Comparison of WiFi-based indoor positioning techniques

Szabolcs Karsai, Zsolt Tóth

University of Miskolc
`karsai1@iit.uni-miskolc.hu`
`tothzs@iit.uni-miskolc.hu`

Abstract

Indoor positioning systems is an actively researched area of computer science. Various techniques, methods and applications were developed in the last decade. This paper gives a brief comparison of trilateriation and fingerprinting techniques. Received signal strength of WiFi access points were used during the tests. Standard, genetic algorithm and simulated annealing based trilateriation methods were implemented. Fingerprinting methods were based on nearest neighbor approach. Correlation coefficient and euclidean distance were used to calculate distance of the actual and the stored measurements. Experimental results showed that approximately 3–4 meter accuracy can be achieved with the tested methods.

Keywords: Indoor positioning, WiFi, trilateration, fingerprinting

MSC: AMS classification numbers

1. Introduction

Finding objects or people is crucial for various applications such as, asset tracking [11] or guidance [6]. Global and Indoor Positioning Systems are usually distinguished. The GPS is one of the most widely used Global Positioning System these days which was designed to determine the position anywhere on the Earth. GPS cannot be used in buildings due to its line of sight requirement. Indoor positioning systems aim to locate objects or people in buildings and they have been actively researched for over a decade.

Various technologies are used for indoor positioning. The first systems were based on infrared and ultrasonic technologies which can require special hardware

development or have significant installation cost. Recent developments are based on RFID or WiFi technologies. Many access control systems use RFID tags for identification due to the low cost of passive tags and simple implementation. In localization systems the users usually have readers and the tags are placed in the building. The user only read the nearby RFID tags and the position is determined by a central server which requires communication. The popularity of the WiFi based localization system can be explained with the low installation cost. These systems use the WLAN for both localization and communication purposes. These systems are cheap to install because the WLAN networks are already established in many cases and the smartphones support WiFi communication.

This paper gives a brief comparison of two basic positioning methods: trilateriation and fingerprinting. These methods are based on received signal strengths information of WiFi access points. The standard trilateriation method is based on distance measurement thus the accuracy of the method depends on the distance calculation. Heuristic methods was used to handle the error of the distance calculation. Fingerprinting technique has two phases: measurement and positioning. During the measurement phase, the characteristics of the received signals are measured in known positions and the results are stored in a database. In the positioning phase, the system compares the actual and the stored measurements in order to determine the current position. Correlation coefficient and Euclidean distance were used to calculate the similarity of the current and the stored measurements.

2. Related work

Positioning and localization is an actively researched area of computer science. The existing solutions can be categorized based on various aspects, such as used technology, accuracy or environment. The indoor and outdoor positioning systems are usually distinguished. Although GPS may be the most well–known and most widely used positioning system, there are other solutions. There are systems for cell phone localization [12, 2] for example the Enhanced 911 [13].

Indoor Positioning Systems become popular about 2000. Since that, numerous indoor positioning systems have been developed [9, 8, 3]. The challenges of these systems are usually caused by the physical structure of the buildings. The existing systems are based on various technologies and principles. This paper compares only received signal strength information based techniques.

LANDMARC [10, 7] is based on RFID technology. The tags are placed in the room and the users carry the readers. The readers are connected to the server via WLAN. The distance between the tags and the reader is not determined precisely. Each reader returns with the detected RFID tags within a given range. Reference tags are used to deal with the dynamical changes of the environment. A k nearest neighbor approach is used for positioning.

The RADAR system [1] is based on received signal strengths which is transmitted by the mobile device. The WLAN network is used for both measurement and communication. The localization service is implemented in the base station which

receives the signals from the mobile devices. The RADAR system has *off–line* and *on–line* phases. In the *off–line* phase, the signal strengths are measured in known positions. In the *on–line* phase, the system is used for positioning. The RADAR system uses both triangulation and signal propagation model.

The Horus system [16, 15] also based on the received signal strength information of WiFi access points. While in the RADAR system the measurements are done by the base stations, in Horus the mobile devices measures the signal strengths. It also has an *off–line* and an *on–line* phases. In the *off–line* phase the measurements are stored in the database which is called radio map. In the *on–line* phase the mobile devices measure the signal strengths then send the measurements to the server. The server use a clustering method to determine the position of the device.

3. Examined techniques

Trilateriation and fingerprinting are the two major techniques of indoor positioning systems. The existing solutions are usually based on one of these techniques. The tested algorithms are detailed below.

3.1. Trilateration

Triangulation and trilateration techniques are both based on the geometrical properties of the triangle. In triangulation, the angles are used to determine the location. Trilateration technique is based on the distance of the transmitter and the receiver. In this paper only trilateration based methods are analyzed.

3.1.1. Distance approximation

The received signal strength was used to calculate the distance of the transmitters and the mobile device. This method was chosen because it does not require additional hardware components and it is easy to implement. On the other hand it can limit the accuracy of the system [5]. Free space path loss was used to calculate the distance. The calculation is based on the signal attenuation and the parameters of the environment were given.

3.1.2. Standard method

The standard trilateriation is based on vector operations. The position of the access points are fix and known thus they can be represented by vectors $(ap_{i,x}, ap_{i,y}, ap_{i,z})$ where ap_i denotes ith the access point and x, y, z stand for the coordinates. Let $l = (l_x, l_y, l_z)$ denote the current location and the distance between the access point and the device is calculated with Euclidean distance. In a n dimensional space $n + 1$ points are required to determine the location of a point. Figure 1 shows both the ideal and the real cases. The ideal case is denoted by the solid line. In the ideal case the point can be determined exactly by the distances from the reference points because the distances are exactly determined. In real cases

the distance calculation is not accurate thus the position cannot be determined exactly. Moreover the different access points can have different errors.
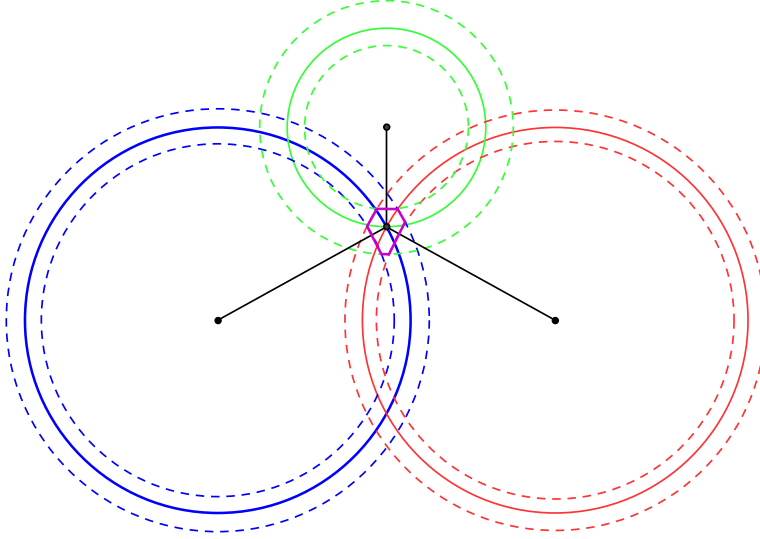

Figure 1: Ambiguity of Trilateriation

3.1.3. Heuristic approaches

The above reviewed standard method has some drawbacks. Firstly, it requires at least four reference points to determine the location. Secondly, it assumes that the distances are accurate, but in the experiments the distance between the device and the access point is approximated so it can contain error.

To deal with these limits the localization problem was transformed into an optimization task. Let us assume that, there are n available reference access points in the current location. The distances can be approximated $\hat{d}(l, ap_i)$. The object function is defined as the square of the difference of the error of the approximated and calculated differences. Equation 3.1 shows the object function. Simulated annealing [14] and genetic algorithm [4] were used to solve the optimization task.

$$E(l_x, l_y, l_z) = \sum_{i=0}^{n} (d(l, ap_i) - \hat{d}(l, ap_i))^2 \to min \tag{3.1}$$

3.2. Fingerprinting

Fingerprinting techniques usually have two phases: an *off–line* teaching phase and an *on–line* positioning phase. A fingerprinting database is used to store measurement of received signal strengths information in known position. The database can

contain one sample or even a time series [15] for each access point. In the *off–line* phase the database is filled. This phase can be costly and it has to be repeated if the environment changes significantly. The accuracy of the system depends on the stored measurements. In the *on–line* phase the location is estimated based on the received signal strengths. Both deterministic and probabilistic estimations can be used. Various pattern recognition techniques are used, such as k–nearest–neighbor, artificial neural network or support vector machines. The tested algorithm was based on the k nearest neighbor approach.

3.2.1. Distance calculation

Correlation coefficient and Euclidean distance were used to calculate the distance of the actual and the received signal strength informations. The database is considered as two matrices. The matrix of the probe points $\mathbf{P}$ which contains the coordinates of the measurements. The measured values are stored in the $\mathbf{M}$ measurement matrix where the columns denote the access points and the rows are the measurements. Hence a measurement is represented by a vector $\mathbf{m}$ which contains the signal strength of the given access point or zero if it is unavailable. These measurements and the probe points are paired.

In the current location the measurements are contained in the $\mathbf{m}_c$ vector. The location is determined based on the similarity of the current measurement $\mathbf{m}_c$ with the stored measurements $\mathbf{M}$. Correlation coefficient was used to determine the similarity of two measurements and it is shown in Equation 3.2 where $\mathbf{m}_i(k)$ is the kth value of the $\mathbf{m}_i$ vector, $\overline{\mathbf{m}_i}$ is the average and $s_{\mathbf{m}_i}$ is the deviation of the $\mathbf{m}_i$ measurement. Let $\mathbf{c}$ denote a vector of correlation coefficients. Equation 3.3 shows the estimation of the current location. The method has a parameter which is the τ threshold value of correlation coefficient. If the correlation coefficient does not exceed the τ threshold, then it is considered zero thus the point is omitted.

$$cor(\mathbf{m}_i, \mathbf{m}_j) = \frac{\sum_{k=0}^{n}(\mathbf{m}_i(k) - \overline{\mathbf{m}_i})(\mathbf{m}_j(k) - \overline{\mathbf{m}_j})}{s_{\mathbf{m}_i} s_{\mathbf{m}_j}} \qquad (3.2)$$

$$\mathbf{l}_c = \frac{\mathbf{cP}}{\overline{\mathbf{c}}} \qquad (3.3)$$

Euclidean distance was used to determine the most similar neighbors. The measurements determine a point in a n dimensional space where n is the number of the access points. Each measurement determines a point and the similarity is determined base on the distance of the points.

4. Experimental results

A web service and a mobile client were developed to measure the accuracy of the above detailed methods. Positioning interface was defined on the server side. The tested methods implemented this interface and they were realized as web services. The mobile client was an Android application and it used HTTP requests and

JSON messages to communicate with the server. Both the server and the client side were implemented in Java.

The methods were evaluated from the point of view of accuracy. The accuracy was calculated as the difference of the determined and the actual position in meters. The measurements were performed in the same positions with the different methods. The tested methods were characterized by the minimum, maximum, average, variance and the 75th percentile of the error. Best and worst cases are represented by the minimum and the maximum error. Although average and variance are widely used to describe measurements, their interpretation can be difficult due to the outlying values. Thus the 75th percentile is used to determine the precision of the algorithms.

4.1. Test environment

The evaluated methods were tested in the first floor of the Institute of Information Science at the University of Miskolc. The building is approximately 48 meters long and 28 meters wide. There are two corridors and a hall on the floor. Two access points were installed at the end of the corridors and another two were placed in the hall for the measurements.

4.2. Trilateriation

Table 1 shows the error of the measurements with trilateriation techniques. The highest accuracy was achieved with the standard trilateriation method. The genetic algorithm based method performed poorly in our experiments but it was tested with only a couple setups. The performance of the genetic algorithm based method could be improved with the tuning. Simulated annealing based heuristic method performed well. The best average accuracy was around 3.6 meters and the 75 per cent of the measured errors were below then 4.75 meters. Experimental results shows that the standard method can be enhanced with heuristic searches which could improved the accuracy of the system.

Method	Min	Max	Avg	Var	75th
Standard	0.7083	10.0874	4.4716	2.7488	5.0971
Genetic Algorithm	0.9526	12.3559	5.0436	3.7166	5.3973
Simulated Annealing	0.8409	7.4148	3.621	2.3094	4.75

Table 1: Errors of Trilateriation methods

4.3. Fingerprinting

Tested fingerprinting method uses correlation coefficient and euclidean distance to determine the similarity of the current and the stored measurements. The tests were performed with three different setups of the correlation threshold and one

setup for the euclidean distance. The experimental results are summed up in Table 2. When the distance was calculated with the correlation coefficient, the accuracy of the algorithm increased with the threshold. On the other hand the error exceeded the 5 meters during the tests. Approximately 3 meters accuracy was achieved when euclidean distance was used.

Method	Min	Max	Avg	Var	75th
0.7	3.3178	31.009	9.0999	8.687	8.2954
0.8	1.55	31.3002	6.6689	8.7572	5.3044
0.9	1.58	8.9208	4.4432	2.0059	5.13
Euclidean	0.34	18	4.1415	5.1047	3.0192

Table 2: Errors of Fingerprinting method with different thresholds

5. Conclusion

Trilateriation and fingerprinting based indoor positioning techniques were reviewed in this paper. Only WiFi signal strengths were used during the measurements. The WLAN network was used for both positioning and communication. Three different trilateriation methods were tested. The standard method was enhanced with genetic algorithm and simulated annealing. The analysis of the fingerprinting technique was only focused on nearest neighbor search based solutions. Correlation coefficient and Euclidean distance were used to determine the similarity of the stored and the current measurements.

The tested methods were implemented as web services and the client was an Android device. Experimental results were compared from the point of view of accuracy. The error of the positioning methods was calculated as the difference of the actual and the determined positions. The methods were evaluated based on the minimum, maximum average, variance and 75th percentile of the errors. Approximately 3-5 meters accuracy was achieved with the tested methods.

References

[1] Paramvir Bahl and Venkata N Padmanabhan. Radar: An in-building rf-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 775–784. IEEE, 2000.

[2] Mike Y Chen, Timothy Sohn, Dmitri Chmelev, Dirk Haehnel, Jeffrey Hightower, Jeff Hughes, Anthony LaMarca, Fred Potter, Ian Smith, and Alex Varshavsky. Practical metropolitan-scale positioning for gsm phones. In *UbiComp 2006: Ubiquitous Computing*, pages 225–242. Springer, 2006.

[3] Kevin Curran, Eoghan Furey, Tom Lunney, Jose Santos, Derek Woods, and Aiden McCaughey. An evaluation of indoor location determination technologies. *Journal of Location Based Services*, 5(2):61–78, 2011.

[4] Lawrence Davis et al. *Handbook of genetic algorithms*. Van Nostrand Reinhold New York, 1991.

[5] Eiman Elnahrawy, Xiaoyan Li, and Richard P Martin. The limits of localization using signal strength: A comparative study. In *Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on*, pages 406–414. IEEE, 2004.

[6] Giuseppe Ghiani, Fabio Paternò, Carmen Santoro, and Lucio Davide Spano. Ubicicero: A location-aware, multi-device museum guide. *Interacting with Computers*, 21(4):288–303, 2009.

[7] Kyuwon Han and Sung Ho Cho. Advanced landmarc with adaptive k-nearest algorithm for rfid location system. In *Proceedings of the 2nd IEEE International Conference on Network Infrastructure and Digital Content (ICNIDC'10)*, pages 595–598, 2010.

[8] Hakan Koyuncu and Shuang Hua Yang. A survey of indoor positioning and object locating systems. *IJCSNS International Journal of Computer Science and Network Security*, 10(5):121–128, 2010.

[9] Hui Liu, Houshang Darabi, Pat Banerjee, and Jing Liu. Survey of wireless indoor positioning techniques and systems. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(6):1067–1080, 2007.

[10] Lionel M Ni, Yunhao Liu, Yiu Cho Lau, and Abhishek P Patil. Landmarc: indoor location sensing using active rfid. *Wireless networks*, 10(6):701–710, 2004.

[11] Asil Oztekin, Foad M Pajouh, Dursun Delen, and Leva K Swim. An rfid network design methodology for asset tracking in healthcare. *Decision Support Systems*, 49(1):100–109, 2010.

[12] Jeongyeup Paek, Joongheon Kim, and Ramesh Govindan. Energy-efficient rate-adaptive gps-based positioning for smartphones. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*, pages 299–314. ACM, 2010.

[13] Jeffrey H Reed, Kevin J Krizman, Brian D Woerner, and Theodore S Rappaport. An overview of the challenges and progress in meeting the e-911 requirement for location service. *IEEE Communications Magazine*, 36(4):30–37, 1998.

[14] Peter JM Van Laarhoven and Emile HL Aarts. *Simulated annealing*. Springer, 1987.

[15] Moustafa Youssef and Ashok Agrawala. Handling samples correlation in the horus system. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 2, pages 1023–1031. IEEE, 2004.

[16] Moustafa A Youssef, Ashok Agrawala, and A Udaya Shankar. Wlan location determination via clustering and probability distributions. In *Pervasive Computing and Communications, 2003.(PerCom 2003). Proceedings of the First IEEE International Conference on*, pages 143–150. IEEE, 2003.