon deployment and observation points

To evaluate the techniques, we conduct 3 different movement scenes consisting of a total of 30 room transitions. Namely,

- HW - R1 - HW - SC - HW - R2 - HW - SC - HW - R3 - HW
- HW - R1 - HW - R3 - HW - SC - HW - R2 - HW - SC - HW
- HW - SC - HW - R1 - HW - R2 - HW - R3 - HW - SC - HW

Since each room has several observation points, as illustrated in Figure 2, the occupancy points of a room in one scene are different from those of the same room in other scenes.

For the purpose of classifier performance checking, we provide information on the actual location when measuring, that is, some form of ground truth against which we compare the measured values and their classification. The ground truth values are collected semi-manually by pushing a physical button of the phone's headset to trigger a confirmation dialog box of current position that the user has to confirm. Such value is stored together with the device's clock time. When a user switches from one room to another one, he is expected to press the button and select the new entered room, among the possible ones. Such real location value is used to evaluate the performance by comparing to classification outputs.

3.2 Metrics

The True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) are the basic ingredients of the evaluation metrics to quantify the performance of a classification method. These represent the number of actual (Positive) or non-occurred (Negative) events that are correctly or incorrectly recognized. Based on these, we have two metrics:

1. Accuracy

The accuracy metric is the number of correct predictions made, divided by the total number of predictions; that is,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

2. F-measure

The F-measure metric gives the harmonic mean of the number of correct estimated events divided by the total number of estimated events (a.k.a. precision) and the number of correct estimated events divided by the total number of real or relevant events (a.k.a. recall). Precision, Recall, and F-measure are defined in Equations (5), (6) and (7), respectively.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F - measure = \frac{(1 + \beta^2) * recall * precision}{\beta^2 * recall + precision} \quad (7)$$

4 Results and discussions

4.1 Results: Classification approach comparison

Several approaches can be used to map the beacon's RSS into a location. The first step we take is to compare three well-known classification techniques to identify the most promising one. These are Nearest Neighbor (NN), Nearest Beacon (NB), and one based on the cosine distance.

We consider the NB as a baseline as this is the simplest approach. It works by associating the location of a device with the strongest RSS signal among all the available signals. We also consider the Euclidean NN for the classification technique. NN makes use of the straight-line distance between two points (i.e. reference and query signals). Finally, the cosine distance works by translating RSS into elements of a vector space and measuring the angle among them considered as a degree of similarity between pair of vectors.

The comparison of these approaches using the test data of scene-2 is shown in Figure 3. The baseline is about 82% F-measure (i.e.: indicated by the strongest signal). This is higher than the Euclidean NN classification. The reason for the NB remaining stable in the graph is because the approach immediately sets the location based on the strongest signal. Hence, different window sizes do not affect the performance as no low pass filter is applied. One can see that the Euclidean NN and the cosine distance provide around 69% and 83% F-measure, respectively, without passing through a low pass filter. Such a filter could relatively enhance the performance of both approaches. The cosine distance undoubtedly outperforms the NN for all window sizes. It also exceeds the baseline for a meaningful interval, namely for time windows of size up to 25 units. There is unnoticeable benefit of considering window sizes larger than 25.

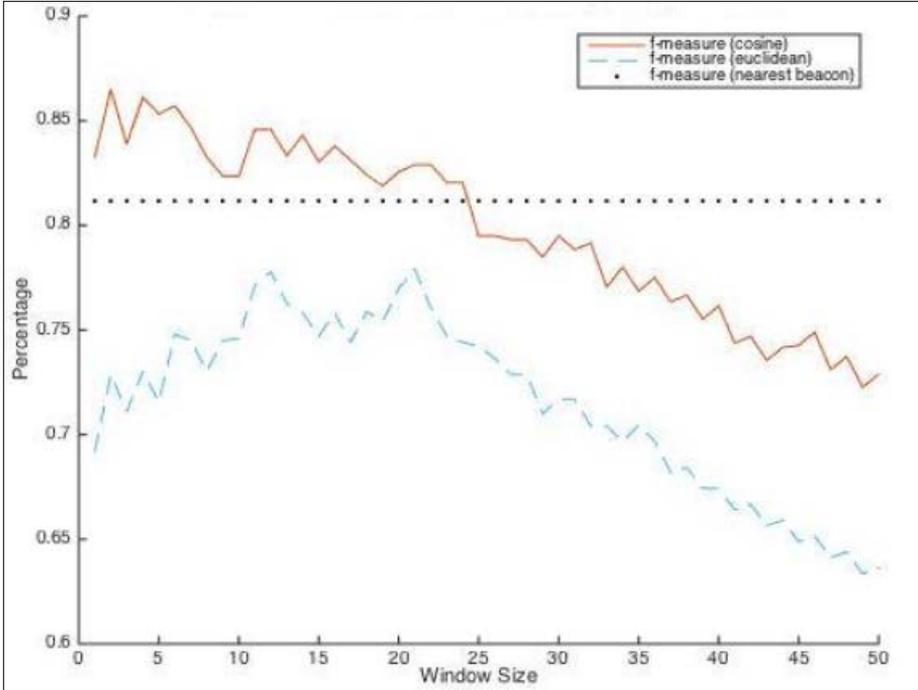


Fig. 3 Comparison of F-measure for increasing time windows

The average performance of the three scenes consisting of 30 room transitions is summarized in Table 2. We evaluate the approaches by considering either only the RSS or a combination of RSS and steps count. The NB approach, however, only takes the RSS into account. Interestingly, the steps count feature does not appear to bring an increase in the accuracy of classification. It even somewhat reduces the performance of both the Euclidean NN and cos-distance approaches. The Euclidean NN method is the least performing in terms of F-measure, with only a value of 68.85%. We take the F-measure into account since our dataset consists of unbalanced classes (the hallway, for example, is less occupied than other classes). Then the classification performance is only partially described by the accuracy metric.

Table 2 Performance of classification techniques

Method	Accuracy (%)	F-Measure (%)
Euclidean NN	81.83	68.85
Nearest Beacon	87.50	75.75
Cos-distance	89.90	80.41
Euclidean NN (+ steps count)	81.62	68.51
Cos-distance (+ steps count)	89.78	80.34

4.2 Results: Cosine distance

Given the comparison results, we focus on the cosine similarity on vectors as the classification approach.

We consider each observation point as a location reference. We have m -observation points with readings from n -beacons. So, we can define a $m \times n$ prototype matrix as:

	b_1	b_2	b_3	...	b_n
v_1	RSS_1^1	RSS_1^2	RSS_1^3	...	RSS_1^n
v_2	RSS_2^1	RSS_2^2	RSS_2^3	...	RSS_2^n
...
v_m	RSS_m^1	RSS_m^2	RSS_m^3	...	RSS_m^n

Each prototype vector v_i is computed as the median of all instances measured at an observation point. The median is taken to represent the observation point and the area it is in. From the practical point of view, we find that some beacons in the hallway (HW) do not contribute to these reference vectors. Therefore, there is an opportunity to reduce the dimensionality of the vector space by discarding unused beacons. To verify the validity of such reduction, we perform a classification experiment with and without these beacons taken into account in the vector space. The result is shown in Figure 4. Both accuracy and F-measures slightly decline for some window sizes, up to 13 units. For the rest, no different performance is noticeable, thus indicating that such space reduction does not hinder the quality of the classification in any significant way.

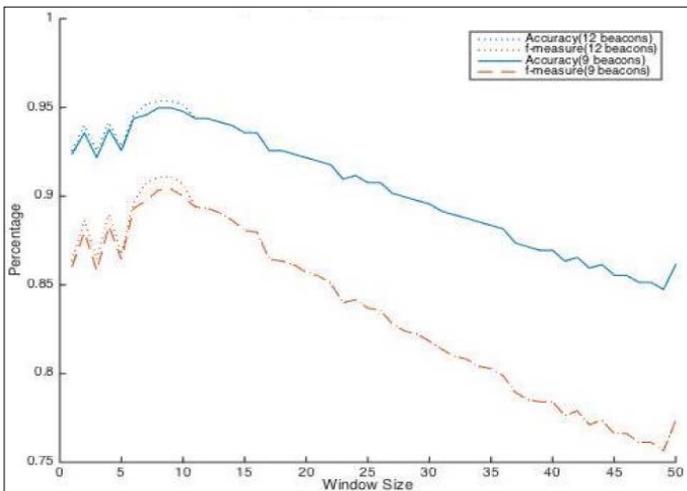


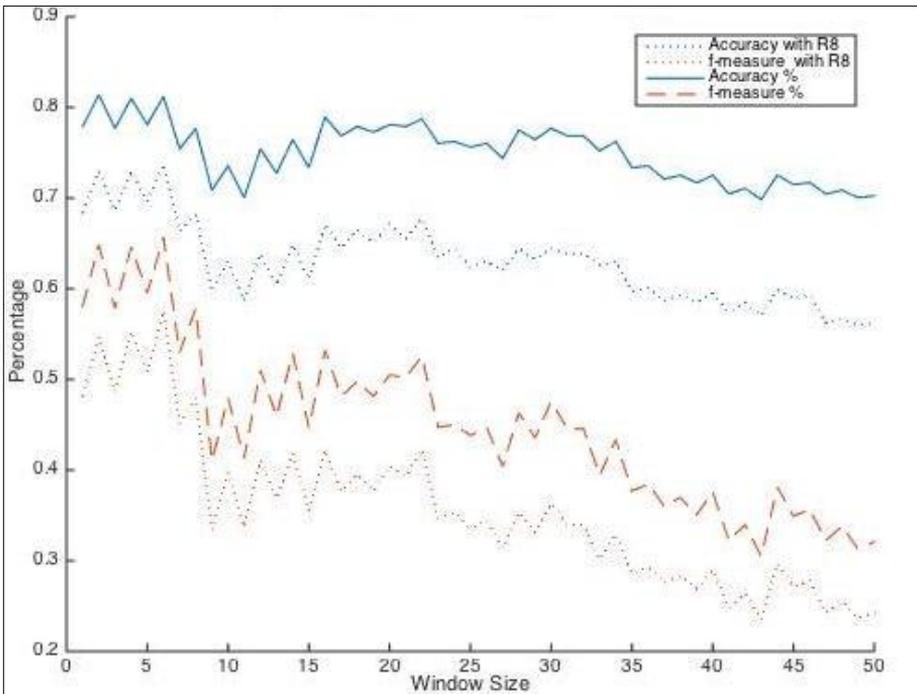
Fig. 4 Impact of discarding beacons in Hallway

Table 4 Classification with cosine distance approach

Scene	Accuracy (%)	F-Measure (%)
1	81.20	65.66
2	93.35	85.32
3	94.38	89.31
Average	89.64	80.10

4.3 Discussion

Our experiments showed that counting steps does not improve the room-level localization for an office space. There is a potential benefit to use steps-count feature for hierarchical classification, such as classifying binary states (i.e. in moving- or static-state) followed by more detail room classification (e.g. static inside room R2).

**Fig. 5** Impact of discarding prototype vector v_8

The experiment shows that the cosine distance with appropriate thresholding and time windowing offers very good classification results, going up to 89.78%. Furthermore, the contribution of the signals from areas which are rarely used is low and can even

be disregarded in the classification process. The reason is that there is a delay between the signal being transmitted from beacons and it being received by a mobile phone. Both period of signal propagation time and listener's sampling interval take part in such delay. Hence, in the case of the place where people rarely drop-by for longer periods (in Hallway, for example), the system is simply not gathering sufficient information from landmarks installed in that place. In other terms, the analysis also suggests where there should be more and where less density of beacons.

The experiment also shows that based on vector evaluation as expressed in Equation 3, we can differentiate between vectors that have similar direction and the others. Vectors v_1 and v_8 are two vectors which have a very similar direction to the vectors that represent other classes, but we cannot ignore both of them. We argue that v_1 is the vector that can't be removed as it represents the Hallway area which is interconnected with all other areas. Thus, it has a similar direction with the other vectors representing other classes/spaces. With respect to v_8 , this is intended to represent class R2 from a fixed observation point in the R2. It is expected to only have similar direction to other vectors representing class R2, but this is not the case. Hence, we decide this to be a removable vector.

5 Related work

Several efforts to obtain reliable location information through Radio Frequency (RF) technologies have been proposed. In [5], Liu *et. al.* estimate location of a tracking tag by comparing the signal strength received by the tag to the RSS of reference tags, both relative to active RFID readers. However, the readers price is expensive and the requirement of carrying additional tag make this inconvenient for our office scenario. Other approaches make use of existing Wi-Fi infrastructures. Some of them take advantage of the signal propagation model [6], of sniffing packets of wireless traffic from occupant devices [7], or by constructing radio map fingerprints [8]. Such approaches provide localization information, though suffer from being dependant on specific physical building construction, require major changes in the access points (e.g. for sniffing communication packets), and have the prerequisite of extensive radio map fingerprints collection, respectively. In other words, they are all hardly portable and scalable. Even though several studies propose methods for building fingerprint maps automatically, these usually utilize some additional resources (e.g.: accelerometer [9, 10]), making system performance really dependent on the step detection accuracy and consuming additional battery resources.

Bluetooth Low Energy standard is designed for marketing purposes. Though it has the potential to be used in personal localization, studies in this field have not been as intensive as Wi-Fi-based localization. In their study, Faragher *et. al.* [11] have shown potential improvements in BLE-based fingerprinting over Wi-Fi technology. Nevertheless, BLE configurations is set with high advertising rates and high transmission powers. This is certainly impractical in daily use due to power constraints, even though we have shown that BLE-based communication is more energy-efficient than Wi-Fi [12]. Room-level localization with BLE beacons has been experimented also by [13], utilizing RSS as feature for room classification. The claim is that 83% accuracy is achieved by implementing KNN and Decision tree techniques. The authors of [14] asserted that 97.22% accuracy could be achieved even only using the strongest beacon signals. However, in their experiment, they only considered two

rooms and did not take more adjacent rooms. Taking more adjacent rooms will become a more challenging case since beacon signals from neighbors can be discovered and would interfere in the localization.

They were also not actively moving in the Hallway while taking experiments, instead, they take data in static points in the Hallway. Moreover, these authors are not interested in the battery life of the beacons and accept frequent battery replacement [14] or even add wired power sources to their Arduino-based beacons [13]. In contrast, authors in [15] have experimented with a special purpose BLE on chip that exploits power harvester in supplying energy to the chip. Nevertheless, this system is intended to demonstrate the possibility of auto-powered indoor localization system that is almost impossible to be used daily due to its dependencies on specific hardwares.

In this study, we propose a scalable and portable room-level localization system, with bearing in mind the energy saving configuration such as increasing the density of battery-powered beacons deployment as a compensation for broadcasting power reduction. We arbitrarily (i.e. without any predefined rules) collect RSS as reference in rooms, similar to [10], but instead of matching this to step detection, we analyze the collected references based on cosine similarity approach. To the best of our knowledge, none of the earlier indoor localization systems have considered cosine similarity in choosing proper vectors and in classifying a room location. In our case, where a room is represented by multiple reference beacons (thus consisting many dimensional data), the properties of received signals are more appropriate to be evaluated in the cosine similarity rather than the Euclidean distance approach.

6 Conclusion

We introduced a room-level indoor localization system and its implementation. The system is based on the emerging BLE standard offering an overall affordable and scalable solution. We address the problem of location accuracy based on three lightweight approaches to classification. We provide an analysis of signal strength vectors in terms of vector component and vector direction in order to support the cosine distance approach, as experimental evidence indicates it as the most suited one.

According to our experimental results in an office building, the cosine distance approach provides greater result by 12% and 5% F-measure over the Euclidean nearest neighbor and the nearest beacon approaches, respectively. Overall, we achieve room-level localization with accuracy of up to 90%.

The approach presented here promises to be highly portable and potentially be brought to other buildings with varying indoor configurations and higher number of rooms. Next step we plan to undertake involves the automatic elimination of vectors that do not help effectiveness in terms of classifying a label. Collecting more instances will also give a benefit to investigate more general patterns and exploit probability-based approaches.

References

1. Georgievski, I., Degeler, V., Pagani, G. A., Nguyen, T. A., Lazovik, A., Aiello, M.: Optimizing energy costs for offices connected to the smart grid. *IEEE Trans. Smart Grid.* **3**(4), 2273-2285 (2012)
2. Nguyen, T. A., Aiello, M.: Energy intelligent buildings based on user activity: A survey. *Energy and buildings.* **56**, 244-257 (2013)

3. Setz, B., Nizamic, F., Lazovik, A., Aiello, M.: Power management of personal computers based on user behaviour. In: Proceedings of 5th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS). 1-8. IEEE (2016)
4. Apple Inc: Getting Started with iBeacon. <https://developer.apple.com/ibeacon/Getting-Started-with-iBeacon.pdf>. 1-11 (2014)
5. Ni, L. M., Liu, Y., Lau, Y. C., Patil, A. P.: LANDMARC: Indoor location sensing using active RFID. *Wireless networks*. **10**(6), 701-710 (2004)
6. Bose, A., Foh, C. H.: A practical path loss model for indoor WiFi positioning enhancement. In: Proceedings of 6th International Conference on Information, Communications & Signal Processing. 1-5. IEEE (2007)
7. Lu, X., Wen, H., Zou, H., Jiang, H., Xie, L., Trigoni, N.: Robust occupancy inference with commodity WiFi. In: Proceedings of Wireless and Mobile Computing, Networking and Communications (WiMob). 1-8. IEEE (2016)
8. Bahl, P., Padmanabhan, V. N.: RADAR: An in-building RF-based user location and tracking system. In: Proceedings of 19th Annual Joint Conference of the IEEE Computer and Communications Societies. **2**, 775-784. IEEE (2000)
9. Jiang, Y., Pan, X., Li, K., Lv, Q., Dick, R. P., Hannigan, M., Shang, L.: Ariel: Automatic wi-fi based room fingerprinting for indoor localization. In: Proceedings of the ACM Conference on Ubiquitous Computing. 441-450. ACM ((2012)
10. Yang, Z., Wu, C., Liu, Y.: Locating in fingerprint space: wireless indoor localization with little human intervention. In: Proceedings of the 18th annual international conference on Mobile computing and networking. 269-280. ACM (2012)
11. Faragher, R., Harle, R.: An analysis of the accuracy of bluetooth low energy for indoor positioning applications. In: Proceedings of the 27th International Technical Meeting of the Satellite Division of the Institute of Navigation. 201-210 (2014)
12. Putra, G. D., Pratama, A. R., Lazovik, A., Aiello, M.: Comparison of energy consumption in Wi-Fi and bluetooth communication in a Smart Building. In: Proceedings of Computing and Communication Workshop and Conference (CCWC). 1-6. IEEE (2017)
13. Conte, G., De Marchi, M., Nacci, A. A., Rana, V., Sciuto, D.: BlueSentinel: A first approach using iBeacon for an energy efficient occupancy detection system. In: Proceedings of BuildSys@ SenSys. 11-19 (2014)
14. Lin, X. Y., Ho, T. W., Fang, C. C., Yen, Z. S., Yang, B. J., Lai, F.: A mobile indoor positioning system based on iBeacon technology. In: Proceedings of 37th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC). 4970-4973. IEEE (2015)
15. Liu, Q., Pawelczak, P., Zuniga, M., Ijntema, W., Drif, A.: Wiploc: Perpetual indoor localization with rf wireless power transfer. Under submission