

A Dynamic Indoor Location Model for Smartphones Based on Magnetic Field: A Preliminary Approach

Carlos E. Galván-Tejada¹(✉), Jorge I. Galván-Tejada¹,
José M. Celaya-Padilla², J. Rubén Delgado-Contreras³,
Vanessa Alcalá-Ramírez¹, and Luis Octavio Solís-Sánchez²

¹ Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas,
Ave. Ramón López Velarde 801, 98064 Zacatecas, Zacatecas, Mexico
ericgalvan@uaz.edu.mx

² Laboratorio de Innovación y Desarrollo Tecnológico en Inteligencia
Artificial (LIDTIA), Universidad Autónoma de Zacatecas,
Ave. Ramón López Velarde 801, 98064 Zacatecas, Zacatecas, Mexico

³ Instituto Tecnológico Superior Zacatecas Sur (ITZaS),
Tlaltenango, Zacatecas, Mexico

Abstract. Due to an increase interest for providing services based on user location, several indoor location approaches based on mobile devices have been proposed recently. This paper focuses on the use of a novel crowdsourcing approach for indoor location of a mobile device that uses social collaboration to improve the accuracy and magnetic field signal as information source using feature extraction and a deterministic method that allows us to include information from new users that improves the fitness of the model. Four phases were included in the methodology: Raw data collection, Data pre-process, Feature extraction and Social collaboration. An experiment was successfully carried out to test the proposed methodology. On the whole, good results were obtained on computational cost, recalculation time and accuracy improvement.

Keywords: Indoor positioning · Crowdsourcing · Social collaboration · ILS

1 Introduction

Recently there has been increased interest for providing services based on user location. These services are characterized by integrating the location or position of a mobile device with other value-added information to a user [5]. However, to achieve this, there are several challenges, such as the proposal of a model of accurate Indoor Location System (ILS) based on mobile devices, which does not require a dedicated infrastructure and uses a minimal amount of computer resources.

To tackle the mentioned challenges, several indoor location approaches based on mobile devices have been proposed, for instance, (i) *inertial based*: This approach is characterized by using the inertial sensors such as accelerometer and gyroscope that are used to estimate the location of the user. The accelerometer can be used to determine the changes in the user's position produced when an acceleration is detected in one or more axes. While the gyroscope can be used to detect the changes in the direction to improve the location estimation [10, 13], (ii) *camera based systems*: This approach uses the camera of the mobile device to capture information from the user's location (e.g. an image, video, markers or codes). The captured information is then compared with the reference information that was previously collected [11, 17]. The main disadvantage of this approach is that it requires high processing capacity and the accuracy of the system is reduced when the quality of captured information presents low resolution or motion blur, (iii) *mobile signal processing*: This approach uses the sensors embedded in a mobile device to detect, measure and capture signals that are emitted by other devices and then conveyed inside the indoor environment (e.g. Wi-fi, Bluetooth), as well as the natural signals that are commonly found in those indoor environments (e.g. magnetic-field, ambient sound). In these systems the position estimation is commonly performed through methods such as "fingerprinting". This latter is composed of two phases: training and position determination. Firstly, a map of the observed signal strength values measured at different locations is recorded during a training phase. Secondly, the signal strength values observed at a user's device are compared with the map values by using proximity matching algorithms, including [2, 15], but not limited to k-Nearest Neighbors (k-NN) [12]. The main disadvantage of this approach is the calibration phase. The calibration phase consist in point-by-point mapping of a given indoor environment, measuring the magnitude and/or direction of a specific signal at each point, and then, using this signal map for location purposes, finding the most similar place in the signal map to the one detected at a given point.

In order to avoid the calibration phase at Indoor Location Systems (ILS) we proposed a novel approach called crowdsourcing, described as a new online distributed problem solving and production model in which networked people collaborate to complete a task [1, 8, 18]. When exploring this approach, several works has been proposed to tackle the indoor location and mapping issue, for instance CrowdInside [1] proposes a data collection using inertial sensor data that is generated by the user's motion, this raw data is then used to construct indoor maps. In other hand, CrowdMap [4] uses a combination of computational vision and other embedded sensors to reconstruct maps, they divide their proposal in 4 steps, including one that process video to find key-frames. Nevertheless, these approaches requires an expensive computational cost phase to integrate the new information that comes from new collaborative users. This paper focuses on the use of a crowdsourcing approach for indoor location of a mobile device that uses social collaboration to improve the accuracy and magnetic field signal as information source using feature extraction and a deterministic method that allows us to include information from new users that improves the fitness of the ILS.

2 Methodology

We modified the methodology proposed by Galvan-Tejada et al. [6] with the aim of using a crowdsourcing approach for indoor location of a mobile device. Although in essence the methodology is similar, feature selection methods based on filters were used and a new phase named “social collaboration” was added as is shown in Fig. 1. This modified methodology is described as follows:

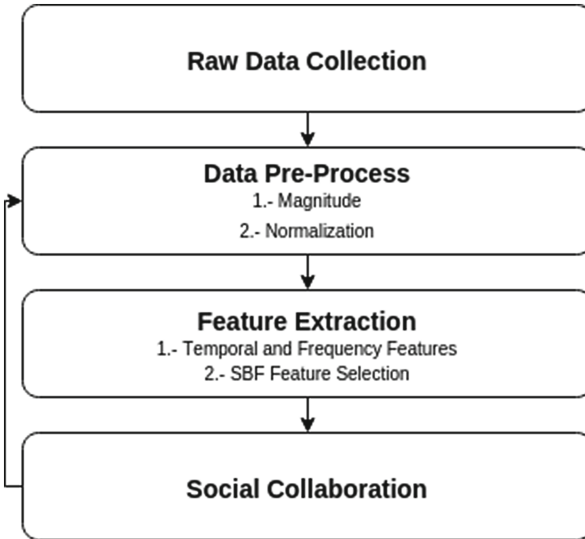


Fig. 1. Four phases proposed methodology.

2.1 Phase 1: Raw Data Collection

At this stage the magnetic field intensity of an indoor environment is captured and stored through a mobile application that uses the magnetometer of a smartphone. This sensor returns three float values, indicating three magnetic field axes in the environment when the lecture is done.

To capture magnetic field measures the user should walk in the indoor environment with the active application on the smartphone. The mobile device must be on hand at waist level and the walking must be done at average walk speed during 10 seconds. The data lectures were collected at a frequency of 100 lectures per second. We named *signatures* to this basic set of data that represent a room.

2.2 Phase 2: Data Pre-process

Once the raw data was collected, a pre-process is needed to get usable data without outliers and anomalies in the final data set. The pre-process is done following the next two steps:

Magnitude. Raw data from the magnetic field is composed by three components. For the sake of simplicity, a single data value of the magnitude was calculated using the Eq. 1.

$$|M| = \sqrt{M_x^2 + M_y^2 + M_z^2} \quad (1)$$

Magnitude allows us to avoid the constraint of smartphone position (i.e. screen position respect to the hand).

Normalization. To avoid the spatial scaling a Z norm normalization was applied to each light lecture signature using the Eq. 2, $z_{i,d}$ where $z_{i,d}$ is the normalized reading, $r_{i,d}$ refers to the i^{th} observation of the signature in dimension d μ_d is the mean value of the signature for dimension d and σ_d is the standard deviation of the signature for dimension d .

$$\forall i \in m : z_{i,d} = \frac{r_{i,d} - \mu_d}{\sigma_d} \quad (2)$$

Equation 2 was applied for all dimensions in R^d .

Once initial magnetic signatures were collected in the rooms and stored in a data set created during the signature collection, feature extraction and selection were carried out to get the initial classification model that allows us to estimate the location of users.

2.3 Phase 3: Feature Extraction

The feature extraction is a process that consists in carrying out an efficient data reduction while preserving the appropriate amount of signal information. In this sense, the feature extraction step allow us to reduce the amount of data to develop the indoor location estimation model. To apply this step we must select the features that will be extracted from the signal, these activities are described below.

Signal Features. To select the signal features that describe the behavior of the magnetic field signal for ILS problem, we used the set features proposed by Galvan-Tejada et al. [6] because it includes first and second order statistical descriptors and Digital Signal Processing (DSP) which are well-known features. This set of features are appreciated in Table 1.

Temporal features were computed from the waveform of magnetic field signal that is generated from the raw data previously recorded, and *Spectral features* extracted from the signal after performing a P-point Fast Fourier Transform to each signature [16].

Feature Selection. To improve the model fitness and getting an indoor location model with less signal features a feature selection process was applied.

To perform a feature selection several techniques have been proposed, for instance filter, wrapper and embedded methods [14]. From these methods,

Table 1. Features extracted.

Features	Temporal Domain	Frequency Domain
Kurtosis	*	*
Mean	*	*
Median	*	*
Standard Deviation	*	*
Variance	*	*
Coefficient of Variation (CV)	*	*
Inverse CV	*	*
1,2,3 Quartile	*	*
1,5,95,99 Percentile		*
Trimmed Mean	*	*
Shannon Entropy		*
Slope		*
Spectral Flatness		*
1-10 Spectrum Components		*

we identified that filter methods have two advantages over other methods: (1) Returns features that can be used with different classifiers and (2) Less computationally expensive. From filter methods we selected the Selection By Filter (SBF) method that was used to the feature selection process. This is a cross validated feature selection strategy [9]. The SBF is set up to fit with a Random Forest function. Once the features are selected, these can be used with several classifiers to get a model to estimate the location, which is the core of an ILS.

2.4 Phase 4: Social Collaboration

To improve the accuracy of the ILS, signatures from different users and sensors must be added to the data set of signatures. Data from single users must be processed with the first two phases and added to the data set of signatures before the phase 3 is carried on. This procedure has the aim to perform a feature extraction and selection from a bigger signatures data set, that includes more variability given the new signatures of different users with different devices.

3 Experiment

The experiment was carried out in a school building shown in Fig. 2. Experimental data is available in the research group website¹. Thirteen (13) rooms were selected considering their distribution. Some spaces share magnetic field data in some parts of rooms (i.e. classrooms that were divided in two).

¹ <http://ingsoftware.reduaz.mx/>.

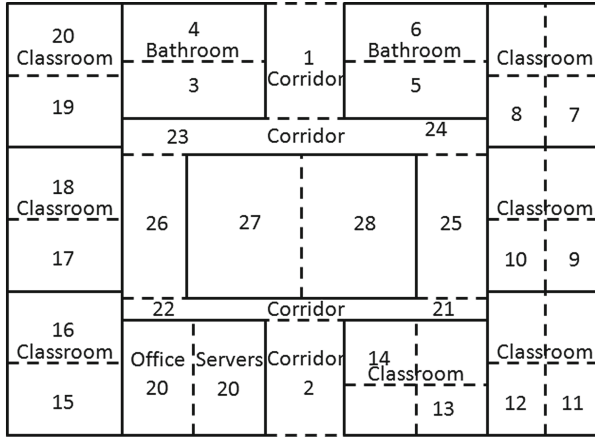


Fig. 2. Distribution plane of school building used in our experiment.

Even when our proposal is a methodology that can be improved adding new signatures, to begin the experimentation and get the initial model to the ILS, we use the Eq. 3 as is proposed by Galvan-Tejada et al., where x is the minimal number of signatures, and N is the number of variables used in the experiment. In our experiment N is equal 598 (13 rooms multiplied by 46 features each). We got a total of 10.22 magnetic signatures, finally uppered to 11.

$$x = \log_2(N) + 1 \tag{3}$$

We choose a Motorola G2, with BOSCH BMC150 magnetic sensor to record the initial magnetic data, using an Android application, developed with the aim to get magnetic sensor data. For the second user, we select a Sony Xperia 7 with MPC Magnetic Field sensors, and finally a third user with a Nexus 7. All sensors were selected to be different with the purpose of test the ILS in a common situation in where we have a diverse type of sensors and users.

After the raw data was collected, the pre-process step was applied using a script in R, which is a freely available language and environment for statistical computing and graphics², because it is multi-plataform and an environment for statistical computing.

After feature matrix is normalized, the SBF is applied using the CARET package in R [7]. The frame work used consist in five phases:

1. Data equalization and normalization.
2. Use the SBF to obtain their meaningful features.
3. Obtain a classification rate using the filtered model with Random Forest.
4. Obtain a classification rate using 46 selected features with Random Forest.
5. Calculate a Confusion Matrix for comparison.

² <http://www-r-project.org/>.

Once the SBF features are selected, a Random Forest classifier composed of 5,000 trees is used to evaluate indoor location model conformed by features selected by filter method. This classifier was proposed by Breiman et al. [3], and selected because it provides tree ensembles that depend on the values of a random feature vector and provides the same distribution to all the trees included in the forest. The decision of this classifier depends on the decision of several trees.

All classification process was performed using Random Forest R package.

In order to improve the fitness of the ILS we proceed to add signatures from the second and third user. The signatures following the pre-process and feature extraction steps. The data from the second user is added first to the data set of features generated from the first user, then, the model is recalculated. Finally, the features from the last user were added to the data set, and the model is recalculated.

4 Results

The SBF fitted model was a 23 feature model. In Table 2 those features included on the model are described.

Table 2. SBF Model Features.

Features	Temporal Domain	Frequency Domain
Kurtosis		
Mean	*	*
Median	*	
Standard Deviation	*	*
Variance	*	*
Coefficient of Variation (CV)	*	*
Inverse CV	*	*
1,2,3 Quartile		
95,99 Percentile	*	
1,5,95,99 Percentile		*
Trimmed Mean		*
Shannon Entropy		*
Slope		*
Spectral Flatness		*
2,8 Spectrum Components		*

After experimental test, the final filtered model from the first user data set, was a 23 feature model of classification, the obtained model can classify correctly 41.11 % of the samples. To improve the accuracy of the model, we added data

from a second user and recalculated the model, this phase was done in less than a second given the deterministic approach to calculate a new model. After adding the data from the second user, the SBF model was the same 23 features model that was acquired with the data from a single user, however the accuracy was improved dramatically to 98.89%. Finally, a third user was added to the data set; again the model was recalculated in less than a second even when the raw data was increased to 1,404,000 data points. After the SBF was applied, the same 23 features model was acquired and the accuracy was improved reaching a non error classification (100%).

To verify the improvement using our proposal we tested a Random Forest (RF) using the 46 features extracted from the magnetic signal. This test was carried on in the 3 cases presented with the SBF approach. In Table 3 is shown the accuracy of a random forest model using 46 features against the 23 model acquired with the SBF selection. For all the cases, RF uses twice the information than the cases of the model acquired with SBF given the number of features needed to develop the model.

Table 3. Accuracy with different number of users.

Users	Random Forest Accuracy	SBF model Accuracy
One User	39.40	41.11
Two User	100	98.11
Three User	100	100

5 Concluding Remarks

In this paper we presented a novel approach for indoor location that uses social collaboration to improve the accuracy and magnetic field signal as information source.

The following important aspects regarding the proposal were identified:

- *Computational cost.* Compared against other ILS proposals, as the case of Genetic Algorithms (GA), the computational cost is lower given that wrapped methods avoid the stochasticity of other methods (as those based in a GA approach).
- *Recalculation Time.* Time became the most important characteristic to take into account when developing a crowdsourcing ILS because the recalibration step is very time consuming. However using a deterministic feature selection method (SBF) allows us to do this task in less than a second.
- *Accuracy Improvement:* Because a diversity applications of an ILS, is a must to consider that will be used by several users with different devices, walking behaviors and morphological characteristics and this implies that ILS must be able to estimate location even with those differences. In this paper we presented how adding new users considering the aforementioned differences, not just can be predicted; moreover, the model increase the accuracy of location estimation of the users.

Future work should take note the possibility of sampling a larger space and try to use different types of feature selection and ways to integrate new information (social collaboration) to the ILS. Additionally other tools should be used to improve the performance of mobile phones and tablets.

References

1. Alzantot, M., Youssef, M.: Crowdinside: automatic construction of indoor floor-plans. In: Proceedings of the 20th International Conference on Advances in Geographic Information Systems, pp. 99–108. ACM (2012)
2. Bilke, A., Sieck, J.: Using the magnetic field for indoor localisation on a mobile phone. Progress in Location-Based Services. Springer, Heidelberg (2013)
3. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
4. Chen, S., Li, M., Ren, K., Qiao, C.: Crowd map: Accurate reconstruction of indoor floor plans from crowdsourced sensor-rich videos. In: 2015 IEEE 35th International Conference on Distributed Computing Systems (ICDCS), pp. 1–10. IEEE (2015)
5. D’Roza, T., Bilchev, G.: An overview of location-based services. BT Technol. J. **21**(1), 20–27 (2003)
6. Galván-Tejada, C.E., García-Vázquez, J.P., Brena, R.F.: Magnetic field feature extraction and selection for indoor location estimation. Sensors **14**(6), 11001–11015 (2014)
7. Jed Wing, M.K.C., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z.: The R Core Team: caret: Classification and Regression Training (2014), r package version 6.0-24. <http://CRAN.R-project.org/package=caret>
8. Jiang, Y., Xiang, Y., Pan, X., Li, K., Lv, Q., Dick, R.P., Shang, L., Hannigan, M.: Hallway based automatic indoor floorplan construction using room fingerprints. In: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 315–324. ACM (2013)
9. Kuhn, M.: Variable selection using the caret package (2012). URL: <http://cran.cermin.lipi.go.id/web/packages/caret/vignettes/caretSelection.pdf>
10. Li, F., Zhao, C., Ding, G., Gong, J., Liu, C., Zhao, F.: A reliable and accurate indoor localization method using phone inertial sensors. In: Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pp. 421–430. ACM (2012)
11. Mulloni, A., Wagner, D., Barakonyi, I., Schmalstieg, D.: Indoor positioning and navigation with camera phones. Pervasive Comput. IEEE **8**(2), 22–31 (2009)
12. Ni, L.M., Liu, Y., Lau, Y.C., Patil, A.P.: Landmarc: indoor location sensing using active rfid. Wirel. Netw. **10**(6), 701–710 (2004)
13. Pratama, A., Widyawan, H.R.: Smartphone-based pedestrian dead reckoning as an indoor positioning system. In: 2012 International Conference on System Engineering and Technology (ICSET), pp. 1–6, Sept 2012
14. Saeys, Y., Inza, I., Larrañaga, P.: A review of feature selection techniques in bioinformatics. Bioinformatics **23**(19), 2507–2517 (2007)
15. Storms, W., Shockley, J., Raquet, J.: Magnetic field navigation in an indoor environment. In: Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), pp. 1–10. IEEE (2010)
16. Tsai, W.H., Tu, Y.M., Ma, C.H.: An FFT-based fast melody comparison method for query-by-singing/humming systems. Pattern Recogn. Lett. **33**(16), 2285–2291 (2012). <http://www.sciencedirect.com/science/article/pii/S016786551200284X>

17. Werner, M., Kessel, M., Marouane, C.: Indoor positioning using smartphone camera. In: 2011 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–6. IEEE (2011)
18. Zhang, X., Yang, Z., Wu, C., Sun, W., Liu, Y., Liu, K.: Robust trajectory estimation for crowdsourcing-based mobile applications. *Parallel Distrib. Syst. IEEE Trans.* **25**(7), 1876–1885 (2014)