Abstract—An accurate room localization system is a powerful tool for providing location-based services. Considering that people spend most of their time indoors, indoor localization systems are becoming increasingly important in designing smart environments. In this work, we propose an efficient ensemble learning method to provide room level localization in smart buildings. Our proposed localization method achieves high room-level localization accuracy by combining Hidden Markov Models with simple discriminative learning methods. The localization algorithms are designed for a terminal-based system, which consists of commercial smartphones and Wi-Fi access points. We conduct experimental studies to evaluate our system in an office-like indoor environment. Experiment results show that our system can overcome traditional individual machine learning and ensemble learning approaches.

I. INTRODUCTION

Due to the current importance of context aware services and the growing ubiquitousness of the Internet of Things (IoT), indoor localization has become an interesting research topic. Moreover, with the increase of smartphone devices, indoor localization applications for smartphones have attracted attention. Often, indoor context aware services require higher localization accuracy than outdoor services. Additionally, the algorithmic complexity of a localization application is constrained by the limited computation and power resources on the smartphone. Thus, indoor localization is still considered as an open challenging problem. With the development of smartphones and the availability of more embedded sensors on mobile devices, several indoor localization methods (e.g., Wi-Fi or magnetic field-based fingerprinting, Wi-Fi or Bluetooth-based ranging, etc.) have been proposed. However, due to the widely extended availability of Wi-Fi signals in indoor environments, Wi-Fi radio-based localization has attracted most attention [4]. Wi-Fi received signal strength indicator (RSSI) is the most used parameter for indoor localization. It can be used both in range-based or fingerprinting-based approaches.

Indoor environments provide many different ubiquitous radio signals, such as Wi-Fi, Bluetooth, magnetic field, sound, light, etc. The earth magnetic field (MF) has distortions over space due to the presence of ferromagnetic materials. These MF distortion patterns can be also used to identify indoor locations [1]. Thereby, MF and Wi-Fi observations can be used as radio fingerprints to detect unique locations in indoor environments.

Fingerprinting-based indoor localization systems usually consist of two phases: training phase (off-line) and localization phase (on-line). In the off-line phase, the fingerprint database is built by collecting various types of radio signals in the target indoor environments. In the on-line phase, the observed fingerprint at an unknown location is compared with the stored fingerprints in the fingerprint database to determine the closest match (i.e., prediction). In the on-line phase, any single discriminative learning model can be applied. However, ensemble learning models usually allow the production of better predictive performance compared to single models. Fingerprinting-based methods that build the classification model exclusively based on beforehand observed data (i.e., fingerprint database) are named discriminative learning methods.

We present a novel room-level localization approach by fusing Wi-Fi RSSI, MF and coarse-grained floor plan information in an ensemble discriminative learning model. To achieve high and stable performance, we apply Hidden Markov Models (HMM) and discriminative learning models to integrate Wi-Fi, MF readings and information about transitions between rooms to achieve room level localization. Our approach requires only information about the physical distribution of rooms in the target area. Therefore, a precise floor plan is not needed. Thus, unlike traditional fingerprinting localization approaches, we include room transition information in the localization process. Figure 1 shows an overview of our proposed approach. The indoor localization sys-
Fig. 1: Indoor Localization System Architecture. Integration of HMM with individual learning models.

system tracks room level location of a person holding a smartphone in real-time. It works as a basis for indoor location-based services for IoT. The main contributions of this work are summarized as follows.

- We propose a novel room level localization method based on an enhanced learning model. Our ensemble learning model fuses Wi-Fi, MF and room transition information to achieve high room level localization accuracy.
- We combine a set of individual machine learning methods in an ensemble learning model. Our approach integrates HMM with discriminative learning techniques to achieve high prediction performance. Unlike traditional Wi-Fi fingerprint-based approaches using individual machine learning methods, we include room transition information in the localization process.
- We implement an efficient terminal-based indoor localization system for smartphones that is able accurately track in real-time room level location of a person holding a smartphone. Since our approach provides room level localization, the fingerprint database is built by taking Wi-Fi and MF measurements while walking randomly through the environment, which requires only room-labeled samples that can be collected in a very short time period. Thus, the off-line phase becomes a simple process.
- We perform a set of experiments to validate the performance of our localization method. We evaluate our system in a complex environment along three different moving paths.

The rest of the paper is organized as follows. In Section II we present some related work. The ensemble learning model for indoor localization is reviewed in Section III. Section IV presents the implementation of the terminal-based system. Section V presents the performance evaluation results of our approach. Section VI concludes the paper.

II. RELATED WORK

Pedestrian Dead Reckoning (PDR) has attracted research interest due to the fast development of Inertial Measurement Units (IMUs) in modern smartphones. IMUs can be leveraged to implement pedestrian movement detection such as heading orientation, stride recognition, stride length estimation [10]. Thus, by using IMUs, PDR systems estimate the new location based on the previously determined location. For instance in [3], [9], the heading orientation is determined based on gyroscope measurements, whereas the displacement is estimated from accelerometer readings. In [3], a Heuristic Drift Elimination (HDE) is intended to deal with accumulated errors. However, HDE requires specialized sensors deployed on the foot of the pedestrian. In [14], authors use readings of the gyroscope to recognize physical turns of the pedestrian, whereas the stride length is determined by readings of the accelerometer. PDR systems measure position changes rather than the absolute position, which results in an accumulation of sensor errors over time. Therefore, these systems must consider additional information to deal with this type of errors.

Radio signals are often used to provide indoor localization. Several parameters of radio signals can be leveraged to locate targets. In [24] for instance, the authors propose to use received signal strength indicator (RSSI), whereas in [12] time information related to radio propagation is applied. Radio-based indoor localization can be classified as range-based and range-free methods. Range is defined as the propagation distance from the target to Anchor Nodes (AN). The first stage in range-based localization methods is to calculate the propagation distances, which is called ranging. Then, different positioning algorithms can be used to estimate the absolute locations of the targets, such as trilateration and multilateration [11]. However, unlike outdoor localization, trilateration does not work well in indoor environments because of the presence of obstacles and room partitions [8]. In range-free localization methods, fingerprinting [2] is very often used because of its robustness to multipath propagation. However, it is very time consuming to build up a radio map, which is required to locate the targets in fingerprinting. In [18], a Wi-Fi-based ensemble learning model is proposed. Authors proposed to provide localization by developing a room-based weighted method in an ensemble learning technique. The room detection method relies on a simple average of the coordinates output by a k-NN estimator. Therefore, this process is prone to errors. Typically, ranging requires much less labor efforts than building radio maps for fingerprinting approaches.

Applications of HMM can be considered to provide indoor localization. In [21], authors employed HMM and radio propagation models to reduce the calibration efforts. The system utilizes a discrete probability distri-
bution to derive probable positions. Then, the position is estimated from the most probable estimated positions. In [17], authors include motion measurements (e.g., heading orientation) in the proposed HMM. Thus, the reported accuracy is improved compared to [21]. In [13], authors propose to fuse IMU measurements with wireless signal readings. Then, the candidate position is derived based on the pedestrian’s motion pattern and the most probable wireless signals at that position. Despite authors report good accuracy, the method to determine the transition probabilities is not explained. Moreover, the applicability of the solution is restricted to the accuracy of the pedestrian motion detection method. In [23], authors propose to fuse a RSSI pattern recognition method and HMM to provide indoor localization. The pattern recognition method uses RSSI variation instead of raw RSSI. Then, the transition probabilities of the HMM are derived from the pedestrian trajectories and the pattern recognition method. The pattern recognition method relies on a beforehand built radio map database. Thus, some reference locations are defined through the indoor environment in an off-line phase to collect reference samples. Such collection process could take several hours or days for small or big areas, which is very labor expensive and time consuming.

III. HMM-Discriminative Ensemble Learning Method

Ensemble learning methods intend to improve prediction results by combining several individual machine learning techniques into one predictive model. Thus, ensemble approaches usually achieve better predictive performance compared to single models. We propose an ensemble learning model by combining HMM with discriminative machine learning methods to provide accurate room detection for smartphone users. Figure 2 shows the probabilistic parameters of the proposed HMM. It is worth to notice that this model is suitable for any zone detection (i.e., any subarea in the area of interest). Hereafter, we will refer to room as zone. The proposed HMM is based on the concept of Markov localization [6], which can be described as estimating the state of HMM with controllable state transitions. For the localization problem, we refer to zones as states of the HMM. Thus, our HMM is specified by the following components:

- A set of \( n \) states \( Z = \{z_1, z_2, ..., z_n\} \), where \( z_i \) is the identifier value of the zone \( i \). Thus, the discrete random variable \( x_t \in Z \) represents the hidden state at time \( t \).
- A transition probability matrix \( A \),

\[
A = \begin{pmatrix}
    a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\
    a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{n,1} & a_{n,2} & \cdots & a_{n,n}
\end{pmatrix},
\]

where \( a_{ij} \) represents the probability of moving from zone \( z_i \) to zone \( z_j \).
- A set of observations \( O \),

\[
O = \{(o_1, o_2, ..., o_m)_1, ..., (o_1, o_2, ..., o_m)_r\},
\]

where \( o_i \) is the zone prediction result of the \( i \)-th individual machine learning method. Therefore, \( O \) is a set of \( P_m^r \) permutations with repetitions allowed, where \( r \) is the number of zones and \( m \) is the number of individual machine learning methods used to build the ensemble predictive model. Conditional independence must be assured among the individual machine learning methods. Thus, the random variable \( y_t \in O \) represents the observations at time \( t \).
- An emission probability matrix \( B \) of observation likelihoods, which is also called emission probabilities,

\[
B = \begin{pmatrix}
    b_{1,1} & b_{1,2} & \cdots & b_{1,r} \\
    b_{2,1} & b_{2,2} & \cdots & b_{2,r} \\
    \vdots & \vdots & \ddots & \vdots \\
    b_{n,1} & b_{n,2} & \cdots & b_{n,r}
\end{pmatrix},
\]

where \( b_{i,j} \) represents the probability of an observation \( (o_1, o_2, ..., o_m)_j \) being generated at zone \( z_i \).
- An initial probability distribution \( \pi = \pi_1, \pi_2, ..., \pi_n \) over zones.

The individual learning methods only rely on the latest observed fingerprint of Wi-Fi RSS and MF readings for localization, which may produce incorrect prediction results. However, the HMM can be used to integrate zone transition information and the current observed information (e.g., current observed fingerprint) to improve prediction results. Hereafter, we will refer to our HMM-discriminative learning model as HMM-d model. In any model with hidden variables (e.g., HMM), the task of determining the sequence of variables (e.g.,
independent of each other, it is reasonable to assume that individual machine learning methods are different and result from the

\[ P_n \]

where \( P_n \) and \( F N_n \) are the true positive and false negative rate of the \( n \)-th individual discriminative learning method.

**IV. IMPLEMENTATION**

We have implemented a terminal-based system for accurate indoor localization. The system comprises two main components: a mobile target and many Wi-Fi Anchor Nodes (ANs). The proposed localization algorithms are running on the mobile target. Figure 1 presents the overview of the system. ANs are some commercial Wi-Fi access points deployed at known or unknown locations along the area of interest. To provide the maximum coverage inside the area of interest, the locations of ANs are defined by the boundary corners and the boundary itself. We have adopted D-Link D-635 and D-Link DAP-2553 devices as Wi-Fi ANs in this work. The beacon period is configured to 100\( ms \) for ANs. The mobile target can be any commercial Android smartphone, which supports Wi-Fi RSSI readings and magnetic field sensor readings. We have deployed the localization algorithms in a Motorola Nexus 6 smartphone. Hereafter, we refer to Motorola Nexus 6 as the mobile target (MT). In order to save resources in the smartphone, we set the sampling rate of the magnetic field sensor to 14Hz. However, the Wi-Fi sampling frequency is much lower, i.e., 3Hz. Table I shows the main characteristics of the mobile target used in this work. Additionally, it is necessary to have coarse-grained information about the area of interest. The system requires information related with zone distribution and physical connections among zones (i.e., zone transition information). The system reports the location of the target in real time. Zone information is also included in the coarse-grained floor plan (i.e., how the area of interest is split in zones). Figure 3 shows the zone definition and transition model in our system. The basic assumption to compute matrix \( A \) is that the likelihood of staying at the same zone is higher than the likelihood of going to another one. Thus, the transition probability matrix \( A \) was empirically defined as follows:

\[
A = \begin{bmatrix}
0.6 & 0.4 & 0 & 0 & 0 & 0 & 0 \\
0.1 & 0.4 & 0.1 & 0.1 & 0.1 & 0.1 & 0 \\
0 & 0.4 & 0.6 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.2 & 0.6 & 0 & 0 & 0.2 \\
0 & 0.4 & 0 & 0 & 0.6 & 0 & 0 \\
0 & 0.4 & 0 & 0 & 0 & 0.6 & 0 \\
0 & 0 & 0.4 & 0 & 0 & 0 & 0.6 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.6 
\end{bmatrix}
\]

To ensure conditional independence between the individual learning methods in HMM-d, we setup three completely different discriminative machine learning techniques for the zone prediction method. KStar [5].

---

**Table I: Mobile Target Specifications**

| Specification | Value |
|---------------|-------|
| Model:        | Motorola Nexus 6; OS: Android 5.1.1 |
| CPU:          | Quad-core 2.7 GHz; RAM: 3GB |
| WLAN:         | Wi-Fi a/b/g/n Magnetometer: Res.: 0.15 Ran.: 9830 |
Multilayer Perceptron (MLP) [22] and the J48 decision tree [16] machine learning algorithms were selected. Moreover, each individual machine learning method was trained with independent balanced training dataset, i.e., same number of instances for every class (zone). Since HMM-d is a zone level (i.e., room recognition) detection method, it is not needed to define any fixed survey point to build the fingerprint database. Thus, the fingerprint database for zone prediction is built by taking Wi-Fi and magnetic field measurements while walking randomly through the environment, which requires only zone labeled samples in a very short time period. Thus, the off-line phase becomes a simple process.

Internal 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 of the learning-based algorithm. Since hyperparameters have significant impact on the performance of the learning-based algorithm, we use a nested cross validation technique to adjust them [15]. Nested cross validation techniques define 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 key of hyperparameters. We optimized the global blend percentage ratio hyperparameter for KStar [5], the confidence factor for J48 [16], as well as number of hidden layers and neurons per layer for MLP [22]. Based on the parameter optimization process, we established the optimal hyperparameter values for the classifiers as follows: global percent ratio of 30% for KStar, single hidden layer with 10 neurons for MLP, and confidence factor of 0.25 for J48.

V. PERFORMANCE EVALUATION

A. Measurement Setup

Experiments were conducted in the building of the Institute of Computer Science at the University of Bern. A part of the third floor with an area of $288m^2$ ($18m \times 16m$) was chosen to deploy the localization system. The smartphone is held by a person moving along three different trajectories (Figure 4). The zone detection method is launched every time a new fingerprint measurement is available (i.e., approximately twice per second). We define 8 zones in our environment. Each zone is a wall separated area (i.e., rooms, corridor). Figure 4 presents the physical distribution of zones, ANs, and trajectories. Additionally, to compare HMM-d to another ensemble learning model, we implemented a majority voting-based method. Hereafter, we refer to majority voting-based method as Voting method. The Voting method uses predicted zone labels from KStar, J48, and MLP for the majority voting rule. Further details about majority voting-based methods can be found in [20]. All the algorithms use the same fingerprint database of Wi-Fi RSS and MF readings, which have been measured during the data collection procedure. The individual predictors can be regarded as traditional fingerprint-based approaches, while the proposed HMM-d is a new ensemble predictor.

B. Room Level Localization Accuracy

To evaluate our prediction models, we consider three measures: prediction accuracy, F1 score, and processing time. In classification problems, accuracy is the ratio of correctly predicted observation to the total observations. F1 is the harmonic mean of precision and sensitivity. Precision is defined as the number of true positives (TP) divided by TP and the number of false positives (FP): $TP/(TP+FP)$. Sensitivity is defined as the number of TP divided by TP and the number of false negatives (FN): $TP/(TP+FN)$ [19]. Thus, F1 considers both performance measures, precision, and sensitivity. F1 can be written as follows:
Figure 4 shows the accuracy of zone prediction for the five predictors in the three trajectories. Figure 6 presents the F1 score by zone for the learning algorithms. Figure 7 shows the average of prediction processing time for both ensemble models HMM-d and Voting. Due to the hyperparameter optimization process, performance accuracy of the individual learning models (i.e., KStar, J48 and MLP) is higher than 80%. However, results show a clear improvement between the individual learning models and our ensemble learning model HMM-d. As it can be seen in Figure 5, our proposed HMM-d model outperforms KStar, J48, and MLP algorithms in the three tested trajectories. Accuracy of HMM-d is improved by 9.17%, 4%, 9.3%, and 9.2% compared to J48, KStar, MLP, and the Voting method respectively. Unlike Voting, HMM-d is able to combine zone transition information with individual learning methods to improve prediction accuracy.

Considering the F1 score, HMM-d outperforms others in all tested zones (i.e., classes). Therefore, HMM-d outperforms Voting, KStar, J48, and MLP considering accuracy and robustness. As it can be seen in Figure 5, accuracy of the learning algorithms remains similar (i.e., not significant difference) in all tested trajectories. However, if we consider sensitivity and precision as performance measure, it is clear to notice that zone 2 is the hardest zone to classify correctly (as shown in Figure 5). This result is explained as zone 2 corresponds to the corridor (see Figure 4a). Thus, fingerprints obtained in this zone are very similar to fingerprints obtained in adjacent zones, especially in the border areas. However, HMM-d improves the classification accuracy.
in this zone by 10.2%, 10.4%, 4% and 8% compared to Voting, J48, KStar, and MLP respectively. Thus, by combining HMM, KStar, J48, and MLP, our approach allows the production of better predictive performance compared to individual and ensemble voting models. Unlike Voting, the HMM-d model is able to achieve better prediction performance than individual learning methods in all zones. As shown in Figure[4] the average time of prediction is very similar in both HMM-d and Voting ensemble methods. However, HMM-d reduces the prediction processing time by 0.0125 ms compared to the Voting method. This means that considering transition zone information in the prediction process takes lower computational efforts than processing the voting rule [20] that is used in voting methods. Thus, HMM-d model overcomes Voting, KStar, J48, and MLP methods by accuracy, robustness an processing time.

VI. CONCLUSIONS

In this work, we proposed a terminal-based room level indoor localization system, which integrates discriminative learning techniques in an ensemble learning method using ubiquitous Wi-Fi and magnetic field fingerprints. Our proposed ensemble learning model achieves high prediction performance by combining less accurate individual discriminative learning models. We define rooms as zones, and adopted the Hidden Markov Model to integrate information about transition probabilities between zones and discriminative learning methods. Thus, this work presents a probabilistic-based system to achieve high room level localization accuracy of a smartphone user, who is moving in a multi room indoor environment. We evaluated our system in a complex real-world indoor environment. Evaluation results indicate that our proposed approach is more accurate and robust than individual learning and majority voting-based models.

ACKNOWLEDGMENT

This work was partly supported by the Swiss National Science Foundation, project no. 154458.

REFERENCES

[1] H. Abdelnasser, R. Mohamed, A. Elgohary, M. Alzantot, H. Wang, S. Sen, R. Choudhury, and M. Youssef, “Semanticslam: Using environment landmarks for unsupervised indoor localization.” IEEE Transactions on Mobile Computing, vol. 15, pp. 1770–1782, 2016.
[2] P. Bahl and V. Padmanabhan, “Radar: an in-building rf-based user location and tracking system.” Joint Conference of the IEEE Computer and Communications Societies. Proceedings, vol. 19, 2000.
[3] J. Borestein and L. Ojeda, “Heuristic drift elimination for personnel tracking systems.” The Journal of Navigation,63(3), pp. 591–606, Oct 2010.
[4] J. Carrera, Z. Li, Z. Zhao, and T. Braun, “A real-time indoor tracking system in smartphones.” Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MMSIM), November 2016.
[5] G. Cleary and E. Trigg, “K*: An instance-based learner using an entropic distance measure.” in I2th International Conference on Machine Learning, 1995, pp. 108–114.
[6] F. Dieter, w. Burgard, and S. Thrun, “Markov localization for mobile robots in dynamic environments,” J. Artif. Int. Res., vol. 11, no. 1, pp. 391–427, Jul. 1999.
[7] G. Forney, “The viterbi algorithm,” Proceedings of the IEEE, vol. 61, pp. 268–278, March 1973.
[8] S. He and G. Chan, “Wi-fi fingerprint-based indoor positioning: recent advances and comparisons.” IEEE Communications Survey Tutorials, 2016, 2016.
[9] F. Hong, H. Chu, L. Wang, Y. Feng, and Z. Guo, “Pocket mattering: Indoor pedestrian tracking with commercial smartphone,” International Conference on Indoor Positioning and Indoor Navigation, 2012.
[10] N. Kakihisi and S. Kamijo, “Pedestrian dead reckoning for mobile phones through walking and running mode recognition,” in The 16th International IEEE Conference on Intelligent Transportation Systems (ITSC), pp. 261–267, Oct 2013.
[11] Z. Li, D. Dimitrova, and T. Braun, “A passive wifi source localization system based on fine-grained power-based triellation,” IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks ( WoWMoM), 2015.
[12] ——, “A time-based passive source localization system for narrow-band signal,” The IEEE International Conference on Communications (ICC), vol. 39, pp. 4599–4605, June 2015.
[13] J. Liu, R. Chen, L. Pei, R. Guinnes, and H. Kussiemi, “A hybrid smartphone indoor positioning solution for mobile lbs,” Sensors, 2012.
[14] A. Mariakakis, S. Sen, J. Lee, and K. Kim, “Sail: Single access point-bases indoor localization." Proceedings of the 12th annual conference on Mobile systems, applications and services Mobisys ’14, pp. 315–328, June 2014.
[15] K.-R. Miller, M. Krauledat, G. Dornhege, G. Curio, and B. Blankertz, “Machine learning techniques for brain-computer interfaces.” BIOMEDICAL ENGINEERING, pp. 11–22, 2004.
[16] R. Quinlan, C4.5: Programs for Machine Learning. San Mateo, CA: Morgan Kaufmann Publishers, 1993.
[17] J. Seitz, T. Vauapel, S. Meyer, B. Gutierrez, and J. Thielecke, “A hidden markov model for pedestrian navigation.” in Proceedings of the positioning navigation and communication, pp. 120–127, march 2012.
[18] D. Taniuchi and T. Maekawa, “Robust wi-fi based indoor positioning with ensemble learning.” IEEE 10th International Conference on Wireless and Mobile Computing, Networking and Communications, pp. 592–597, 2014.
[19] H. Thornburg. (2016, Mar.) Accuracy, precision, recall and f1 score: Interpretation of performance measures.
[20] ——. (2017, Mar.) sklearn.ensemble.votingclassifier. [Online]. Available: http://weka.sourceforge.net/doc.dev/
[21] M. Wallbaum and O. Spaniol, “Indoor positioning using wireless local area networks,” in Proceedings of the IEEE John Vincent Atanasoff International Symposium on Modern Computing., pp. 17–26, Oct 2006.
[22] D. WEKA. (2018, Mar.) Class multilayerperceptron. [Online]. Available: http://weka.sourceforge.net/doc.dev/
[23] N. Yepeng, L. Jianbo, A. Liu, and Y. Bai, “An indoor pedestrian positioning method using hmm with a fuzzy pattern recognition algorithm in a wlan fingerprint system,” Sensors, Sep 2016.
[24] M. Yousef and A. Agrawala, “The horus wlan location determination system; Proceedings of the International Conference on Mobile Systems, Applications, and Services (MobiSys’05), pp. 205–218, 2005.