WIFI-Based Indoor Positioning System

Qinyang Li
Department of Computer Science
The George Washington University
Washington D.C 20052
Qyl121@gwmail.gwu.edu

Qiuyu Zhang
Department of Computer Science
the George Washington University
Washington D.C 20052
qiuyu0805@gwmail.gwu.edu

Abstract

With the rapid development of Internet technology and WIFI coverage increasing year by year, people's demand for indoor WIFI positioning has also increased. Due to severe signal attenuation and multipath effects, general-purpose outdoor positioning facilities (GPS) do not work effectively in buildings. Although there are other technologies, such as Bluetooth, RFID, mobile phone base stations, etc., which can also be used to achieve indoor positioning, but they are not everywhere like WIFI. Hence, we plan to design a WIFI-based Indoor positioning system to solve such a problem. The important factor which could influence the positioning accuracy is the distances among access points (APs) to terminal. Firstly, we accept the RSSI (received signal strength indicator) from the WIFI around us. However, RSSI is affected by many factors such as indoor temperature, humidity, multi-path effect, etc., and data preprocessing is required. After we receive RSSI data, we use Gaussian fitting and Kalman filtering to simply preprocess the data, and the distance between terminal and AP is obtained indoor signal propagation model. Next, we can use Trilateration algorithm to obtain WIFI locations and positioning terminal. Source code available at https://github.com/ramborz/WIFI-positioning-

1 Introduction

WIFI is widely used in many large or small buildings such as houses, airports, shopping malls, libraries, etc. Generally, a WIFI system consists of fixed access points (APs), which are deployed indoors in some locations where are easy to install. The system administrator usually knows the locations of these APs. WIFI fingerprint positioning has been widely used in the indoor positioning field. The weighted K-nearest neighbor (WKNN) algorithm is one of the most widely used deterministic algorithms. We plan to use RSSI based algorithm as the algorithm of WIFI indoor positioning system.

RSSI localization techniques are based on measuring signal strength from a client device to several different access points, and then combining this information with a propagation model to determine the distance between the client device and the access points. Trilateration (sometimes called Multilateration) techniques can be used to calculate the estimated client device position relative to the known position of access points [2].

1.1 Project Purpose

This is group project for CSCI 6221 at 2020 spring class year. The purpose of this project is to become familiar with one or more software paradigms and to develop and implement a software system/application. The instructor for this class is professor Yih-Feng Hwang, yhwang1@gwu.edu

1.2 Relation to machine learning

During the testing phase of this project, the triangulation algorithm is too accurate to work with raw data we collected which contain unseen error in different type of obstacle. Those errors bring a lot of uncertainty to our model of relationship between WIFI signal strength and distance. To solve this problem, the KNN algorithm are used to work on raw data to improve the overall accuracy of our project.

1.3 Related work

Initial online research shows that most search result are companies selling completed product to the public. At same time, some website offers basic concepts about WIFI positioning. There are 4 different way to positioning using WIFI. In this project, the RSSI and Multilateration are used. Another method including Fingerprinting, Angle of arrival, and Time of flight. As for scholarly reviewed research, there are some PHD paper researching about using K-clustering on fingerprinting method which have much higher accuracy than this project [6]. On the other hand, this project is focused on bring this technology to general public use rather than pinpoint accuracy.

2 Rationale

We should be able to access the basic information about the routers. We assume the signal strength can be accessed as a digital value. On the other hand, certain model of devices could display calculated distance between different devices which we need to further investigate.

With digitized signal strength from the routers, we will be able to calculate the distance between the router and the device using certain algorithm. From here there are two functions available.

The first function should use the position of the device as origin of a 3D empty workspace and mark the position of any known router in this workspace. We consider using Euler angles and transfer it into cartesian coordinate system for further use.

The second function should use available router to identify certain known workspace from a workspace database. Then use the information of this workspace to locate the terminal.

3 Issues Encountered

How to Transform WIFI signal into workable distance

How to calculate coordinate of WIFI based on distance

How to properly store collected data for different environment for future use

How to improve positioning accuracy using machine learning

How to implement the program for general use

3.1 RSSI to Distance

The positioning technology based on RSSI has a significantly advantage which is lower costly in surveying or mapping preparations works. The positioning principle is firstly constructing the mapping between RSSI and distance, and then collecting real-time RSS values and calculating the distance according to function (1), and then estimating the coordinate by positioning algorithm such as Trilateration Algorithm.

$$PL(d) = PL(d_0) + 10\eta \log(d / d_0) + X_\sigma \qquad (1)$$

In this function (1), d is the distance of AP and terminal, d0 is reference distance, is distance factor which depends on specific propagation environment, X is zero-mean Gaussian random variable; PL(d) is the RSSI value at position d, PL(d0) is the RSSI value at the reference distance (usually is 1m).

In order to get the distance d, we can transfer function (1) to function (2):

$$\mathrm{d} = 10^{\frac{PL(d0) - PL(d) + X\sigma}{10\eta}} \qquad (2)$$

Because the limitation of space and equipment, it is easy to get PL($d0$) and PL(d) value, but the problem is that how to determine the value of η. One solution is that we test several times and compare the real-time distance with the distance given by the function (1), we estimate the value of n which would compute relatively accurate distance value.

After we get the distance from all the WIFI around us in one spot, we record these data into database and move to another spot. We can use the trilateration algorithm to measure the location(coordinate) of all WIFI, knowing that we have scanned 4 spots. As long as we have the coordinate of WIFI around us, we choose 4 WIFI which have the most powerful signal strength to locate ourselves by using trilateration algorithm again.

3.2 RSSI Processing

Kalman filtering is an algorithm that uses the linear system state equation to observe the data through the system input and output to optimally estimate the system state. Since the observation data includes the influence of noise and interference in the system, the optimal estimation can also be regarded as a filtering process.

Even if the terminal remains stationary, the detected signal strength from the same AP will continue to fluctuate up and down. Kalman filtering can be used to dynamic data to obtain a smooth numerical output, so the Kalman filtering can be used to make the RSSI output smoothness[5].

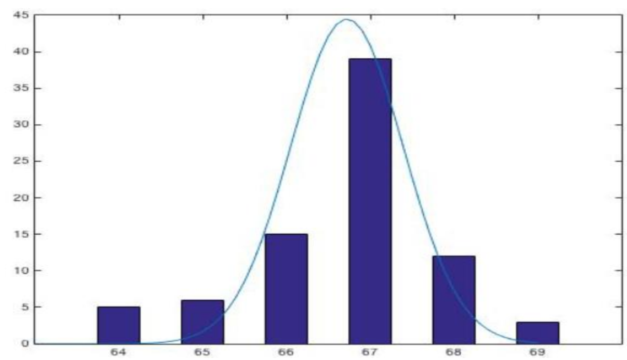


Figure 1. Recorded RSSI Values

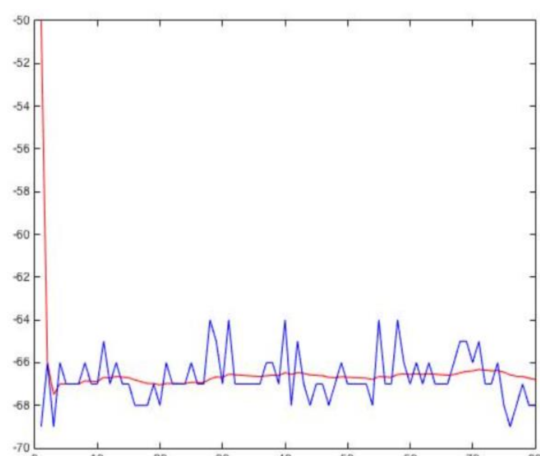


Figure 2. RSSI After Using Kalman Filtering

The blue line is the original RSSI data, floating up and down around -67; the red line is the gently changing data after Kalman filtering. It can be seen from Fig 2 that the RSSI value become smooth after Kalman filtering, which lays the foundation for the next step to improve the positioning accuracy.

3.3 Trilateration Algorithm

Triangular positioning method refers to a mathematical principle, which uses two or more detectors to detect the orientation of the target at different positions, and then uses the principle of triangular geometry to determine the position and distance of the target. Trilateration (trilateration) is a commonly used positioning algorithm. It is used to identify the location of the device. A device constantly emits roaming radio signals that may be picked up by three or more cell towers enabling the triangulation to work.

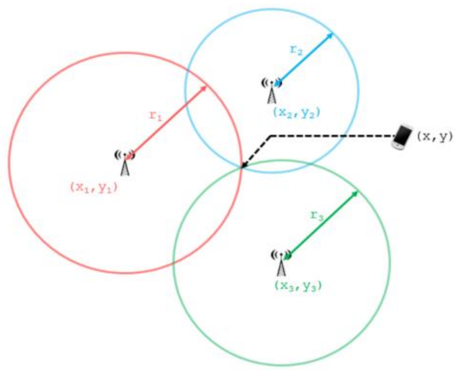


Figure 3. Trilateration Algorithm

On the diagram above, each circle represents all the possible locations of a mobile phone at a given distance (radius) of a cell tower. The aim of a trilateration algorithm is to calculate the (x, y) coordinates of the intersection point of the three circles. Each circle is defined by the coordinates of its center e.g. (x1, y1) and its radius e.g. r1.
Using d1, d2, and d3 as radii, make three circles. According to the Pythagorean theorem, the formula for calculating the position of the intersection point or unknown point is obtained:

$$(x_1 - x_0)^2 + (y_1 - y_0)^2 = {d_1}^2$$

$$(x_2 - x_0)^2 + (y_2 - y_0)^2 = {d_2}^2$$

$$(x_3 - x_0)^2 + (y_3 - y_0)^2 = {d_3}^2 \qquad (3)$$

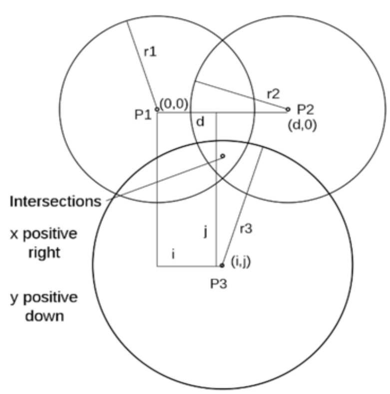


Figure 4. Trilateration Algorithm

Suppose the position of the unknown point is (x, y), let the sphere center coordinate of the first sphere P1 be (0, 0), P2 be at the same longitudinal coordinate, the sphere center coordinate is (d, 0), P3 sphere center coordinate Is (i, j), the three spherical radii are r1, r2, r3, and z is the height of the intersection of the three spheres and the horizontal plane. There are:

$$r_1{}^2 = x^2 + y^2 + z^2 \qquad (4)$$

$$r_2{}^2 = (x-d)^2 + y^2 + z^2$$

$$r_3{}^2 = (x-i)^2 + (y-j)^2 + z^2$$

When z = 0, it means that the three circles intersect into a point on the horizontal plane, and first solve x:

$$x = \frac{(r1^2 - r2^2 + d^2)}{2d} \qquad (5)$$

Transform formula two, substitute z2 of formula one into formula two, and then into formula three to get the calculation formula of y:

$$y = \frac{(r_1{}^2 - r^2{}_2 - x^2 + (x-i)^2 + j^2)}{2j} \qquad (6)$$

3.4 K-Nearest Neighbors Algorithm

The k-nearest neighbors (KNN) algorithm is a simple, easy-to-implement supervised machine learning algorithm that can be used to solve both classification and regression problems. For online RSS vectors s, calculate the distance (such as Euclidean distance) between it and each RSS vector {s1, s2,…, s_M} in the location library . Then it selects the nearest k location data record (a record is a correspondence between an RSS vector and a location).

- For KNN regression, the labels are coordinates x and y, which can be calculated numerically, and the position coordinates of the k records are averaged to obtain the positioning result.

- For the KNN classification, the positioning area is divided into a grid of 1m × 1m × 1m, each grid is regarded as a category, replaced by a grid label, the k network labels are counted and voted, and the grid with more votes is selected as the positioning result .

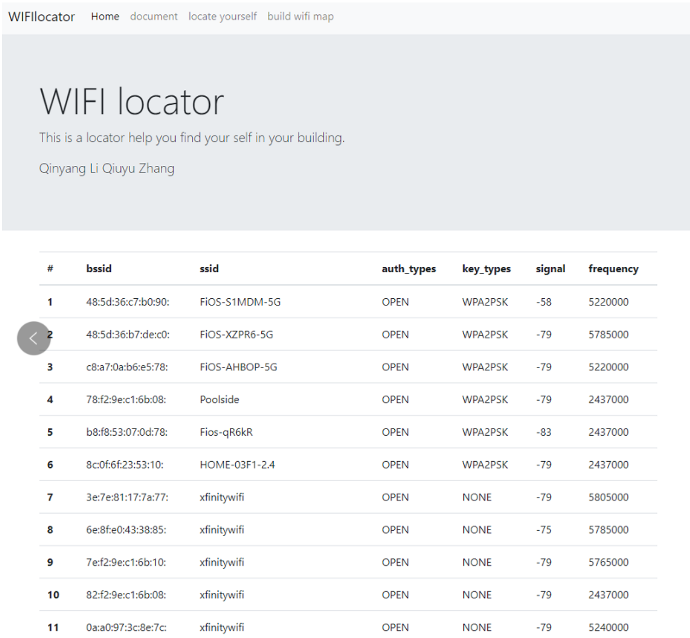
KNN is a lazy learning method. In the above process, it does not need to use the training data to "learn", it available to search directly in the training data when positioning.

3.5 Web Service

All the data we collected are stored in an online MySQL database. Amazon RDS offers free online storage service which we used in this project. In order to run our program without using the command line or any developer tool, we designed a python web application combined with Angular.js to present the function of our program on a webpage. Angular is a platform for building mobile and desktop web applications. We used Angular.js as our main method to communicate with the online database at the same time, we used Bootstrap as the frame of our web page. Finally, all web services will be uploaded and running at Heroku. It is a platform as a service based on a managed container system, with integrated data services, for deploying and running modern apps. If everything worked, everyone could have access to our WIFI locating program and be able to store the information about the WIFI around them.

4 Experimental Design and Result

The experimental environment is my house. The terminal is a laptop. I make my upper right corner of my room as the coordinate origin, the figure 5 is the WIFI Information we scanned:


WIFIlocator Home document locate yourself build wifi map

WIFI locator

This is a locator help you find your self in your building.

Qinyang Li Qiuyu Zhang

#	bssid	ssid	auth_types	key_types	signal	frequency
1	48:5d:36:c7:b0:90:	FiOS-S1MDM-5G	OPEN	WPA2PSK	-58	5220000
2	48:5d:36:b7:de:c0:	FiOS-XZPR6-5G	OPEN	WPA2PSK	-79	5785000
3	c8:a7:0a:b6:e5:78:	FiOS-AHBOP-5G	OPEN	WPA2PSK	-79	5220000
4	78:f2:9e:c1:6b:08:	Poolside	OPEN	WPA2PSK	-79	2437000
5	b8:f8:53:07:0d:78:	Fios-qR6kR	OPEN	WPA2PSK	-83	2437000
6	8c:0f:6f:23:53:10:	HOME-03F1-2.4	OPEN	WPA2PSK	-79	2437000
7	3e:7e:81:17:7a:77:	xfinitywifi	OPEN	NONE	-79	5805000
8	6e:8f:e0:43:38:85:	xfinitywifi	OPEN	NONE	-75	5785000
9	7e:f2:9e:c1:6b:10:	xfinitywifi	OPEN	NONE	-79	5765000
10	82:f2:9e:c1:6b:08:	xfinitywifi	OPEN	NONE	-79	2437000
11	0a:a0:97:3c:8e:7c:	xfinitywifi	OPEN	NONE	-79	5240000

Figure 5. WIFI Information

The four APs' coordinate have been stored in database before. In this spot, the x-coordinate value is 2.5267 m, the y-coordinate value is 3.6186 m. The accepted RSSI data was processed and the inverse calculation of signal intensity attenuation model was conducted, here, the η is 2.9, PL(d_0) is 36 dBm. X_σ is ignored because of the simplicity of indoor arrangement. The positioning result is shown as figure 6:

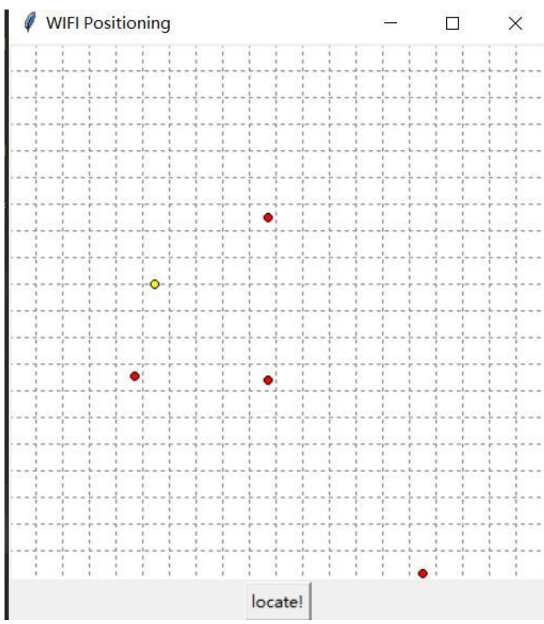


Figure 6. Positioning

5 Conclusion and Future Work

Trilateration algorithm can actually positioning terminal with the error of 1-2 m due to the fluctuates value of WIFI signal strength.

In the future work, we consider refining our algorithm by implementing KNN algorithm in order to improve positioning accuracy and reduce algorithm complexity. Our KNN algorithm used a manually tested the best K value. It is possible to improve algorithm and automatically update K value based on every situation.

Also, we plan to use Heroku which is a platform as a service based on a managed container system, with integrated data services and a powerful ecosystem, for deploying and running modern apps. However, we were unable to deploy our whole program to Heroku at this time. Due to It provides little information to debug and some compatibility of some modules we used during development. We designed a html base 3d presentation module for displaying the location of WIFI router and the user. However, it is not compatible with the current bootstrap web frame so we have to display our result as a terminal window.

References

[1] Yang, C., & Shao, H. R. (2015). WiFi-based indoor positioning. IEEE Communications Magazine, 53(3), 150-157.

[2] Yang, J., & Chen, Y. (2009, November). Indoor localization using improved rss-based lateration methods. In GLOBECOM 2009-2009 IEEE Global Telecommunications Conference (pp. 1-6). IEEE.

[3] Youssef, M. A., Agrawala, A., & Shankar, A. U. (2003, March). WLAN location determination via clustering and probability distributions. In Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003.(PerCom 2003). (pp. 143-150). IEEE.

[4] Shi, Y., Long, Y., Lu, F., Xu, Z., Xiao, X., & Shi, S. (2017, November). Indoor RSSI trilateral algorithm considering piecewise and space-scene. In 2017 IEEE International Conference on Smart Cloud (SmartCloud) (pp. 278-282). IEEE.

[5] Chen, Z.Zou, H. & Jiang, H. et al. (2015) Fusion of WiFi, smartphone sensors and landmarks using the Kalman filter for indoor localization, Sensors, 15, 715.

[6] Woodman, O. J. (2010, September). Pedestrian Localisation for Indoor Environments. Retrieved April 30, 2020, from http://www.cl.cam.ac.uk/research/dtg/www/publications/public/ojw28/Main-PersonalRedist.pdf