

Mobile User Location Tracking with Unreliable Data

Samia Zouaoui and Abdelmalik Bachir

Abstract We present an architecture of a location tracking system based on the deployment of completely passive monitors that capture WiFi messages transmitted by mobile users for reasons other than localization, such as those messages transmitted for connectivity management (e.g. probe request messages). Being completely passive, our system has the main advantage of being potentially able to track any WiFi equipped devices without the device contributing to the tracking or even being aware of it. However, the feature of complete passivity comes with the challenge of getting accurate localization at regular time intervals, because some devices may not transmit any WiFi message if they are not being actively used. In addition, some messages transmitted by the device might not be captured by monitors due to many reasons such as collision with another message, temporarily changing channel conditions, or software glitches due to the driver of capturing system. The missing of those messages affects the location accuracy of our tracking system because sometimes, the system has to rely on messages captured by less than three monitors. Therefore, we present two techniques to compensate for that missing data by estimating the current position of the user based on its previous positions. The first technique is called Direction and it targets selecting the most probable current position that minimizes the direction change compared to the past positions. The second method is called Speed; it takes as the most probable position the one that leads to the least speed change compared to previous speeds. Both Direction and Speed are inspired from the assumption that humans tend not to make abrupt changes in their speeds and directions while moving under normal circumstances. We evaluate our proposed techniques in comparison with Dead Reckoning technique by simulation with computer generated mobility data and with real mobility data from the CRAWDAD project. By using NS3 we evaluated our techniques with log-normal and indoor propaga-

S. Zouaoui (✉)

LINFI Laboratory, Department of Computer Science,
University of Biskra, P.O. Box 145 RP, 07000 Biskra, Algeria
e-mail: s.zouaoui@univ-biskra.dz

A. Bachir (✉)

LMA Laboratory, Department of Mathematics,
University of Biskra, P.O. Box 145 RP, 07000 Biskra, Algeria
e-mail: a.bachir@univ-biskra.dz

© Springer International Publishing Switzerland 2016
S. Chikhi et al. (eds.), *Modelling and Implementation of Complex Systems*, Lecture Notes in Networks and Systems 1, DOI 10.1007/978-3-319-33410-3_19

tion models. NS3 simulations on both log-normal and indoor propagation models show that both methods can lead to satisfactory results and missing of data can be compensated by the proposed heuristics.

Keywords Location tracking · Trilateration · WiFi · NS3 simulation

1 Introduction

The development of high capacity and performance computer systems coupled with the democratization of storage and the improvement of data manipulation methods has encouraged many data gathering and analytics applications with the goal of understanding various phenomena. One of the emerging and most prominent of these applications is the collection and analysis of crowd movements in confined areas where various techniques other than WiFi-based one have been used with more or less success [1].

However, with the increasing deployment of WiFi hot-spots and the democratization of smart-phones, tablets, and other hand-held WiFi-equipped devices, it has become possible to collect data on those users by capturing the signals transmitted by their devices. Indeed WiFi devices transmit some management frames from time to time searching for preferred access points and seeking association with them with the goal of accelerating hand-off between those access points. Those management frames are transmitted without encryption and thus can be captured and analyzed by any WiFi device with monitoring capabilities. Those frames are always transmitted with the same physical identifier (the MAC address of the device) and thus can be used track the movement of the user.

One of the most straightforward way to track the movements of a user is to perform proximity based localization where the user is deemed to be located near the access point with which it is associated or with the one from which it is receiving the strongest signal. While this proximity-based localization is sufficient in some tracking applications, having a better accuracy is required for a larger amount of applications. Therefore, there is a need for finer-grained WiFi-based location tracking systems.

Although there has been extensive research in the area of WiFi localization, particularly in indoor environment, with more or less satisfactory results due to the challenging constraints of radio propagation characteristics, location tracking causes much more challenges particularly in the case of completely passive tracking where the user does not need to cooperate or help the tracking nor does it even need (technically) to be aware of the tracking process. These additional challenges are mainly due to the irregularity and the small amount of suitable WiFi frames expected from the users. Those frames can be more or less frequent depending on the activity of the user with his hand-held device. If the device is not actively used it may go to sleep mode and refrain from transmitting messages for an extended period of time. The other challenge is the possibility of missing messages at the monitors. If an area is

overcrowded, the high number of frames captured by each monitor can lead to saturation thereby causing the missing of a number of those frames. These challenges make the tracking even more challenging than traditional localization systems where the user is receiving regular powerful signals from anchors which help averaging them and getting a better estimate of the signal strength.

In this paper, we present an architecture of a completely passive location tracking system based on the capture of WiFi frames transmitted by mobile users. We discuss the expected performance of multilateration which is one of the most practical techniques for WiFi devices. We run simulation with NS3 to assess the performance of the position tracking in a general propagation model (log-normal) and an indoor propagation model (IUT-R P.1238). We consider the problem of missing data that occurs at the monitors and propose two techniques to compensate for those missing data to estimate the current position of the user based on its previous position. The first technique is called Direction and it aims to select the most probable current position that minimizes the direction change among past positions. The second method is called Speed; it takes as the most probable position the one that leads to the least speed change compared to previous speeds. Both Direction and Speed are inspired from the assumption that humans tend not to make abrupt changes in their speeds and directions while moving. NS3 simulations on both log-normal and indoor propagation models show that both methods can lead to satisfactory results and missing of data can be compensated by the proposed heuristics.

2 Mobile User Location Tracking

2.1 Localization Versus Location Tracking

Localization is one of the technical areas that have received increasing attention in recent years due to the boom for location-based services that mobile users can benefit from and the wealth of localization applications in WiFi, ad-hoc, and sensor networks [2, 3]. Localization is defined as the process of determining the position of a mobile device at a given time. The localization process involves the use of wireless signals, to be exchanged with the non located node, in order to get some physical measurements that help in inferring the node's position.

Location tracking is a system that can follow the user mobility by measuring user movements (sequence of locations) over a period of time [4]. It can be achieved by logging the user's historical locations. Tracking applications are numerous including understanding shopping behaviors in malls, schools, public safety, disaster areas, airport, museums, campuses, and exhibitions [1].

There are two types of location tracking: active and passive. In active location tracking, the user performs positioning as in traditional localization technologies and then shares its positions to the tracking system. In passive location tracking, the user does not participate in the tracking procedure. The difference between those two

tracking types is that active tracking may lead to better accuracy as the user receives regular messages from anchors nodes thus leading to a better localization. However, this comes with the constraint of requiring the user to actively collaborate with the tracking system.

2.2 Indoor Location Tracking

Location tracking can be applied in two contexts: outdoor and indoor. For outdoor or (LOS: Line-Of-Sight) localization, GPS (Global Positioning System) is the most famous active location tracking system used. It works very well in open sky. However, its performance drop in other environments because GPS signals can be blocked by buildings, thick forests and other types of physical obstacles like walls, roofs, floors, etc. Thus, GPS does not work well in indoor or NLOS (Non-Line-Of-Sight) environments due to the complex structure and dynamic nature of indoor environments that affect the wireless signal propagation characteristics making it complex and hard to model. Multipath interference is a problem that exists in indoor environment which happens when the transmitted signal from a satellite is reflected due to barriers such as buildings or trees. Weak signals also affect the accuracy of the position.

2.3 User Detection Accuracy

An indoor environment is quite different from an outdoor environment. The propagation of a wireless wave can be influenced by some factors that would affect the accuracy of the location estimation of mobile users. In an indoor environment, walls, furniture, or walking people will change the propagation of the wireless wave and introduce variance to the wireless signal received by the user [3]. The RSS (Received Signal Strength) is usually quantified by RSSI (Received Signal Strength Indicator) which is a value that can be read from the wireless radio device. The accuracy of the measure provided by the RSSI is affected by the following factors.

- The Access Point (AP) may be blocked by an object, thus the received signal strength by AP from a terminal may be lower than it should be. Therefore, relying only on the RSSI to estimate a mobile user location becomes unreliable.
- Different environments have different levels of interference. The noise in one environment may be higher than in another due to the existence of many wireless devices transmitting electromagnetic waves.
- There could be refraction, reflection, diffraction, absorption, and scattering of radio signals, which causes the signal strength to be weakened.
- The signal strength can be affected by multipath fading or shadow fading.

3 WiFi-Based Location Tracking

The use of WiFi has many attractive features such as: (i) existing low-cost hardware, (ii) large-scale deployment of WiFi, (iii) free software, (iv) no need for sophisticated special hardware, (v) no need for users to install applications or even be aware of the passive location tracking.

3.1 *Frame Types*

WiFi networks use radio technologies called IEEE 802.11 to provide secure, reliable, fast wireless connectivity. A typical WiFi set-up includes one or more access points (APs) and one or more clients. An AP broadcasts its SSID (service set identifier, or “network name”) via packets that are called beacons, which are usually broadcast every 100 ms. The beacons are transmitted at 1 Mbit/s, and are of relatively short duration and therefore do not have a significant effect on performance.

A mobile user running WiFi transmits many types of frames: data, control, and management. While data frames are most likely to be encrypted, management frames are transmitted in clear and thus reveal the identity of the user which can be used to track its movements.

All of these frames contain a frame header, which includes the source and destination MAC addresses. It also contains information such as beacon interval and Service Set Identifier (SSID), which is the name of the WLAN. The SSID is important for a terminal to know which network it is trying to establish a connection with. Management frames perform supervisory functions. They are used for the purpose of establishing a connection between an AP and a terminal. A terminal in a WLAN, with multiple APs deployed, may move around and as a result, the terminal may need to switch association from one AP to the next using management frames. They perform the following operations: (i) join and leave wireless networks, and (ii) move associations from access point to access point. In addition to management frames, control frames are used to coordinate data frame exchange. Although, location tracking can be done on any type of frames, we focus on management frames as they are transmitted in clear without encryption.

3.2 *Wireless Modes*

Most wireless users only use their wireless cards as a station to an AP. In managed mode, the wireless card and driver software rely on a local AP to provide connectivity to the wireless network. Another common mode for wireless cards is ad-hoc mode. Two wireless stations that want to communicate with each other directly can do so by sharing the responsibilities of an AP for a limited subset of wireless LAN services.

Ad-hoc mode is used for short-term connectivity between stations, when an AP is not available to provide connectivity.

Many wireless cards also support master mode, where the wireless card provides the services of an AP when paired with the appropriate software. Managed mode allows to configure a wireless card to connect to an AP. Finally, wireless cards support monitor mode functionality. When configured in monitor mode, the wireless card stops transmitting data and sniffs the currently configured channel, reporting the contents of any observed packet to the host operating system. This mode is useful for completely passive location tracking systems as the entire contents of wireless packets, including header information can be analyzed [5].

4 Positioning Approaches

Positioning systems can be classified according to the measurement techniques they employ to determine the user's location. There are many approaches: triangulation, multilateration, area-based, and fingerprinting [6–9]. In this paper, we focus on multilateration as it is the most practical among localization approaches.

In the multilateration, the localization is based on turning RSSI measures into distances from the mobile user to the anchors. The conversion from RSSI to distance is based on a path loss model (also called radio propagation or attenuation model) which predicts the loss in signal strength in function of the distance between the source and destination nodes. The loss in signal strength is caused by (i) distance, (ii) multipath (reflected, diffracted, or scattered copy of the transmitted signal) and (iii) shadowing (blockage of signal due to obstacles). To predict the loss in signal in different environments, different propagation models have been developed and can be categorized into two classes; theoretical (deterministic) and experimental (statistical) models. Theoretical models try to simplify the complex behavior of path loss, multipath and shadowing using mathematical models [7]. A widely used model is the log-normal path-loss that predicts the path loss a signal encounters inside a building or densely populated areas over distance. Mathematically, the received power over a distance d between the transmitter and the receiver according to the log-normal model is given in (1). We have:

$$P_r(d) = P_r(d_0) - 10n \log \left(\frac{d}{d_0} \right) - X_\sigma \quad (1)$$

where d_0 is the reference distance generally taken equal to 1 m, n is path-loss exponent, X is a Gaussian random noise variables of average for 0 (dBm) and standard deviation of σ (dBm). $P_r(d)$ is received signal power (dBm) and $P_r(d_0)$ is the received signal power at the reference distance (dBm). Multilateration algorithms aim at providing a good estimate of the user location given the exact anchor locations and distances the user to each anchor. Multilateration requires at least three non collinear anchors to be able to estimate the position of the user. To estimate the position of

the tracked user, the monitors obtain RSS measures and turn them into distances required to apply the multilateration algorithm. These distances can be obtained by solving (1) for d which results in:

$$d = d_0 * \exp \frac{P_r - P_r(d_0) - X_\sigma}{10 * n} \tag{2}$$

Let (x, y) be the coordinates of Monitor i and d_i the distance between the user and Monitor i . We have the following:

$$\begin{cases} d_1^2 = (x_1 - x)^2 + (y_1 - y)^2 \\ \dots \\ d_n^2 = (x_n - x)^2 + (y_n - y)^2 \end{cases} \tag{3}$$

Equation (3) can be rewritten to as:

$$\mathbf{AX} = \mathbf{b} \tag{4}$$

where

$$\mathbf{A} = \begin{pmatrix} 2(x_1 - x_n) & 2(y_1 - y_n) \\ \vdots & \vdots \\ 2(x_{n-1} - x_n) & 2(y_{n-1} - y_n) \end{pmatrix}, \mathbf{X} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{5}$$

$$\mathbf{b} = \begin{pmatrix} x_1^2 - x_n^2 + y_1^2 - y_n^2 + d_1^2 - d_n^2 \\ \vdots \\ x_{n-1}^2 - x_n^2 + y_{n-1}^2 - y_n^2 + d_{n-1}^2 - d_n^2 \end{pmatrix} \tag{6}$$

By adopting the minimum variance estimation method, the coordinates (x, y) of the user can be calculated. We have:

$$\mathbf{X} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \tag{7}$$

Note that multilateration is a very efficient technique. Its main weakness is caused by the inefficiency of the RSSI measure to be turned into a distance, because this depends on the knowledge of the environment constraints that may change from one region to another closed region and from time to time.

5 Location Tracking with Unreliable Data

5.1 *The Problem of Missing Data*

When using monitors to capture packets transmitted by a user, it is not uncommon that a capture is missed by a monitor. This can be caused by many factors such as obstacles obstructing the signal, hardware problems at the radio transceiver, saturation due to a high number of packets being captured, etc. To evaluate the amount of those missed captures, we run experiments with three monitors placed in an office environment in a Professional Education Institute. We installed three monitors M1, M2, and M3 in various locations of the office. In the experiments scenario, we let a user move in the office and transmit packets from time to time, and let every monitor capture those packets and measure their corresponding RSSIs. At the end of the experiments, the monitors collected 497 packets in total. Monitors 2 and 3 observed a loss of 6 and 41 packets respectively, which makes the total loss rate of 9.46 %.

5.2 *The Effect of Missing Data*

With missing RSSI readings, it becomes difficult to estimate the location of the user. We consider the case where there are two RSSI readings which could be turned into two distances. Therefore, (4) will not necessarily have a unique solution. The results of that equation will depend on the positions of the two circles centered at the two monitors with ranges as distances obtained from the RSSI readings. There will be multiple cases: no solutions when the two circles do not intersect, infinity of solutions if the two circles are the same, one solution if the two circles touch at a single point and two solutions if the two circles intersect at two different points.

5.3 *Estimating Position with Missing Data*

We consider the case where there are two solutions and aim at finding the best methods to eliminate the unlikely location and keep the most probable one. For the multilateration technique, we use two heuristics to estimate the most probable position of the user. These metrics are as follows.

5.3.1 **Direction Method**

We assume that humans are less likely to make abrupt changes in their movements. Therefore, in our selection of the most suitable point, we take the one with the least direction changes among potential candidate points. We calculate the movement

vectors of all potential candidates and take the one with the minimum direction change. Mathematically this reduces to taking the vector the maximum *cosine* value with the previous movement vector. Assume that the user was at Location L_{-2} , then L_{-1} , and we want to eliminate L_0 or L'_0 the two potential current locations resulting from the intersection of the two circles. We take the location estimate \hat{L} which results in the minimum cosine value among the following:

$$\hat{L} = \begin{cases} L_0 & \text{if } |\cos(L_{-2}L_{-1}, L_{-1}L_0)| < |\cos(L_{-2}L_{-1}, L_{-1}L'_0)| \\ L'_0 & \text{otherwise} \end{cases} \quad (8)$$

5.3.2 Speed Method

We assume that humans are likely to change the pace of their movements abruptly. Therefore, we take the point that is closest to the history of the speed of movement of users. Technically we calculate the distances of all potential candidates from the current point. For all these points we calculate the corresponding velocities and take the point whose the corresponding velocity is closest to the previous speed. The basic idea is to measure the minimum distance between the mobile node and the two points of intersection of two circles by using Euclidean distance. If we assume that the user was at Location L_{-2} (resp. L_{-1} , L_0 , L'_0) at time t_{-2} (resp. t_{-1} , t_0 , t_0), we calculate the velocities v_0 and v'_0 and compare them to the previous velocity v_{-1} .

$$v_{-1} = \frac{\|L_{-2}L_{-1}\|}{t_{-2} - t_{-1}}, v_0 = \frac{\|L_{-1}L_0\|}{t_{-1} - t_0}, v'_0 = \frac{\|L_{-1}L'_0\|}{t_{-1} - t_0} \quad (9)$$

where $\|X\|$ is the norm of the vector X . Thus, the user is assumed to be at the location that minimizes the difference in velocity. We have:

$$\hat{L} = \begin{cases} L_0 & \text{if } |v_{-1} - v_0| < |v_{-1} - v'_0| \\ L'_0 & \text{otherwise} \end{cases} \quad (10)$$

5.3.3 Dead Reckoning

Is a localization technique proposed in [10]. In Dead Reckoning, nodes are localized during a time interval called checkpoint. There are two localization phases in Dead Reckoning. The first phase is called initialization phase during which a node is localized using the multilateration mechanism. A node remains in the initialization phase until it localizes using the multilateration mechanism. The subsequent localization phase is called sequent phase. In this phase, a node localizes itself using only two anchor nodes. Bezouts theorem [11] is used to estimate the node's locations. Let (x, y) be the position of an unknown node and (a_1, b_1) , (a_2, b_2) be the position of

two of its neighboring anchor nodes. Moreover, let the distance between an unknown node and the respective anchor nodes be d_1 and d_2 , respectively. Then

$$\begin{cases} (x - a_1)^2 + (y - b_1)^2 = d_1^2 \\ (x - a_2)^2 + (y - b_2)^2 = d_2^2 \end{cases} \quad (11)$$

After solving the Eq. (11), the algorithm estimates two positions $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$. Next, the node computes the correction factors (Cf_1 and Cf_2) to select one of the two estimated positions P_1 and P_2 . The correction factor is computed by using $P(\hat{x}, \hat{y})$ which the position of the node using multilateration in the first time. After that, it use the previous position at the checkpoint t_i to estimate its location in the next checkpoint at t_{i+1} .

$$\begin{cases} Cf_1 = \sqrt{(\hat{x} - x_1)^2 + (\hat{y} - y_1)^2} \\ Cf_2 = \sqrt{(\hat{x} - x_2)^2 + (\hat{y} - y_2)^2} \end{cases} \quad (12)$$

The correct position of the node is P_1 if ($Cf_1 < Cf_2$). Otherwise, it will be P_2 . This is because, the calculated position $P(\hat{x}, \hat{y})$ always deviate from the actual position by a small margin.

5.4 Error Estimation

To evaluate the accuracy of our location tracking system, we calculate the estimation error between the real location L and the estimated one returned by our system \hat{L} . We have:

$$\epsilon = \|L - \hat{L}\| \quad (13)$$

6 Simulation

To evaluate the performance and the accuracy of our location tracking methods, we run extensive simulations with NS3 which is a discrete-event network simulator where the simulation core and models are implemented in C++. NS3 is open source and licensed under the GNU GPLv2 license and therefore has benefited from a growing community base which contributed to adding more radio propagation models and network protocols. Since its release in 2008 it is one of the most important and widely used network simulation tools. Creating a NS3 simulation consists of four basic steps. These basic component types of a network are nodes, applications, net devices, channels and topology helpers [12, 13]. An important part of any wireless network simulation is the appropriate choice of the propagation loss model to be used

Table 1 Description of real mobility DataSets used

	Environment size	Environment	User type	Sample size
CRAWDAD DataSet1 [14]	400 m ²	Office	Mobile robot	1689
CRAWDAD DataSet2 [15]	15 m × 36 m	Office	User	60
KIOS DataSet3 [16]	500 m ²	Office	User	96
CRAWDAD DataSet4 [17]	50 m × 75 m	Office	User	180

to model the performance of a wireless network. These models are needed for the simulator to compute the signal strength of a wireless transmission at the receiving stations. There are a variety of such models in NS3.

The indoor radio propagation model we used with NS is implemented according to the description of ITU-R P.1238. We considered two cases ($\sigma = 0$ and 1) to reflect various environments. The area that we used in our simulations is an office $10 \times 20 \times 10$ building with concrete windows. This building has one floors and an internal 20×2 grid of rooms of equal size.

For the user mobility, we considered computer generated mobility data and real mobility traces. In the computer generated data, we considered an ideal mobility model where the user is simulated to move along constant direction with a constant speed. We also considered the indoor mobility model that comes with the NS3 package. For the real mobility data, we considered 4 data sets from the CRAWDAD and KIOS projects. The main characteristics of these datasets (DS) are summarized in Table 1.

For all these mobility scenarios, we use three monitor places at non collinear positions. For each position of the user we simulate the transmission of a message from the user that will be captured by the monitors. Each monitor that captures a message, reads its RSSI and turns it into a distance that is used to estimate the user position according to one the techniques tested: Dead Reckoning, Distance, and Speed.

To simulate unreliable data, we introduce random losses at the monitors. We introduce a probability of missing a message for each monitor. Initially we assume that only one monitor misses a message at a time so there are always at least two other monitors receiving the same message, and we aim to estimate the position of the user based on the available two RSSI measures.

6.1 Location Tracking Without Missing Data

In the case ideal mobility model and $\sigma = 0$, all the positions can be correctly estimated and the estimated positions perfectly match the real ones. In a little more complex situation where $\sigma = 1$, the estimated positions do not match the real ones even with an ideal mobility model (see Fig. 1). However, as we notice in the same

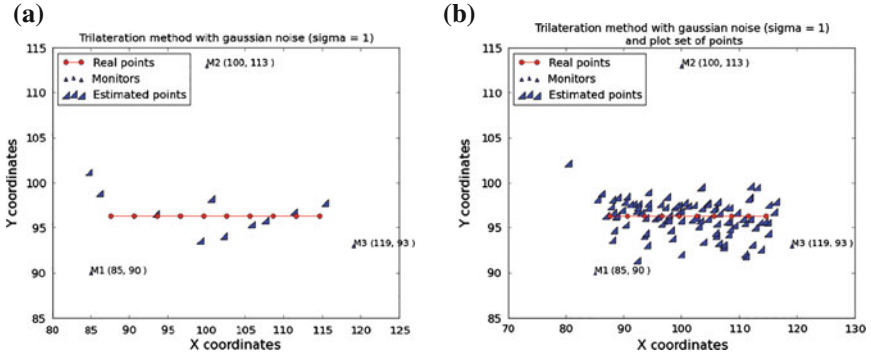


Fig. 1 Coordinates of estimated position calculated based on RSS measures. Ideal mobility with $\sigma = 1$. **a** Each position is used to calculate 1 coordinate estimate. **b** Each position is used to calculate 10 coordinate estimates

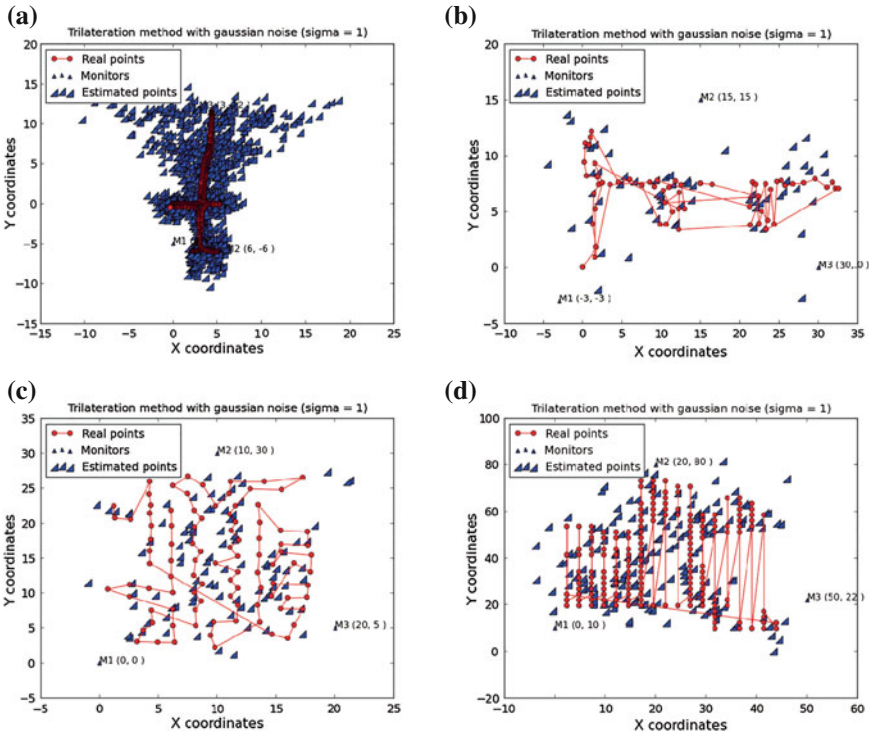


Fig. 2 Position coordinates from real mobility data and their corresponding position estimated without missing data and with $\sigma = 1$. **a** DataSet1. **b** DataSet2. **c** DataSet3. **d** DataSet4

Table 2 Summary of the obtained results

DS	Pos.	Method	$\sigma = 0$		$\sigma = 1$		Trilateration	
			Mis. = 74 %	Mis. = 99 %	Mis. = 74 %	Mis. = 99 %	$\sigma = 0$	$\sigma = 1$
1	1689	Direction	0.121511	0.264769	1.645680	1.602990	0.0026119	1.6618
		Speed	4.268930	5.645070	4.865160	5.904530		
		Dead reckoning	0.706925	0.849137	1.789630	2.006390		
2	60	Direction	1.526920	1.946580	3.763000	3.848790	4.74E-06	2.3797
		Speed	4.692470	5.564970	5.818530	6.378680		
		Dead reckoning	3.564433	4.127600	4.578430	4.997270		
3	96	Direction	0.837704	0.611903	2.547160	2.084700	4.45E-06	2.14042
		Speed	1.858930	0.327840	4.009810	2.300220		
		Dead reckoning	2.018800	2.108680	3.117430	2.823130		
4	180	Direction	5.537220	3.412730	6.861800	6.921910	1.13E-05	4.77129
		Speed	8.400200	8.615250	9.903870	10.850200		
		Dead reckoning	3.929080	9.361340	7.637980	9.929080		

CDF for DataSet1

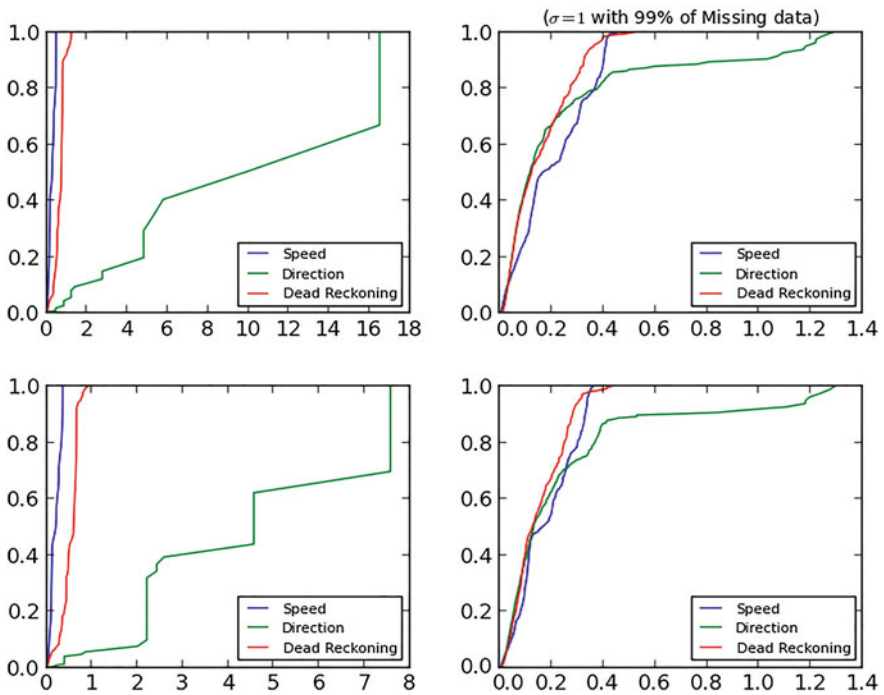


Fig. 3 CDF with DataSet1

figure, the difference between these values quantified by the by the estimation error is low.

In the case of real mobility data and $\sigma = 1$, we plotted the graphs in Fig. 2 to show the effect of RSSI fluctuations of the estimation of positions in the case of real mobility data. The mean errors in position estimation with these scenarios are summarized in Table 2.

6.2 Location Tracking with Missing Data

We run various scenarios for missing data. In the first one, Missing0 already discussed above, there are no missing data and all the messages transmitted by the user will be captured by the monitors. In the second and third scenarios only a proportion of the messages are captured by the monitors. For example, in Missing74 and Missing99 scenarios, 74 and 99 % of the localization estimations are based on the RSSI readings of only two monitors, respectively. For various setting, we plot the CDF of the error in position estimation expressed in meters as shown in Figs. 3, 4, 5 and 6.

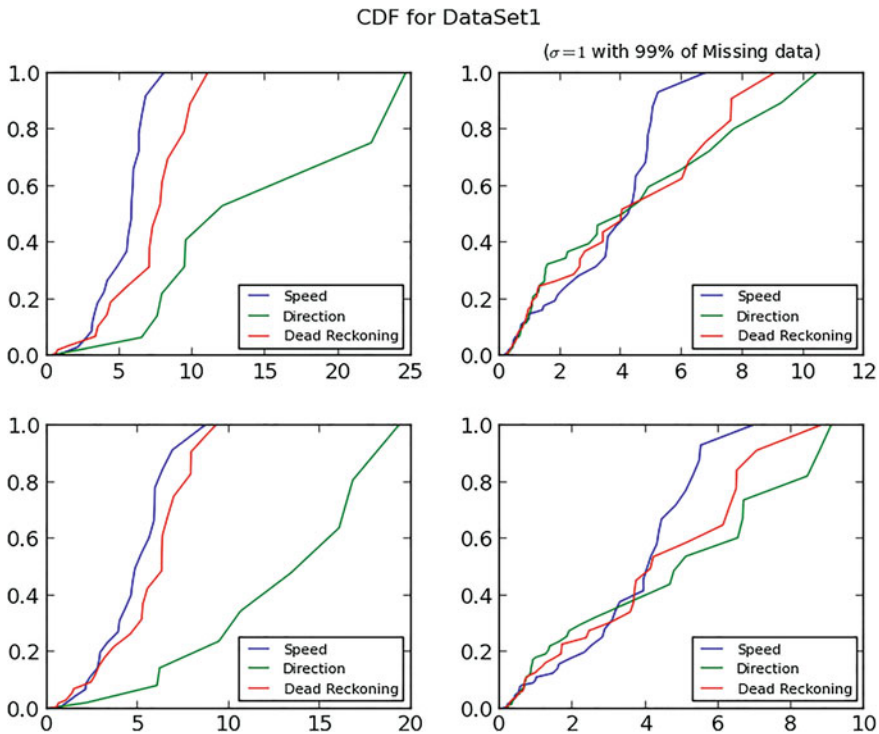


Fig. 4 CDF with DataSet2

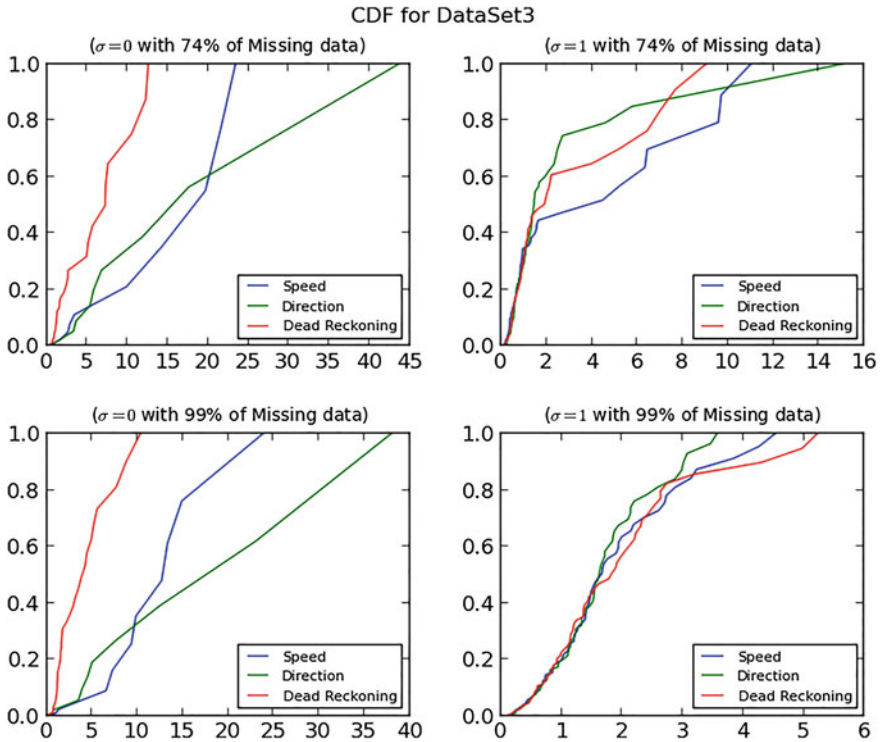


Fig. 5 CDF with DataSet3

In general, we also notice that when σ increases the error in position estimation also grows as explained earlier in the case of no missing data. We also show that Direction method achieves the best position estimation followed by Dead Reckoning followed by Speed and that there is not a big difference in position estimation when the ratio of missing data increases from 74 to 99%. In some situations such as in DataSet3 (Fig. 5), we show that Speed achieves good position estimate. The reason is that in DataSet3, the user movements are regular. For the DataSet4, we show that all methods return higher errors in position estimation. This is because user movements are not regular as the users seems to make vertical movements due to the large time interval between recorded positions. There is another reason why position estimates are less good in DataSet4 compared to the other data sets, which is due to the placement of monitors that farther away from the positions to be estimated. We also show that Direction method is the best in situation where we have short intervals between positions recording such as in DataSet1 (Fig. 3).

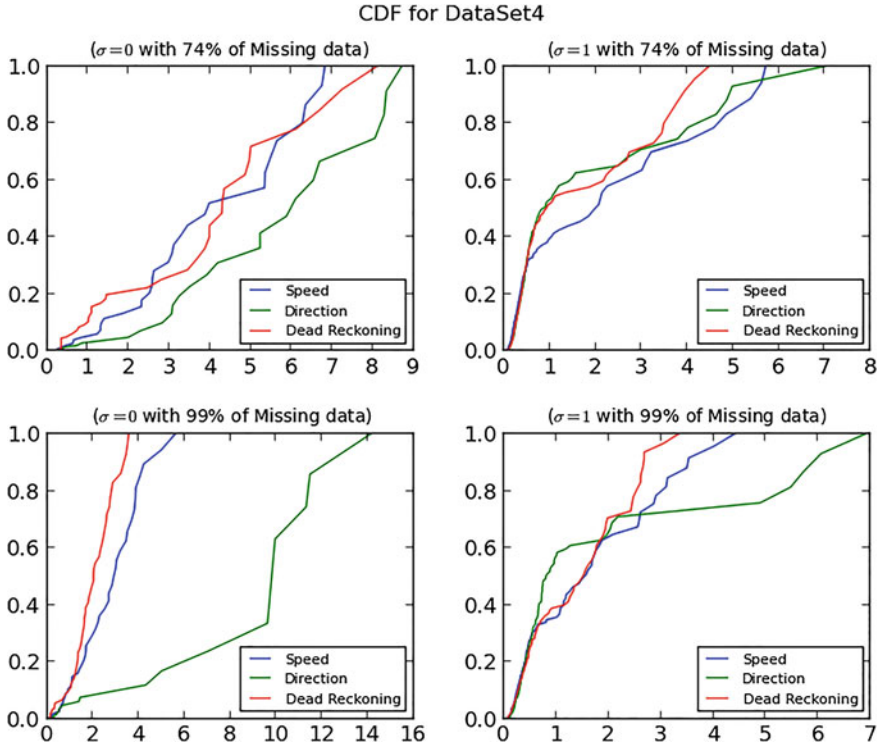


Fig. 6 CDF with DataSet4

7 Conclusions

We have presented a completely passive location tracking system that allows finding WiFi equipped mobile user locations. We analyzed the performance of the proposed system in a general environment governed by a log-normal path loss model with both real and computer generated data sets. We presented the problem of missing data that occurs at the monitors which we validated with experiments based on WiFi-enabled Raspberry Pi monitors. To cope with the problem of missing data, we proposed two heuristics in relation with the direction and speed change of mobile user with the assumption that humans users are likely to keep nearly constant speed and direction in their movements. NS3 simulations on both a general path-loss model and an indoor model on both computer generated and real mobility data have shown that Direction method achieves better results in general compared to Speed and Dead Reckoning.

Acknowledgments This work is funded in part by the Algerian Ministry of Higher Education and Scientific Research under contract B*01420130132. We gratefully acknowledge the use of wireless data from the CRAWDAD archive at Dartmouth College.

References

1. Larson, J.S., Bradlow, E.T., Fader, P.S.: An exploratory look at supermarket shopping paths. *Int. J. Res. Mark.* **22**(4), 395–414 (2005)
2. Schiller, J., Voisard, A.: *Location-Based Services*. Elsevier (2004)
3. Lasla, N., Younis, M.F., Ouadjaout, A., Badache, N.: An effective area-based localization algorithm for wireless networks. *IEEE Trans. Comput.* **64**(8), 2103–2118 (2015)
4. Zhang, D., Xia, F., Yang, Z., Yao, L., Zhao, W.: Localization technologies for indoor human tracking. In: 2010 5th International Conference on Future Information Technology (FutureTech), pp. 1–6. IEEE (2010)
5. IEEE standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements Part 11: wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007), pp. 1–2793, March (2012)
6. Cisco Systems. Location tracking approaches. In: *Wi-Fi Location-Based Services 4.1 Design Guide*. Cisco Systems (2008)
7. Chuan-Chin, P., Chuan-Hsian, P., Lee, H.-J.: Indoor location tracking using received signal strength indicator. In: *Emerging Communications for Wireless Sensor Networks*. InTech (2011)
8. Guvenc, I., Abdallah, C.T., Jordan, R., Dedeoglu, O.: Enhancements to rss based indoor tracking systems using kalman filters (2003)
9. Lasla, N.: *Toward an efficient localization system for wireless sensor networks*. PhD Thesis, USTHB (2015)
10. Rashid, H., Turuk, A.K.: Dead reckoning localisation technique for mobile wireless sensor networks. *IET Wirel. Sens. Syst.* **5**(2), 87–96 (2015)
11. Bezouts theorem. <http://www.sjsu.edu/faculty/watkins/bezout.htm>. Accessed 28 Jan 2016
12. NS3 project. <https://www.nsnam.org/docs/release/3.9/manual.pdf>. Accessed 15 July 2015
13. Rampfl, S.: Network simulation and its limitations. In: *Proceeding zum Seminar Future Internet, Innovative Internet Technologien und Mobilkommunikation und Autonomous Communication Networks*, vol. 57 (2013)
14. Parasuraman, R., Caccamo, S., Baberg, F., Ogren, P.: CRAWDAD dataset kth/rss (v. 2016-01-05) (2016). <http://crawdad.org/kth/rss/20160105>
15. King, T., Kopf, S., Haenselmann, T., Lubberger, C., Effelsberg, W.: CRAWDAD dataset mannheim/compass (v. 2008-04-11) (2008). <http://crawdad.org/mannheim/compass/20080411>
16. Panayiotou, C., Laoudias, C., Piche, R.: Device self-calibration in location systems using signal strength histograms. *J. Location Bas. Serv.* **7**(3), 165–181 (2013). KIOS WiFi RSS dataset (PDF Download Available). doi:10.1080/17489725.2013.816792. Accessed 6 Feb 2016
17. Bauer, K., Anderson, E.W., McCoy, D., Grunwald, D., Sicker, D.C.: CRAWDAD dataset cu/rssi (v. 2009-05-28) (2009). <http://crawdad.org/cu/rssi/20090528/text>