Fingerprint Database Variations for WiFi Positioning

Nelson Acosta, Juan Toloza, Carlos Kornuta
INCA/INTIA - School of Exact Sciences – UNICEN, TANDIL - Argentina
nacosta@exa.unicen.edu.ar
Post-Doctoral Fellow, CONICET, INCA/INTIA - School of Exact Sciences – UNICEN,
TANDIL - Argentina
jmtoloza@exa.unicen.edu.ar
Doctoral Fellow, CONICET, INCA/INTIA - School of Exact Sciences – UNICEN,
TANDIL - Argentina
ckornuta@exa.unicen.edu.ar

ABSTRACT

Indoor positioning systems calculate the position of a mobile device (MD) in an enclosed environment with relative precision. Most systems use WiFi infrastructure and several positioning techniques, where the most commonly used parameter is RSSI (Received Signal Strength Indicator). In this paper, we analyze the fingerprinting technique to calculate the error window obtained with the Euclidian distance as main metric. Build variations are presented for the Fingerprint database analyzing various statistical values to compare the precision achieved with different indicators.

Keywords

Indoor Positioning, RSSI, Fingerprint.

Council for Innovative Research

1. INTRODUCTION:

In recent years, WLANs (Wireless Local Area Network) have had an exponential growth, with most public spaces providing access to wireless hotspots. This has resulted in an increasing interest in location services. One of the main features offered by these technologies is that they allow users to receive services based on their geographic location. Some examples include interactive maps of malls and museums, college campus guided maps, patient monitoring systems in hospitals and/or nursing homes for the elderly [1]. The use of GPS is not possible, since it does not work in enclosed spaces because it requires a direct and unobstructed line of sight between the receiver and a minimum of three satellites [2] [3].

Calculating the relative position of a mobile device is the process through which information is obtained about the position of such device in relation to landmarks on a predefined space [4]. The techniques used to calculate a position in an enclosed space are classified, based on the sensor technology used, in: Time of Arrival (ToA), Angle of Arrival (AoA), and Received Signal Strength Indicator (RSSI). In the case of RSSI, there are two main models used to calculate the position of the MD: signal propagation model and empirical model. The former is based on the application of mathematical models that determine signal behavior and its degradation while it propagates. The empirical model calculates the position based on parameters that are stored in a database. Within the latter, the algorithm that is most widely used for calculating positions is the Fingerprint algorithm [5].

In this paper, we present the development of a positioning system based on the Fingerprinting technique that introduces various metrics to calculate the position of the MD. Also, certain improvements are implemented, such as increased accuracy and lower error in relation to position, and use metrics for the different variations of the Fingerprint database are presented.

This paper is organized as follows: Section 2 describes the state of the art, Section 3 presents the technique used, Section 4 details the experiments carried out, Section 5 analyzes the metrics used and the results obtained during positioning, and Section 6 analyzes variation use metrics for the fingerprint database. Finally, Section 7 describes the conclusions and future work.

2. STATE OF THE ART

Several developments have been carried out in the area of device location using RSSI. One of the first approaches in this context is the RADAR system [6], which combines two methods – an empirical model using Fingerprint and a mathematical model that takes into account signal propagation. It obtains a mean accuracy within 2-3 meters. In [7], the authors of the RADAR system implement improvements, such as multipath and interference, with the purpose of analyzing and reducing problems that are inherent to the nature of the signal. They also analyze environmental changes during the experimental phase. The accuracy obtained is less than 2 meters.

In 2002, the Fingerprint technique is used to calculate the position of a MD in combination with a probabilistic model [1], which uses a probabilistic technique that calculates the probability that the device is at a certain position within the radio map, based on Bayes algorithm. The system achieves an accuracy of 1.5-3 meters.

In 2003, the LEASE system [8] offers a framework based on the Fingerprint technique that achieves an accuracy of 2.1 m. The Ekahau commercial system [9] is a positioning software that consists in an administrator, a server, and a client. The administrator records the strength of the signal at the access points (APs), and creates the positioning database. The client sends requests to the server, and the server responds by calculating the position. System accuracy is 2-3 meters. The Aeroscout commercial system [10] has 4 main components: positioning tags, mobile devices, a calculation engine, and a receiver. To calculate the position, a hybrid technique is used combining the ToA technique and RSSI processing. The system achieves a relative accuracy of 2-3 meters.

In 2003, a signal propagation model [11] is used; this model obtains a third-degree polynomial regression to calculate the position of the MD, achieving an accuracy of 1-3 meters. In 2004, the Fingerprint method is used with two schemes [12]: the first one is based on changing the number of sampling points using the Euclidean distance as metric, while the second one uses the enhanced Euclidean distance and the standard deviation of sampling sets as weights. The system achieves a relative error of 2-3 meters.

In 2006, a system is presented that uses Fingerprint and the Euclidean distance method with an enhancement algorithm that uses fuzzy logic [13]; an accuracy of 4.47 m is obtained first and then, with fuzzy logic, it is improved to 3 m.

In 2008, a calculation method based on a clustering algorithm is used [14], called Cluster Filtered KNN (K-Nearest Neighbors); results are then compared to those obtained with the KNN calculation method. The system achieves an accuracy of 2 meters.

In 2008, a positioning system is implemented that is based on a multilayer neural perceptron [15] that uses AP RSSI values as metric, and then compares the results obtained with the WKNN and traditional KNN algorithms. The system achieves a relative accuracy of 2 meters.

In 2010, a technique called PKNN (Predicted KNN) [16], based on the historical positioning data of a mobile device, is used. The technique achieves a 33-percent improvement in relation to the traditional KNN algorithm, with an accuracy of 1.3 meters.

Also in 2010, the Fingerprint technique is used in combination with a propagation model [17] to calculate the position of a MD. The propagation model is defined based on the physical features of the environment, calculating the Wall Attenuation Factor (WAF). The proposed system uses a filter to improve accuracy, and achieves an error that is below 1.8 meters.

In 2011, a Fingerprint positioning system is presented that is based on the LDPL (Long Distance Path Loss) propagation model to calculate the distance between the pattern vector and the vectors stored in the database, without using the Euclidean distance [18]. The system introduces an improvement of the order of 10% in relation to the KNN algorithm, with an error of the order of 2 meters.

3. FINGERPRINT TECHNIQUE

The Fingerprint technique is based on measuring the intensity of the signal received, RSSI (Received Signal Strength Indicator), by a MD from various APs on a given location and applying to those intensities an algorithm that allows calculating the position of the MD. This technique requires training, where each of the power vectors is sampled. [19].

First, a radio map [20] must be designed, which is a pattern map containing the specific positions within an enclosed space and an RSSI strengths vector with all the strengths or intensities of the APs reached at each position. Creating a radio map involves:

- At each position, signal strength (RSSI) values are shown, putting together a strength vector whose dimension depends on the number of visible APs.

- For each sector in the area that can receive the signal from N access points, an RSSI vector is obtained from each AP.

The coordinates that represent the location of each sampling point and the strength vectors obtained are stored in a database.

To calculate the position of the MD, the values from all visible APs are captured from that same position. The acquired values are then compared with the values stored in the database to obtain the coordinates that represent the location of the device [21].

Calculation algorithms correlate the values obtained between the location information and the Fingerprint database to determine the relative position of the MD. There are two classes of algorithms: probabilistic and deterministic.

The most widely known deterministic method is that of the "nearest neighbor." In this case, the strength vector is compared with the vectors stored in the database in order to determine their proximity, which is calculated based on the distance between vectors. The most commonly used distance metric is the Euclidian distance. The vector that is closest to the input value determines the location of the MD.

The main disadvantage of this method is that it requires calculating the distance from the pattern vector to each of the vectors stored in the database, which increases computation time. In this sense, the contribution of this work is mainly focused on reducing this time by adding a pre-processing stage for the vectors stored in the Fingerprint database and then working with the simplified version of the database.

4. EXPERIMENTS AND DESIGN OF THE DATABASE

Even though experiments are done all over the campus of the National University of the Center of the Province of Buenos Aires, the data analyzed in this paper are only those collected in the fellowship area of the INTIA/INCA research institute. The area spans over approximately 36 m2. To carry out measurements and design the radio map, the area is divided along a coordinate axis (row, column) (Figure 1); each region of the map has a separation of 1 meter from the previous point.

The Fingerprint database was built with 36 sampling points, and 6 points were selected for the calculation phase.

Figure 1. Coordinates location in the map

Data were captured with IWLIST on Ubuntu 8.04, and were processed with WifiPos, a tool developed in C language in a Windows environment. The data capture process for building the radio map (Figure 1) is as follows:

1. Positioning of the MD at a coordinates point in the map.

2. Scanning and capturing RSSI values for 180 seconds to stabilize sampling signal.

3. Scanning, capturing and storing the RSSI and SSID values corresponding to the signal of the various APs that are within reach at that position for 90 seconds.

4. Moving the device to the next coordinates point in the map, and repeating step (1) if it is not the last location.

Once the data are sampled, the Fingerprint database is built to then apply the KNN algorithm during the calculation phase. The goal is to design a database that allows minimizing calculation time by carrying out a statistical and mathematical pre-processing of the data.

Given a set of RSSI values obtained from each AP, the following calculations are performed at each coordinates point:

- **Maximum Values:** The maximum value of RSSI reached by the AP is obtained.

- **Minimum Values:** The minimum value of RSSI reached by the AP is obtained.

- **Average:** The arithmetic average of the values obtained is calculated.

- **Mode:** The mode is calculated with the total of the samples obtained.

- **Interquartile Pair:** data are sorted in ascending order and then divided in 4 sets with equal number of elements. Both end quartiles are removed, and the following are calculated from the inner quartiles:

    - average

    - arithmetic mean

    - **Interquartile Pair Mean Value:** The mean value of the interquartile pair is obtained.

This way, instead of working with the entire database and comparing against all vectors for each position of the MD, it is possible to compare only against the results of the vectors that were processed.

5. RESULT ANALYSIS

During the calculation phase, the MD is positioned at a point in the map (Figure 1) and "pattern vectors" are obtained for the strengths of the visible APs for a period of 90 seconds. Then, with these vectors and each of the strength vectors previously stored (maximum values, minimum values, average, mode, and interquartile pair mean value), the Euclidean distance is calculated by obtaining the distances to each coordinates point.

The position of the MD is determined as the pair of coordinates with the shortest distance between the training set (Fingerprint database) and the input data pattern.

Multiple tests are carried out and the analysis of errors is sorted as follows, considering a total of 89 calculation vectors obtained at each calculation point.

First, a quantitative analysis of the hits obtained with each technique is performed.

Figure 2 shows the analysis corresponding to position 6.3 and the mode technique. It can be observed that there are 71 matching vectors for which the MD is located at point 6.3, with a hit percentage of 80%. The 18 remaining vectors yield points 2.6, 6.6 and 1.6 as their results.
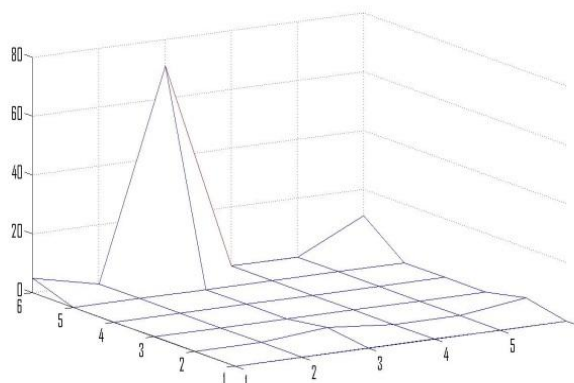

Figure 2. Surface chart for point 6.3

Figure 3 shows a distribution of the positions calculated for the vectors obtained at position 2.3. The percentage of hits for the different techniques are: mode 85%, quartiles mode 69%, average 79%, and mean quartiles value 69%. At points 6.2, 1.2, 6.3 and 6.6 there are false positives which, in the worst of cases, represent a 15% with the "mean quartiles value" technique.
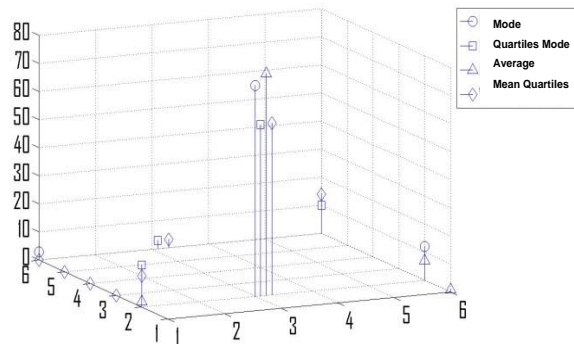

Figure 3. Vector distribution chart for point 2.3

Going back to Figure 1, it should be noted that there are no walls adjacent to the capture point that could attenuate the strength and reduce the value of RSSI.

Secondly, considering the actual position of the device and the position calculated by each of the techniques, it is possible to determine the maximum error to ensure an accuracy of 95% by implementing the techniques.

Figure 4 presents the maximum errors (measured in meters) obtained with each technique for the central positions of the radio map. At both points, using the average and mode techniques, a maximum error of 2 meters is obtained, while with the quartiles mode technique this error is 4 meters.
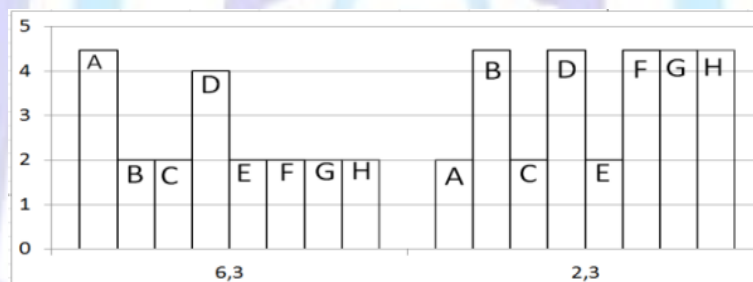
Figure 4. Surface chart for point 6.3. A: Maximum, B: Minimum, C: Mode, D: Quartiles Mode, E: Average, F: Quartiles Average, G: Mean Value, H: Quartiles Mean Value.

6. MODIFIED FINGERPRINT DATABASE USE ANALYSIS

The WiFiPos tool allows calculating the position of the MD by calculating the Euclidean distance between the vector obtained when positioning the device and the pre-processed values of the different techniques presented.

If there are 100 vectors with RSSI values for the various APs, the time required to position a MD is reduced by a factor of 100x. This is because the data collected are pre-processed offline during the previous stage.

With this proposal, the time required to position a MD can be significantly reduced when the system is working online.

7. CONCLUSIONS AND FUTURE WORK

A comparative analysis of applying some variations in the composition of the Fingerprint database is presented. The work area was very small and with many internal obstacles. At this point of the project, we are capable of positioning the MD with an accuracy of two meters if away from walls, and four meters if it is close to a wall.

The use of a summary Fingerprint database allows reducing computational load and maintaining positioning efficacy.

In the future, efficacy will be analyzed in large indoor and outdoor areas. Tests will be carried out using neural networks and fuzzy clustering on the same data.

ACKNOWLEDGMENTS

REFERENCES

[1] A. M. Ladd, K. E. Bekris, G. Marceau, A. Rudys, L. E. Kavraki, and D. S. Wallach, Robotics-based location sensing using wireless Ethernet. ACM International Conference on Mobile Computing and Networking (MOBICOM'02), New York, 2002.

[2] G. M. Djuknic and R. E. Richton, Geolocation and assisted GPS, IEEE Computer. Vol. 34, No. 2, pp: 123-125, 2001.

[3] J. Toloza, N. Acosta, A. de Giusti. An approach to determine magnitude and direction error in GPS system. Asian Journal of computer science And Information Technology. Vol. 2, No. 9, Pp: 1-5. 2012.

[4] T. S. Rappaport, J. H. Reed, and B. D. Woerner, Position location using wireless communications on highways of the future, IEEE Communic. Vol. 34, No., pp: 33-41, 1996.

[5] K. Pahlavan, X. Li, and J. P. Makela, Indoor geolocation science and technology, IEEE Communications Magazine., Vol. 40, No. 2, pp: 112-118, 2002.

[6] P. Bahl and V. N. Padmanabhan. RADAR: An in-building RF based user location and tracking system, IEEE Infocom 2000. Vol.2, No. 1, pp: 775-784. 2000.

[7] Paramvir Bahl, Venkata N. Padmanabhan, Anand Balachandran. Enhancements to the RADAR User Location and Tracking System, Microsoft Academic Research. Technical Report. Vol 1, No. 1, pp: 1-13. 2000.

[8] P. Krishnan, A. S. Krishnakumar, Wen-Hua Ju, Colin Mallows, Sachin Ganu. A System for LEASE: Location Estimation Assisted by Stationary Emitters for Indoor RF Wireless Networks. IEEE Infocom 2004. Vol.1, No. 1, pp: 1-11. 2004.

[9] Ekahau, Ekahau positioning engine 2.0; 802.11 based wireless LAN positioning system, Ekahau Technology, Internal Report, www.eukahau.com, 2012.

[10] Aeroscout, aeroscout positioning engine 2.0, aeroscout Technology, Internal Report, www.aeroscout.com, 2012.

[11] Y. Wang, X. Jia, and H.K Lee. An indoors wireless positioning system based on wireless local area network infrastructure. 6th Int. Symp. on Satellite Navigation Technology Including Mobile Positioning & Location Services, Melbourne, 2003.

[12] Kamol Kaemarungsi and Prashant Krishnamurthy, Modeling of Indoor Positioning Systems Based on Location Fingerprinting. IEEE Infocom 2004. Vol.1, No. 1, pp: 1-11. 2004

[13] A. Teuber, B. Eissfeller, WLAN indoor positioning based on Euclidean distances and fuzzy logic, Proceedings of the 3rd Workshop on Positioning, Navigation and Communication (WPNC'06), Munich, Germany, 2006.

[14] Jun MA, Xuansong LI, Xianping TAO and Jian LU. Cluster Filtered KNN: WLAN-Based Indoor Positioning Scheme. IEEE Infocom. Vol.1, No. 1, pp: 1- 8. 2008.

[15] Shih-Hau Fang, Tsung-Nan Lin. Indoor Location System Based on Discriminant-Adaptive Neural Network in IEEE 802.11 Environments. IEEE transactions on neural networks, vol. 19, no. 11. pp: 1973-1978. 2008.

[16] Shahrzad Khodayari, Mina Maleki, Elham Hamedi. A RSS-based Fingerprinting Method for Positioning based on Historical Data. Symposium on , Ottawa, 2010.

[17] Yuhong Liu and Yaokuan Wang. A Novel Positioning Method for WLAN Based on Propagation Modeling. Progress in Informatics and Computing (PIC), IEEE International Conference on Shanghai. Vol.1, No. 1, pp: 397 - 401. 2010.

[18] Sukhoon Jung, Choon-oh Lee, and Dongsoo Han. Wi-Fi Fingerprint-based Approaches Following Log-Distance Path LossModel for Indoor Positioning. Intelligent Radio for Future Personal Terminals (IMWS-IRFPT), IEEE MTT-S International Microwave Workshop Series on. Daejeon. 2011.

[19] Brachmann, F. A comparative analysis of standardized technologies for providing indoor geolocation functionality, Symposium on Computational Intelligence and Informatics (13th CINTI), 2012 IEEE, Hungary, Budapest 2012.

[20] M. A. Youssef, A. Agrawala, and A. U. Shankar, WLAN location determination via clustering and probability distributions, in Proc. IEEE International Conference on Pervasive Computing and Communications, Fort Worth Texas, New York, 2003

[21] P. Enge and P. Misra, Special issue on GPS: The global positioning system, Proceedings of the IEEE. Pp: 3-172. 1999.

Author' biography with Photo

P.h.D. Héctor Nelson Acosta was graduated at the Universidad Nacional del Centro de la Provincia de Buenos Aires (Argentine), and got the Ph.D. degree at the Autonomous University of Madrid (Madrid, Spain). He is the director of the Instituto de Investigación en Tecnología Informática Avanzada (INTIA) at the Universidad Nacional del Centro de la Provincia de Buenos Aires since the 1998. He is working as a proffesor since 1993 at the Computer and Systems Department, at the Universidad.

P.h.D. Juan Manuel Toloza was graduated in march 2013 Universidad Nacional de La Plata (Argentine). He have a Post-Doctoral Fellow of Consejo Nacional de Investigaciones Científicas y Técnicas since 2013. He works at the Instituto de Investigación en Tecnología Informática Avanzada at the Universidad Nacional del Centro de la Provincia de Buenos Aires. His current research insterest includes signal procesing, data analysis from sensors, robotics and accuracy GPS. He is teacher assistant in Computer`s Architecture and Digital Techniques.

Degree in information systems Carlos Antonio Kornuta was graduated in march 2012 at the Universidad Nacional de Misiones (Argentine). Doing his P.h.D. computational and industrial mathematics at the Universidad Nacional del Centro de la Provincia de Buenos Aires (Argentine). He have a scholarship of the Consejo Nacional de Investigaciones Científicas y Técnicas since 2012. His current research insterest includes Telecommunications, wireless positioning systems, security informatics, computer forensics. He is teacher assistant in data communications and Computer`s Architecture and Digital Techniques .