Machine Learning-based Real-Time Indoor Landmark Localization

Zhongliang Zhao, Jose Carrera, Joel Niklaus, and Torsten Braun

Institute of Computer Science, University of Bern, CH-3012 Bern, Switzerland,
zhao@inf.unibe.ch, carrera@inf.unibe.ch, joel.niklaus@students.unibe.ch,
braun@inf.unibe.ch

Abstract. Nowadays, smartphones can collect huge amounts of data from their surroundings with the help of highly accurate sensors. Since the combination of the Received Signal Strengths of surrounding access points and sensor data is assumed to be unique in some locations, it is possible to use this information to accurately predict smartphones' indoor locations. In this work, we apply machine learning methods to derive the correlation between smartphones' locations and the received Wi-Fi signal strength and sensor values. We have developed an Android application that is able to distinguish between rooms on a floor, and special landmarks within the detected room. Our real-world experiment results show that the Voting ensemble predictor outperforms individual machine learning algorithms and it achieves the best indoor landmark localization accuracy of 94% in office-like environments. This work provides a coarse-grained indoor room recognition and landmark localization within rooms, which can be envisioned as a basis for accurate indoor positioning.

Keywords: Machine learning, indoor localization, real-time landmark detection

1 Introduction

High localization accuracy within buildings would be very useful - in particular, large complex buildings like shopping malls, airports and hospitals would be well served by this feature. It would make orientation within these highly complicated structures much easier and would diminish the need for big floor maps scattered all around these buildings. However, walls, roofs, windows and doors of the buildings greatly reduce the GPS signals carried by radio waves, which leads to a severe accuracy loss of GPS inside buildings.

Different solutions already exist for indoor localization of mobile devices such as Pedestrian Dead Reckoning (PDR) and Wi-Fi fingerprinting based methods [1] [2]. In PDR the future location of a smartphone user is predicted based on the current location, and the movement information derived from the inertial sensor measurements. In Wi-Fi fingerprinting, the Received Signal Strength (RSS) values of several access points in range are collected and stored together with the

source: https://doi.org/10.24451/arbor.19708 | downloaded: 19.3.2024

coordinates of the location. A new set of RSS values is then compared with the stored fingerprints and the location of the closest match is returned.

In contrast to outdoors, building interiors normally have a large number of different Wi-Fi access points constantly emitting signals. By scanning the area around the device, we can measure the received signal strength of each of the nearby access points. Because there are typically so many of them, we presume that the list of all these values combined is unique at every distinct point in the building. Furthermore, we can strongly assume that these values are also constant over time as the access points are fixed in place and are constantly emitting signals of the same strength. Of course, there may be occasional changes, for instance if the network is remodeled, but we expect these changes to be infrequent.

In this way, we can collect lots of labeled location data of the building. However, because each data point may contain a very large number of Wi-Fi access point RSS measurements and magnetic field measurements, the data is very complex. Therefore, we propose using supervised Machine Learning (ML) methods to process this large amount of collected data. By training a classifier (supervised learning algorithm such as K-Nearest-Neighbor) on the collected labeled data, rules can be extracted. Feeding in the actual live data (RSS values, magnetic field values, illuminance level, etc.) of a moving user, the trained classifier can then predict the user's location on a coarse-grained level. We propose to apply machine learning methods, both individual predictors and ensemble predictors, to solve this task due to the large amount of features that are available in indoor environments, such as Wi-Fi RSS values, magnetic field values, and other sensor data. We expect that ensemble predictors can outperform the individual machine learning algorithms to discover patterns in the data, which can then be used to differentiate between different rooms and regions within the detected rooms.

The rest of the paper is organized as follows. In Section 2 we present some related work in indoor localization and landmark detection. Section 3 describes the used machine learning models, including the individual and ensemble ones, as well as the considered features to conduct the indoor landmark localization task. Section 4 presents implementation and experiment details. Section 5 discusses the performance results of our approach. Section 6 concludes the paper.

2 Related Work

Various machine learning-based approaches that use fingerprinting to estimate user indoor locations have been proposed. Machine learning-based indoor localization can be classified into generative or discriminative methods, which build the machine learning model using a joint probability or conditional probability respectively [1] [2]. K-Nearest-Neighbor (KNN) is the most basic and popular discriminative technique. Based on a similarity measure such as a distance function, the KNN algorithm determines the k closest matches in the signal space to the target. Then, the location of the target can be estimated by the average of the coordinates of the k neighbors [3]. Generative localization methods apply

statistical approaches, e.g., Hidden Markov Model [4], Bayesian Inference [5], Gaussian Processes [6], on the Wi-Fi fingerprint database. Thus, the accuracy can obviously be improved by adding more measurements. In [6] for instance, Gaussian Processes are used to estimate the signal propagation model through an indoor environment. There is a limited number of works that have focused in reducing off-line efforts in learning-based approaches for indoor localization [7] [8] [9]. These approaches reduce the off-line effort by reducing either the number of samples collected at each survey point or the number of survey points or both of them. Then, a generative model is applied to reinforce the sample collection data. In [7] for instance, a linear interpolation method is used. In [8], a Bayesian model is applied. In [9], authors propose a propagation method to generate data from collected samples. In [2], authors combine characteristics of generative and discriminative models in a hybrid model. Although this hybrid model reduces offline efforts, it still relies on a number of samples collected from fixed survey points (i.e., labeled samples) along the environment. Therefore, to maintain high accuracy, the number of survey points shall be increased in larger environments. Thus, collecting samples from numerous survey points will become a demanding process, which makes the system unsuitable to large environments. In [10], authors validated the performance of different individual machine learning approaches for indoor positioning systems. However, they rather compare the results without any deep analysis of the performance difference. Moreover, they did not discuss how ensemble learning approaches could be used to enhance system performance.

In this work we present and analyze the performance of different individual predictors as well as ensemble predictors for the indoor landmark localization problem. This work could also be used as a basis of indoor tracking systems to firstly locate the target with a coarse-grained accuracy using indoor landmark localization, which then triggers the real-time localization algorithm to locate the object around the detected landmarks. The located landmark can also be used to correct the localization failures like the kidnapped robot problem [11].

3 Machine Learning-based Indoor Landmark Localization

An indoor landmark is defined as a small area within a room. The aim of the indoor landmark localization system presented in this work is to improve the accuracy of indoor landmark recognition using machine learning approaches. We do this by excluding all the possible locations of the user within the room if the system predicts the others by using landmarks. Thus, when a landmark has been recognized, the indoor positioning system can use the identified coarse-grained locations to optimize the positioning accuracy, such as revising positioning errors.

3.1 Algorithms

In this section, we shortly describe the machine learning algorithms that are used in this work to perform the room landmark localizations.

Naive Bayes (NB) classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features.

K-Nearest Neighbors (KNN) is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space.

Support Vector Machine (SVM) is a supervised learning model with associated learning algorithms that analyze data used for classification and regression. Given a set of training examples, each is marked as belonging to one or the other category. An SVM training algorithm builds a model that assigns new data measurements to one category or the other, making it a non-probabilistic binary linear classifier.

Multilayer Perceptron (MLP) is a class of feed-forward artificial neural network. An MLP consists of at least three layers of nodes. Except for the input nodes, each node is a neuron that uses a nonlinear activation function.

Voting is one of the simplest ensemble predictors. It combines the predictions from multiple individual machine learning algorithms. It works by first creating two or more standalone prediction models from the training dataset. A Voting classifier can then be used to wrap the models and average the predictions of the sub-models when asked to make predictions for new data.

3.2 Features

In a machine learning-based classification task, the attributes of the classes are denoted as features. Each feature is describing an aspect of the classes. In our case features are our measurements, for instance an RSS value. To deliver good machine learning prediction accuracy it is very important to select the right attributes/features and to also modify certain features or even create new features out of existing features.

Wi-Fi RSS values provide the core data as they contribute the most to the performance of the ML methods. The smartphone scans the surrounding Wi-Fi access points, obtains and registers the RSS values of each access point. Wi-Fi RSS values depend on the distance between the smartphone and the Wi-Fi access points. Normally, the Wi-Fi RSS values in our datasets were between -20 dBm and -90 dBm.

Magnetic Field (MF) The device's sensors measure the magnetic field in the device's coordinate system. As the user walks around, the orientation of the device may change all the time. Therefore, we have to collect all possible values from every orientation in every point in the training phase. This would result in a huge amount of data and the training performance would be inaccurate.

Light sensors might also be helpful to identify rooms. For instance, a room facing a window will clearly be brighter than one surrounded by walls only. As shown in Section 5 this does improve the prediction accuracy. However, these assumptions are not stable, as the illuminance level might change over time. Therefore, it is better to work with light differences instead of absolute values.

4 Implementation and Experiments

This section explains how the indoor room landmarks are defined and presents details about how to make the room landmark localization using ML algorithms.

4.1 Room and Landmark Recognition

A room landmark is defined as a small area within a room, and room landmark fingerprint database inlcudes the Wi-FI RSS, MF measurements, and illuminance level data measured within that small area. In the room recognition phase we distinguish several rooms on the same floor. In the landmark recognition phase we distinguish several landmarks inside the detected room. Therefore, we define two landmarks in a small room with size of 3x3 meters, and four landmarks in a normal office-sized room (5x5 meters). In a big room (7x7 meters) we define five landmarks, one in each corner and one in the center, as shown in Figure 1.

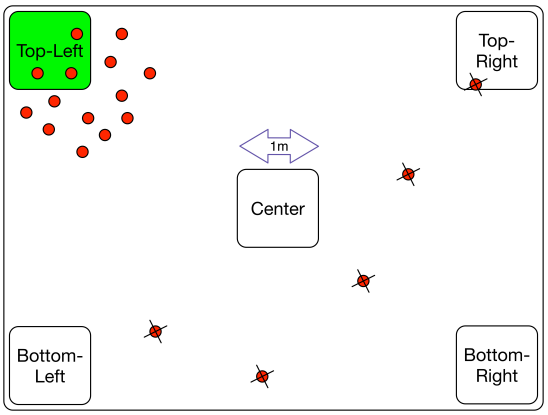


Fig. 1. Five landmarks and the collection of red points (location examples) predicted by the indoor landmark localization system.

4.2 System Architecture

Figure 2 shows the data flow and the different components of our developed Android app. Sensor and Wi-Fi RSS values are measured by the smartphone and received by the app. We then perform the data training process offline in a PC to pass the collected data to the Model Training component, which applies different machine learning algorithms to build the models. The trained models are then optimized and transfered to the app on the smartphone for online experiments.

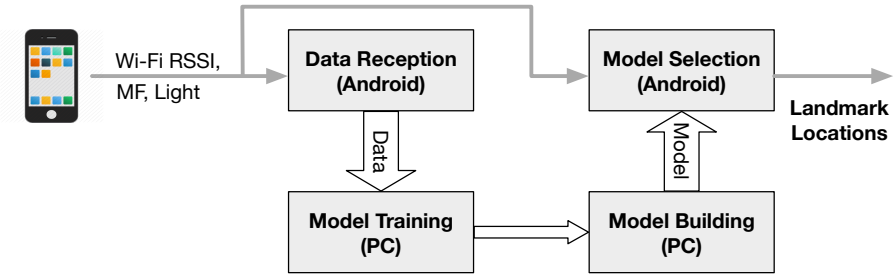


Fig. 2. The architecture of the implemented Android app.

Considering that the landmark detection accuracy can be influenced by some environmental parameters, we conduct some experiments to determine how parameters such as AP position or number of APs influence the accuracy of the Wi-Fi-based fingerprinting landmark detection approach. Additionally, we perform experiments to show how the accuracy is improved by considering additional features such as magnetic field (MF) values and light illuminance level readings. As shown in Figure 3, we define 9 wall separated areas in our experiment environment. Hereafter, we refer to these areas as rooms. In our experiments, we do not need to know the locations of the APs, while only the fingerprints of Wi-Fi RSSI, MF readings, and illuminance level readings are recorded during the data collection process.

Parameters of learning-based algorithms are optimized from training data. Additionally, certain algorithms also have parameters that are not optimized during the training process. These parameters are called hyperparameters, which have significant impact on the performance of the learning-based algorithm. Therefore, we use a nested cross validation technique to adjust them. The nested cross validation technique defines an inner and outer cross validation. The inner cross validation is intended to select the model with optimized hyperparameters, whereas outer cross validation is used to obtain an estimation of the generalization error. Ten-fold cross validation was applied on both inner and outer cross validation. The classifiers were optimized over a set of hyperparameters. We optimized the global blend percentage ratio hyperparameter for KNN, kernel type function for SVM, number of hidden layers and neurons per layer for MLP. Based

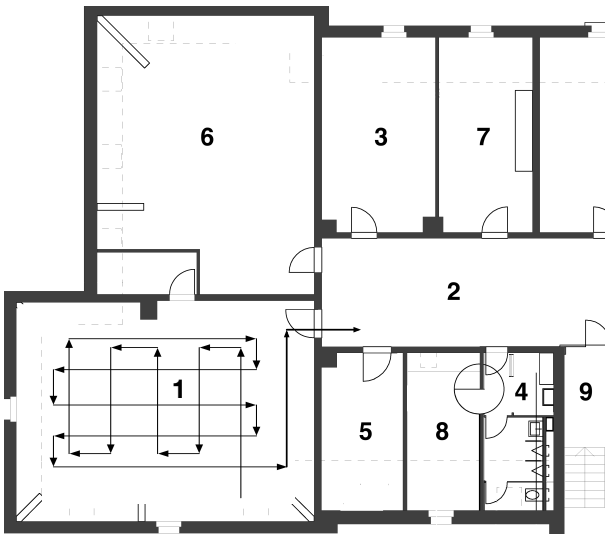


Fig. 3. Experiment scenario and data collection path.

on the parameter optimization process, we established the optimal hyperparameter values for the classifiers as follows: global blend percent ratio of 30% for KNN, single order polynomial kernel, $c = 1$, $\gamma = 0.0$ for SVM, and single hidden layer with 10 neurons for MLP.

4.3 Datasets

To test the room landmark detection performance, we performed experiments on the third floor of the Computer Science building of the University of Bern, as shown in Fig. 3. During the experiments, we collected 14569 data points in total, 3061 data points were collected from the biggest room (1) and 514 data points were from the smallest room (4). Collecting the training dataset takes around 50 minutes. With the collected data, we build models with different data: the first one builds the fingerprint using only collected Wi-Fi RSSI data, the second one using Wi-Fi RSS together with MF readings, and the third one with Wi-Fi RSS, MF readings, and illuminance level readings.

As described before, to build the landmark fingerprint database, we ask a person to walk randomly around each room holding the phone in his/her hand. Landmark fingerprint database entries must be collected equally distributed along the whole area in each room. The data collection rate is only constrained by computational capabilities of the Wi-Fi sensor of the smartphone. Thus, in our experiments every data measurement was collected at a rate of 3 entries/second. Because our approach does not need to predefine any survey point, the time needed to build the landmark fingerprint database is proportional to the number of collected instances multiplied by the instance collection rate.

5 Results

5.1 Indoor Landmark Localization Accuracy

This section discusses the accuracy of the landmark detection model when different classifiers and features are used. When comparing their performance, it is impossible to define a single metric that provides a fair comparison in all possible applications. We focus on the metrics of prediction accuracy, which refers to the percentages of correct room recognition and landmark localization within the detected room. Landmark definition is described in Section 4.1.

At first we use only Wi-Fi RSS values as inputs to machine learning algorithms. Figure 4 shows the classification accuracy of different predictors when different numbers of Wi-Fi RSS values are used. As we can see, starting from 5 RSS values, more RSS inputs increase the prediction accuracy for most of the predictors. Nevertheless, after 7 Wi-Fi RSS values are used, the improvement of adding more RSS values is almost negligible in all tested classifiers, and some of the predictors even got reduced accuracy when additional RSS values are considered. We think that the signal interferences may be the reason for the worse performance when more than 7 Wi-Fi RSS values are utilized. Therefore, we take 7 Wi-Fi RSS as the default configuration for the following experiments.

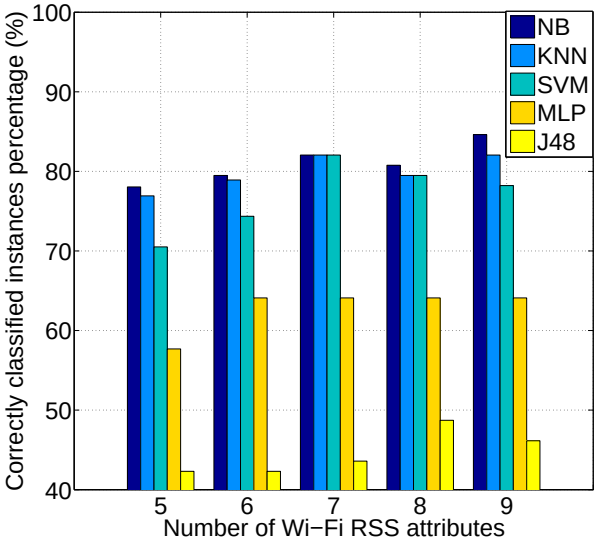


Fig. 4. Landmark prediction performance with different numbers of Wi-Fi RSS values.

Next, we compare the classification accuracy when using only Wi-Fi RSS, Wi-Fi RSS plus MF, and Wi-Fi RSS plus MF and illuminance levels. Figure 5 shows the performance evaluation of the selected classifiers obtained with different feature combinations. The best performance is reached by the Naive Bayes classifier, which achieves 90.13% of instances correctly classified if the fingerprint

is composed by Wi-Fi RSS, MF readings, and illuminance levels. By using Wi-Fi RSS, MF readings, and illuminance levels in the room landmark recognition, the accuracy is improved in all tested classifiers.

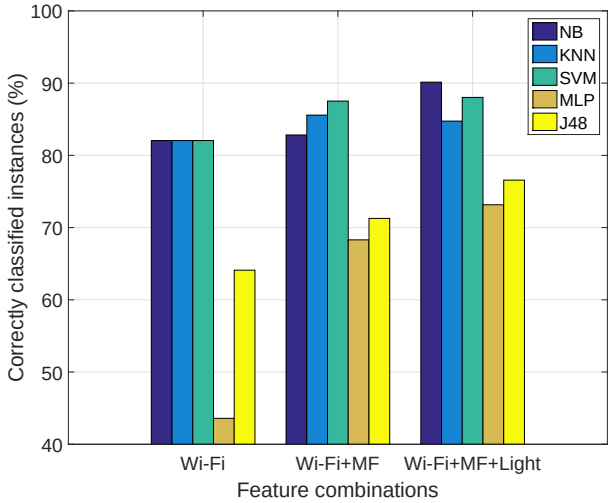


Fig. 5. Landmark prediction performance when using different features.

As mentioned before, hyperparameters have significant impacts on the performance of the learning-based algorithm. Figure 6 shows the performance of the selected classifiers with the hyperparameters optimized. The classifiers are all fed with Wi-Fi RSS plus MF and illuminance levels. As we can see, compared to results in Figure 5, all the classifiers have improved performance, and MLP even reaches an accuracy of 92.08%. We also include the results of Voting, which combines the prediction results of MLP, Naive Bayes, KNN, and SVM using majority vote. It shows that Voting can reach an accuracy of 94.04%.

5.2 Result Analysis

In indoor environments, Wi-Fi RSSI and MF measurement vary dependent on locations. However, these values will remain similar on nearby positions. For example, on locations close to landmark borders, high similarities will be observed on the RSS values. These similarities could lead to misclassification problems. From Figure 5 we can see that KNN and SVM outperform others in terms of accuracy when Wi-Fi RSS and MF readings are used. This is because KNN is an instance-based learning algorithm, which uses entropy as a distance measure to determine how similar two instances are. Thus, this method is more sensitive to slight variations upon the instance as unity. J48 builds the classification model by parsing the entropy of information at attribute level. It means that J48 measures entropy in the attribute domain to decide which attribute goes

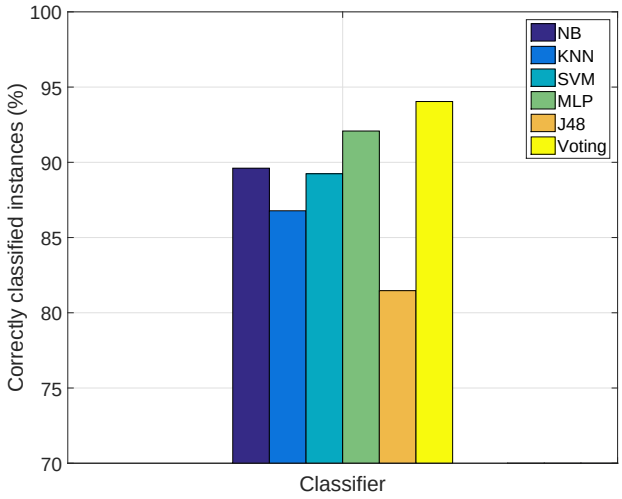


Fig. 6. Landmark prediction performance of individual predictors with optimized hyperparameters and Voting ensemble predictor.

into a decision node. Therefore, the classification model is prone to misclassification in the landmark detection problem. When the illuminance level is included as input feature to predictors, Naive Bayes outperforms others. This is because the feature of illuminance level is completely independent from other radio signal measurements, which fits with Naive Bayes' strong assumptions about the independence of each input variable.

To further explain how the Voting predictor improves the performance of individual predictors, we show the confusion matrix of room recognition using MLP, SVM, Naive Bayes (NB), and Voting in Tables 1 - 3. We can observe that room 2 is correctly identified 527 times by MLP, 632 times by SVM and 393 times by NB. As a consequence, SVM seems to be better in predicting room 2 as compared to other two predictors. Furthermore, NB does not seem to have less misclassification of class b compared to other two predictors. Analyzing the results from the above-mentioned tables, MLP has misclassified class b 138 times, NB 272 times, and SVM only 33 times. From the confusion matrix of Voting, as shown in Table 4, we can see that the Voting ensemble predictor adopts behaviors of different individual predictors. For instance, it adopts the good behavior of MLP and Naive Bayes, which leads to a much better prediction accuracy for room 2. This can be observed from the only two misclassifications of room 2 as room 9 as shown in Table 4. Unfortunately, it still has problems in some classifications. For instance, there are 116 misclassification of room 8 as room 4, which is probably due to a higher weight assigned to MLP instead of SVM. In general, it can be observed that the Voting ensemble predictor improves the accuracy, while there are still difficulties to distinguish small rooms that are next to each other, as room 8 and 4 depicted in Figure 3.

Table 1. The confusion matrix for MLP with optimized hyperparameters.

a	b	c	d	e	f	g	h	i	<-- classified as
2164	0	0	0	0	0	0	0	0	a = 1
0	527	0	0	0	0	0	0	138	b = 2
0	25	865	0	0	0	0	0	0	c = 3
0	9	0	154	0	0	0	128	0	d = 4
0	43	0	0	642	0	0	0	0	e = 5
0	25	0	0	44	1249	0	0	0	f = 6
0	6	0	0	0	0	1064	0	1	g = 7
0	0	0	149	0	0	0	498	0	h = 8
0	0	0	51	0	0	0	64	778	i = 9

Table 2. The confusion matrix for SVM with optimized hyperparameters.

a	b	c	d	e	f	g	h	i	<-- classified as
2164	0	0	0	0	0	0	0	0	a = 1
10	632	0	0	0	0	0	0	23	b = 2
0	55	799	0	0	36	0	0	0	c = 3
0	147	0	144	0	0	0	0	0	d = 4
0	43	0	0	642	0	0	0	0	e = 5
204	69	32	0	0	1013	0	0	0	f = 6
0	126	0	0	0	0	945	0	0	g = 7
0	0	0	119	0	0	0	528	0	h = 8
0	0	0	57	0	0	0	7	829	i = 9

Table 3. The confusion matrix for Naive Bayes with optimized hyperparameters.

a	b	c	d	e	f	g	h	i	<-- classified as
2149	0	0	0	15	0	0	0	0	a = 1
0	393	9	153	0	0	0	110	0	b = 2
0	0	876	0	0	14	0	0	0	c = 3
0	130	0	135	0	0	0	26	0	d = 4
1	0	44	0	628	0	0	12	0	e = 5
0	0	101	0	0	1217	0	0	0	f = 6
0	59	12	0	0	0	973	0	27	g = 7
0	0	0	119	0	0	0	528	0	h = 8
0	0	0	0	0	0	0	64	829	i = 9

Table 4. The confusion matrix for the Voting ensemble classifier.

a	b	c	d	e	f	g	h	i	<-- classified as
2164	0	0	0	0	0	0	0	0	a = 1
0	663	0	0	0	0	0	0	2	b = 2
0	25	865	0	0	0	0	0	0	c = 3
0	32	0	143	0	0	0	116	0	d = 4
0	43	0	0	642	0	0	0	0	e = 5
0	69	0	0	0	1249	0	0	0	f = 6
0	43	0	0	0	0	1027	0	1	g = 7
0	0	0	119	0	0	0	528	0	h = 8
0	0	0	0	0	0	0	64	829	i = 9

6 Conclusions and Future Work

This work analyzes the performance of 5 common individual predictors and 1 ensemble predictor in indoor landmark localization to distinguish rooms on a floor and special landmarks using machine learning methods. We have validated the performance of the system using different smartphone sensor measurements, such as Wi-Fi RSS, MF readings, and illuminance levels. Evaluation results show that the Voting ensemble predictor achieves the best indoor landmark localization accuracy of 94%. In the future, we will further optimize the hyperparameter cross-validation procedure and integrate this work with an indoor tracking system to firstly locate the target with a coarse-grained accuracy, which then triggers the tracking algorithm to track the object around the located landmarks.

7 Acknowledgements

This work was supported by the Swiss National Science Foundation #154458.

References

1. S. He and S. Chan *Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons.* IEEE COMMUNICATIONS SURVEYS & TUTORIALS, 2016.
2. R. W. Ouyang and A. K. S. Wong and C. T. Lea and M. Chiang *Indoor Location Estimation with Reduced Calibration Exploiting Unlabeled Data via Hybrid Generative/Discriminative Learning.* IEEE Transactions on Mobile Computing, 2012.
3. B. Lakmali and W. Wijesinghe and K. de SIva and K. Liyanagama and S. Dias *Design, Implementation & Testing of Positioning Techniques in Mobile Networks.* The 3rd International Conference on Information and Automation for Sustainability.
4. Z. Ghahramani *An Introduction to Hidden Markov Models and Bayesian Networks.* International Journal of Pattern Recognition and Artificial Intelligence, 2001.
5. O. Bousquet and U. von Luxburg and G. Ratsch *Bayesian Inference: An Introduction to Principles and Practice in Machine Learning.* Fresenius Environmental Bulletin 20(5), 2004.
6. B. Ferris and D. Hahnel and D. Fox *Gaussian Processes for Signal Strength-based Location Estimation.* Procedures of Robotics Science and Systems, 2006.
7. X. Chai and Q. Yang, *Reducing the calibration Effort for Probabilistic Indoor Location Estimation.* IEEE Tran. Mobile Computing, vol 6, no 6, pp. 649-662, 2016.
8. D. Madigan and E. Einahrawy and R. Martin and W. Ju and P. krishnan and A. Krishnakumar, *Bayesian Indoor Positioning Systems.* IEEE. INFOCOM, vol 2, pp. 1217-1227, 2005.
9. S. Liu and H. Luo and S. Zou, *A Low-Cost and Accurate Indoor Localization Algorithm Using Label Propagation based Semi Supervised Learning.* Fifth International Conference Mobile Ad-Hoc and Sensor Networks, pp. 108-111, 2009.
10. D. Mascharka and E. Manley, *LIPS: Learning Based Indoor Positioning System using mobile phone-based sensors* 2016 13th IEEE Annual Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, 2016, pp. 968-971.
11. J. Carrera and Z. Zhao and T. Braun and Z. Li and A. Neto, *A real-time robust indoor tracking system in smartphones* Elsevier Journal of Computer Communications. DOI: http://dx.doi.org/10.1016/j.comcom.2017.09.004