Applying Classifiers in Indoor Location System

Gabriel Villarubia, Francisco Rubio, Juan F. De Paz,
Javier Bajo, and Carolina Zato

Abstract. Research in indoor location has acquired a growing importance during the recent years. The main objective is to obtain functional systems able of providing the most precise location, identification and guidance in real time. Currently, none of the existing indoor solutions have obtained location or navigation results as precise as the ones provided by the analog systems used outdoor, such as GPS. This paper presents an indoor location system based on Wi-Fi technology which, from the use of intensity maps and classifiers, allows effective and precise indoor location.

Keywords: indoor location system, classifiers, Wi-Fi.

1   Introduction

Nowadays, indoor location is one of the most promising research fields in mobile computing [7], and it is possible to find several studies. To obtain precise information about the location of objects or people inside building can provide very useful information to develop services. Some of the most relevant of these services are those oriented to identification-based access control, location-based security, location-aware computing, etc.. These services are deployed in indoor environments, such as hospitals, factories, shopping malls, or even as complementary systems to GPS (Global Positioning System). The main challenge in current indoor location systems is to obtain precise measures with a reasonable cost in infrastructure.

Gabriel Villarubia · Francisco Rubio · Juan F. De Paz · Carolina Zato
Department of Computer Science and Automation, University of Salamanca,
Plaza de la Merced, s/n, 37008, Salamanca, Spain
e-mail: {gvg,paco_rg,fcofds,carol_zato}@usal.es

Javier Bajo
Departamento de Inteligencia Artificial, Facultad de Informática,
Universidad Politécnica de Madrid, Madrid, Spain
e-mail: jbajo@fi.upm.es

The main reason to explain why it is not possible to find effective indoor solutions is due to technical and financial reasons. If we take into consideration a GPS system, we simply need a physical device that connects to a finite number of satellites in open spaces. However, in an open space, it is required to make use of a technological infrastructure with a considerable number of fixed devices that are used as beacons or readers, and notably improve the cost of the system. Amongst the technologies that are currently used most in the development of Real-Time Location Systems are, RFID (Radio Frequency IDentification) [8] [10], Wi-Fi y ZigBee [8] [9]. Nowadays, some of the existing indoor location engines make use of 802.11 technology also known as Wi-Fi, and have acquired a growing importance due to the reduced and reasonable cost of these systems compared to the alternatives. Moreover, Wi-Fi provides a good number of advantages: easy deployment and is integrated in most of the current electronic devices, wide range of existing networks installed in several buildings and locations.

This paper presents a study aimed at obtaining the design of an innovative Wi-Fi indoor location engine, able of providing location-aware Information of people or objects inside a building. The use of Wi-Fi technology for location systems is mainly based on intensity maps constructed from RSSI levels in different zones. The maps are used as a basis to obtain classifications. The classifiers use the data of the maps and the data obtained from the devices to determine the position of a person inside a building. The system presented in this paper has been adapted to be installed in resource-constrained devices. In this way it is not necessary to connect to databases to obtain information for the classification, which notable improve the performance of the system and the fault tolerance.

This article is divided as follows: section 2 describes the state of the art; section 3 presents the proposed model; section four describes the results obtained and the conclusions respectively.

2   Background

There exist three main algorithms that are used by real-time location systems to determine the location of the mobile nodes (tags): Triangulation, Fingerprinting and Multilateration [4]. Triangulation allows obtaining location coordinates by means of the calculus of the length of the sides of the triangle from the angles of the received signal in each of the antennas, which requires at least 3 reference points. Fingerprinting, also known as signpost or symbolic location, is based on the study of the characteristics of the each of the location zones, obtaining measurements of the radiofrequency characteristics and estimating the influence area of each tag. [5]. Multilateration estimates the distances between readers and tags, taking into account parameters as RSSI (Received Signal Strength Indication) or TDOA (Time Difference Of Arrival) [6], in such way that the intersection of the estimated distances from each tag to three or more fixed nodes determines the point where the tags are identified. Multilateration provides better results that triangulation when it is used outdoors, but its performance is notably reduced

indoors. The reason is fluctuation of the RSSI levels in indoor environments due to the presence of different elements (people, objects, animals). Besides, multilateration is based on the estimation of distances, and it is necessary a previous estimation of the RSSI values, which is difficult because the RSSI values change constantly.

Location techniques based on triangulation and multilateration cannot be considered as efficient because the signals suffer several attenuations caused by the elements present in the rooms. In these cases, it is recommended to use heuristics for the classification process and perform training to obtain location algorithms that improve the precision. The heuristics allow the collection of possible measurements in specific positions and to use these measurements to calculate the most probable position.

3 Proposed Reasoning System

Classification techniques facilitate the association of cases to the existing groups. The behavior of these algorithms is similar to the algorithms used for clustering, but they are too much simple in most of the cases. The existing techniques can be grouped in the following categories depending on the nature of the algorithms:

- Decision rules and decision trees.
- Probabilistic models: Naive Bayes [11], Bayesian networks [1] [2] [3].
- Fuzzy logic: K-NN (K-Nearest Neighbours), NN (K-Nearest Neighbours).
- Function search: Sequential Minimal Optimization (SMO) [12].
- Artificial neural networks.

Specifically, the proposed system has integrated a bayesian network to estimate the probabilities of belonging to the points previously scanned in the intensity maps. The intensity maps obtained indoors are created using the parameters shown in Table 1. Each of the rows in Table 1 contains the information of all the Wi-Fi networks scanned in that moment and identified by a coordinate (x,y). The rows can contain more or less columns depending on the scanned intensities.

Table 1 Format of the measurements in the intensity maps

x	y	SSID	BSSID	RSSI	SSID	...

The scanned measurements, represented using the format shown in Table 1, are used to obtain the distances that are used by the bayesian network which is trained and configured for further estimations. In order to build Bayesian networks, it is first necessary to establish search mechanisms that can generate the DAG (Directed Acyclic Graph) using a set of heuristics that can reduce the number of combinations and generate the final Bayesian network. There are various Bayesian network search mechanisms, including tabu search [1], conditional independence

[2], K2 [1], HillClimber [1], TAN (Tree Argumented Naive Bayes) [3]. In this work, we have used conditional independence. This algorithm is based on the calculation of the conditional Independence test for the variables to generate a DAG that can obtain the probability estimates. If the variables being studied are independent, it will not be possible to generate a Bayesian network with good results, but in this work this is not a problem with the independence of the variables. The attribute for the classification in the bayesian network is defined as the union of x and y.

4   Results and Conclusions

To analyze the overall performance of the system, we defined a case study in the University of Salamanca. The first step was to calibrate the map for the first floor of the Physics building at the Faculty of Sciences. During the calibration process, we obtained different signal measurements in the corridors of the floor. The classifier was constructed from these data. The total number of measurements was 380, as shown in Figure 1a. The area of the floor of the building was approximately 1700m2. The floor was equipped with 4 fixed access points spread across the floor of the building. Moreover, some signals emitted by access points that went out and lit corresponding to offices and laboratories were detected. The access points were the infrastructure nodes available in the university, and it was not necessary to introduce any additional access points to reduce the system error. Figure shows a screenshot obtained from the mobile device used for the experiment.

Fig. 1 A) First floor of the Physics building. B) Mobile interface of the system.

The system was evaluated obtaining measures in different points and estimating the position. We obtained 332 new measurements and estimated the corresponding positions. The results obtained are shown in Table 2. The absolute mean error obtained is shown in the first column and, as can be seen in Table 2, the bayesian network provides the best value for the error, 2,4 meters. The average time to estimate the position is 0,65 milliseconds, which is good enough to be used in a mobile device. It is necessary that the average time has been obtained in a laptop. It has not been possible to measure the time in a mobile device, but the system perfectly worked in the mobile during the experiment.

Table 2 Errors and time to estimate positions using different classifiers

Error meters	Time millisecond	Classifier
2,414	0,652632	.BayesNet
3,040	2,126316	NaiveBayes
3,040	1,957895	NaiveBayesUpdateable
4,147	37,989475	SimpleLogistic
8,890	22,660526	SMO
3,945	0,281579	Ibk
8,299	11,03421	LWL
2,857	14,047368	KStar
9,430	0,078947	AdaBoostM1
3,830	0,281579	AttributeSelectedClassifier
5,035	1,452632	Bagging
4,256	6,481579	ClassificationViaRegression
12,605	0,042105	CVParameterSelection
4,325	0,2	FilteredClassifier
4,306	8,652632	LogitBoost
12,624	7,476316	MultiClassClassifier
12,605	0,052632	MultiScheme
3,826	0,4	RandomCommittee
3,560	0,768421	RandomSubSpace
12,605	0,047368	Stacking
12,510	5,247368	DecisionTable
4,572	1,760526	JRip
13,370	0,044737	OneR
6,488	1,639474	PART
12,605	0,039474	ZeroR
9,430	0,068421	DecisionStump
7,819	0,255263	J48
14,142	90,007896	LMT
3,346	0,415789	RandomForest
2,918	0,092105	RandomTree
2,649	0,194737	REPTree

As a conclusion, it is possible to say that the work presented in this paper facilitates a new technique to locate objects and people using Wi-Fi signals in indoor environment, making use of the existing infrastructure networks. The system does not require the installation of additional access points or any other hardware to improve the system performance. Thus, the implementation cost is notably reduced compared to most of the existing systems. Our future work focuses on the incorporation of new measurements into the algorithm, such as GSM or 3G that can be combined with the Wi-Fi signals and improve the overall performance of the system.

References

1. Bouckaert, R.R.: Bayesian Belief Networks: from Construction to Inference, Utrecht, Netherlands (1995)
2. Verma, T., Pearl, J.: An algorithm for deciding if a set of observed independencies has a causal explanation. In: Proc. of the Eighth Conference on Uncertainty in Artificial Intelligence, pp. 323–330 (1992)
3. Friedman, N., Geiger, D., Goldszmidt, M.: Bayesian Network Classifiers. Machine Learning 29, 131–163 (1997)
4. Glassner, A.: Principles of digital image synthesis. Morgan Kaufmann (1995)
5. Georgé, J.P., Gleizes, M.P., Glize, P.: Emergence of organisations, emergence of functions. In: Symposium on Adaptive Agents and Multi-Agent Systems, pp. 103–108 (2003)
6. De Paz, J.F., Rodríguez, S., Bajo, J., Corchado, J.M.: Multi-agent system for security control on industrial environments. International Transactions on System Science and Applications Journal 4(3), 222–226 (2008)
7. Chen, Y.-C., Chiang, J.-R., Chu, H.-H., Huang, P., Tsuid, A.W.: Sensor-Assisted Wi-Fi Indoor Location System for Adapting to Environmental Dynamics (2011)
8. Razavi, R.S., Perrot, J.-F., Guelfi, N.: Adaptive modeling: An approach and a method for implementing adaptive agents. In: Ishida, T., Gasser, L., Nakashima, H. (eds.) MMAS 2005. LNCS (LNAI), vol. 3446, pp. 136–148. Springer, Heidelberg (2005)
9. Giunchiglia, F., Mylopoulos, J., Perini, A.: The tropos software development methodology: Processes, models and diagrams. In: AAMAS 2002 Workshop on Agent Oriented Software Engineering (AOSE 2002), pp. 63–74 (2002)
10. Tapia, D.I., De Paz, J.F., Rodríguez, S., Bajo, J., Corchado, J.M.: Multi-Agent System for Security Control on Industrial Environments. International Transactions on System Science and Applications Journal 4(3), 222–226 (2008)
11. Duda, R.O., Hart, P.: Pattern classification and Scene Analysis. John Wisley & Sons, New York (1973)
12. John, C.: Platt Fast training of support vector machines using sequential minimal optimization. In: Advances in Kernel Methods, pp. 185–208 (1999)