Peer-reviewed version available at *Sensors* **2020**, *20*, 427; doi:10.3390/s20020427

Article

Crowdsensing Influences and Error Sources in Urban Wi-Fi Fingerprinting Positioning

Cristian-Liviu Leca [1]*, Ioan Nicolaescu [1] and Petrica Ciotirnae [1]

[1] Military Technical Academy 'Ferdinand I', Bucharest, Romania; cristian.leca@mta.ro, ioannic@mta.ro, petrica.ciotirnae@mta.ro

* Correspondence: cristian.leca@mta.ro

Abstract: Wi-Fi fingerprinting positioning systems have been deployed for a long time in location-based services for indoor environments. Combining mobile crowdsensing and Wi-Fi fingerprinting systems could reduce the high cost for collecting the necessary data enabling the deployment of the resulting system for outdoor positioning in areas with dense Wi-Fi coverage. In this paper, we present the results attained in the designing and evaluation of an urban fingerprinting positioning system based on crowdsensed Wi-Fi measurements. We first asses the quality of the collected measurements, highlighting the influence of received signal strength on data collection. We then, evaluate the proposed system by comparing the influence of the crowdsensed fingerprints on the overall positioning accuracy for different scenarios. The evaluation helps gain valuable insight into the design and deployment of urban Wi-Fi positioning systems while also allowing the proposed system to match GPS-like accuracy in similar conditions.

Keywords: crowdsensing; databases; smartphones; urban positioning; Wi-Fi fingerprinting

---

1. Introduction

Mobile phone sensing has attracted much attention in the recent past due to factors, such as their ubiquitous presence in everyday life, the addition of sensors not specific to their original design (barometer, accelerometer, light etc.) allowing them to be used intuitively for mobile sensing. Also, their permanent connection to the Internet using mobile data or a Wi-Fi connection has allowed the usage of the mobile phone sensors in areas such as transportation, tourism, logistics, quality of air monitoring or social networking [1,2]. Mobile phone sensing can have its powers augmented by engaging a large number of participants (*crowd*) into contributing for a small cost or by a common cause (e.g. to monitor road traffic and obtain real-time navigation) by the process of mobile crowdsensing.

We consider that the development of the mobile crowdsensing paradigm is essential to the development of function Wi-Fi fingerprinting positioning systems for urban and outdoor areas. Previous research on fingerprinting positioning is mainly aimed at designing techniques and systems for indoor areas [3,4]. While the development of indoor Location-based-services and positioning systems have been motivated by complex, non-line-of-sight(NLOS) environments and a general unavailability of GPS positioning, urban crowdsensed fingerprinting positioning will compete with GPS positioning for mobile devices equipped with both sensors and will really prove to be a problem-solver for energy and size constrained IoT Devices requiring GPS-like accurate positioning.

In this paper, we apply the mobile crowdsensing technique to prove the feasibility of designing a Wi-Fi fingerprinting positioning system for urban areas. The envisioned system benefits from the ubiquitous presence of Wi-Fi access points (AP) in urban areas, as proven in our previous research [5–7]. The Wi-Fi AP received signal strength (RSS) measurements are done by the help of the 802.11n network interface of the mobile phones used for crowdsensing.

We propose the usage of the crowdsensing technique as opposed to previous research using automated or grid-based fingerprint collection for indoor positioning systems. Crowdsensing is suited for covering urban areas [8] as it offers:

- continuous growth and resolution of the coverage area for the positioning systems;
- coverage of areas restricted from public access (e.g. private homes);
- constant feedback on the APs situation allowing for the removal of inactive, mobile or moved ones;
- high density of fingerprints in areas most frequented by the users resulting in a higher accuracy of the positioning in areas most likely to be required;

The main challenge specific to the proposed solution is to prove that fingerprinting positioning can achieve GPS-like accuracy in similar conditions specific to the urban environment. To address this we thoroughly evaluate the accuracy of the fingerprinting positioning system by using different sub-datasets extracted from our collected data. For the position estimation, we apply the methods proposed in the papers published by Lohan et al. and Mendoza et al. [9,10].

The remainder of the article is organized as follows: Section 2 covers related work on similar positioning systems and highlights our previous work on the subject. Section 3 introduces the data collection method and the motivation behind choosing a specific test area for our research. Section 4 discusses the implications of the crowdsensing technique on fingerprint database generation. In section 5 we present the results of our analysis, highlighting the characteristics of the positioning algorithms used and the influence of the crowdsensed measurements. Section 6 is dedicated to a comparison of urban fingerprinting and GPS positioning accuracy. Section 7 concludes our work and highlights the main contributions of the paper.

2. Related Work

Outdoor positioning is usually based on multilateration or triangulation techniques [11] which work better in line-of-sight propagation conditions. These type of methods are not suited for urban environments affected by fading and NLOS propagation [6]. An approach, presented in [12], aims at estimating the AP position by the help of information contained in the APs name. Indoor positioning systems [13] have also tackled the NLOS problem, one particular promising solution being the process of signal association with particular locations [14–17]. In this class of fingerprinting positioning methods, a position is solely characterized by it's measured signal pattern, consisting of the consecutive collected Wi-Fi measurements. Thus, our proposed crowdsensed fingerprinting scheme is applied similarly without the need of knowing exact AP positions. Our method does not require distance or angle measurements, leading to its suitability for usage in urban areas.

Wi-Fi fingerprinting has traditionally exploited Wi-Fi interfaces equipped on mobile devices not limited to smartphones [18] and the ubiquitous presence of Wi-Fi APs [19]. Many of the early proposed systems [20,21] rely on an initial training phase required for constructing the fingerprint database to be used for the positioning phase. Outdoor positioning system would need significant time and effort for the training phase which crowdsensing systems aim at limiting. Crowdsensed systems are suited for outdoor environments where smartphones automatically collect RSS measurements coupled with GPS obtained locations for the training phase. As an alternative, crowdsourced systems have been proposed for indoor systems which require an active contribution from the user. In crowdsourced systems, the users are required to input their exact position in the positioning area using a dedicated app. [22–26].

Designing a fingerprinting positioning system for urban environments requires knowledge regarding the presence and density of Wi-Fi APs in the coverage area of the system. Our previous work on fingerprinting positioning [6] proved the ubiquity of Wi-Fi APs in urban environments and offers insights regarding the data collection methodology, fingerprint database maintenance and sources of errors specific to the environment such as mobile APs.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2019 doi:10.20944/preprints201911.0186.v1

Peer-reviewed version available at Sensors 2020, 20, 427; doi:10.3390/s20020427

Achieving seamless urban positioning with no constraints on the indoor/outdoor environment conditions is a significant step towards implementing a global indoor positioning systems as proposed in [27].

As highlighted in previous research [9,10], Wi-Fi fingerprinting positioning systems encounter one major problem consisting of challenges in implementing and reproducing results published in previous research papers on the subject. To address this issue we format our data to match the format proposed to be used for standard data-sets in [9] and we use two of the positioning algorithms proposed in [10], the weighted-centroid method and the log-Gaussian distance method.

3. Fingerprint analysis and test data description

The data collected for our research consisted of more than 1 million Wi-Fi measurements collected in Bucharest(Romania) and areas nearby the city by the help of 8 smartphones carried by different users for 6 months. Due to the crowdsensing method, the collected measurements are denser near the city center where the scanning occurred more often. This prompted us to select one of the most frequented areas, shown in Fig. 1 during the crowdsensing step for the evaluation of the collected data on the accuracy of Wi-Fi fingerprinting positioning.

Figure 1. Detail of the area.

The chosen test area has a high density of Wi-Fi measurements and allows for long-term variations in Wi-Fi signal propagation. The area is characteristic to the urban environment of Romania, being placed between two high traffic boulevards while the buildings in the area are tower apartment blocks. Locations of measurements in the test area are shown in Fig. 2.

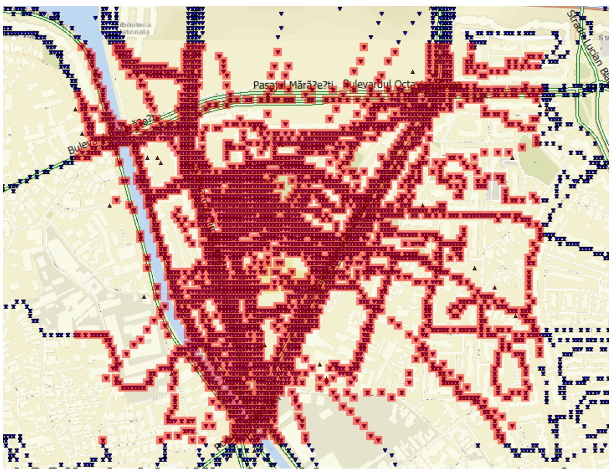

Figure 2. Measurement points in the test area

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2019 doi:10.20944/preprints201911.0186.v1

Peer-reviewed version available at Sensors 2020, 20, 427; doi:10.3390/s20020427

The test area, with a surface of $1km^2$, scanning resulted in 75188 fingerprints of 10072 distinct APs. The average signal level for all fingerprints was -83,85 dBm with a standard deviation of 5.46 dB.

The fingerprints coordinates were transformed from the decimal degree global format to local coordinates referenced to the upper left corner of the test area.

The test area dataset and all subsequent sub-datasets were randomly split between training data - 85 % of the total and test data - 15 % of the remaining fingerprints. The training fingerprints are used for estimating the position of the test fingerprints. The resulting estimated position is compared to the test fingerprint position resulting in a positioning error. Fig. 3 is an example of training and test data positions with locally referenced coordinates expressed in meters.

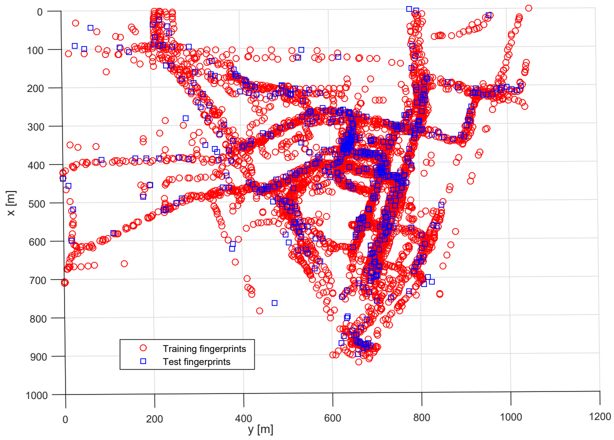

Figure 3. Training and test fingerprint positions

4. Crowdsensing influence on fingerprint database

Fingerprint collection by the crowdsensing technique has implications on the density and the received signal strength characteristics.

The density of fingerprints influences the accuracy of the crowdsensed positioning system and is closely linked to the number of people that have visited a certain area. Fig. 4 displays the density of the gathered fingerprints using an interpolated graph overlayed on the test area map. Each colored dot signifies the presence of a number of APs in the area around it, the color of the dot ranging between blue - one AP and red - 70 APs.

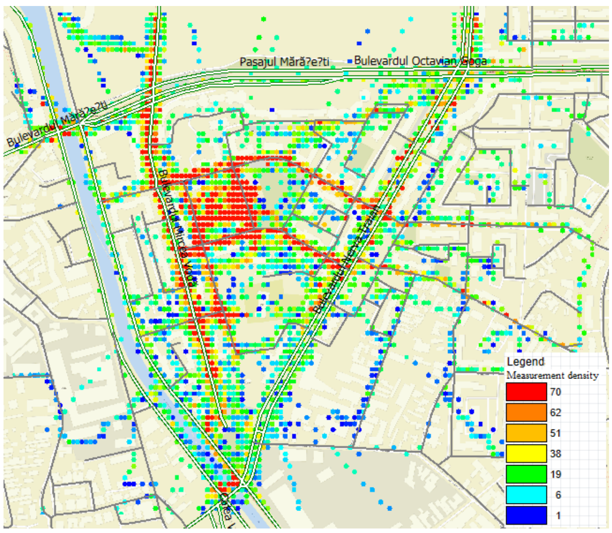

Figure 4. Density of crowdsensed APs in the test area

Peer-reviewed version available at Sensors 2020, 20, 427; doi:10.3390/s20020427

Table 1. Frequency of measurements for the APs in the test area

Frequency	APs measured
1	3507
2	1446
3	791
4	575
5	479
6	384
7	297
8	229
9	202
10 - 19	1137
20 - 29	492
30 - 39	232
40 - 49	108
50 - 59	60
60 - 69	50
70 - 79	34
80 - 89	21
90 - 99	10
100 - 109	11
>110	7

The average signal level (dBm) of the fingerprints in the test area is shown in Fig. 5. The signal map is generated using the interpolation technique with the average being calculated for a 10-meter radius.

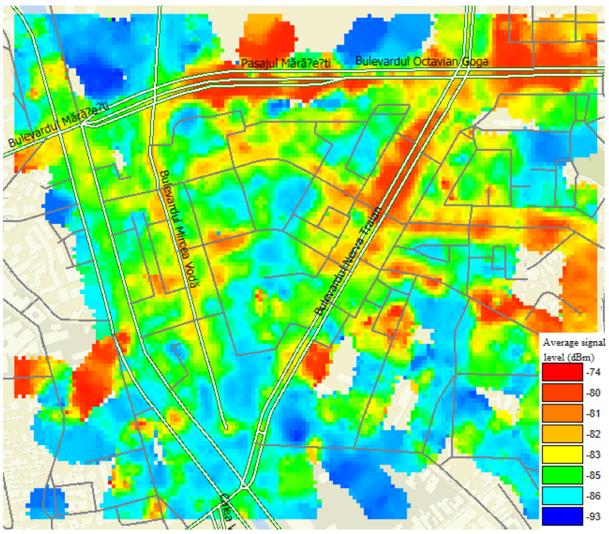

Figure 5. Map of average signal level for fingerprints in the test area

The placement of the AP, for example, upper in a building, and the crowdsensing technique will influence the signal level at ground level where the measurements were collected. APs placed on lower floors or at ground level have the chance to be scanned more often and be seen with a higher signal level, as shown in Tab. 1.

Fig. 6 is obtained by plotting the density of APs scanned less than three times. By comparing Fig. 4 and Fig. 6 we observe that APs scanned less than three times were observed at the margins of the test area where the overall density of measurements is lower. This fact leads to the following observations:

- APs in densely scanned areas are more likely to be themselves scanned more often;

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2019 doi:10.20944/preprints201911.0186.v1

Peer-reviewed version available at Sensors 2020, 20, 427; doi:10.3390/s20020427

- the crowdsensing method offers consistent results for successive scans of the same area;
- APs radio visibility displays a character of stability for repeated scans.

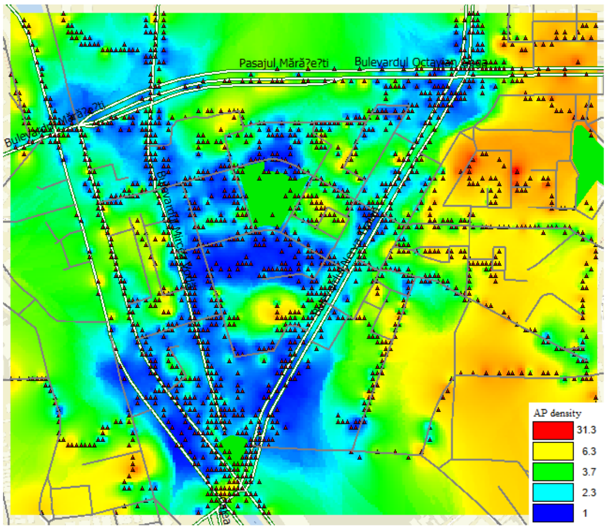

Figure 6. Density map for APs scanned less than three times

Access points that were scanned less than three times can increase the positioning error significantly, as they are most likely placed at a great distance from the measurement points or are outside the test area.

5. Positioning accuracy and significant error sources generated by crowdsensing

The metric that is used for the evaluation of the positioning accuracy of the crowdsensed data is the mean 2D error. The error is calculated as the difference between the estimated position and the known position of the test data sets.

As mentioned, the evaluation of the positioning accuracy in an urban test environment for the crowdsensed data is done by the help of the weighted-centroid method [28] and the log-Gaussian distance method [10]. The initial results for both algorithms when applied to the complete test area fingerprint database show the following mean 2D error:

- 30,017 meters for the weighted centroid method;
- 45,414 meters for the log-Gaussian distance method.

These mean error values are significantly higher than those reported in previous research where the test area has a smaller surface and was placed indoor, for example, approximately 10 meters in reference [9] or approximately 6 meters in reference [29]. When compared to previous results reported in papers evaluating positioning accuracy in urban environments, the attained results using crowdsensed data are similar, for example, reference [30] reports an average positioning accuracy between 23.5 and 36 meters while reference [31] reports errors between 35 and 120 meters. When comparing indoor fingerprinting positioning research with outdoor fingerprinting positioning research results we observe that the mean errors of outdoor systems are by an order of magnitude larger than those of indoor systems.

The mean error of outdoor fingerprinting positioning system can be reduced using database processing, as proposed in references [6,32,33] which publish results closing a 15-meter average positioning error. The fingerprint database processing in the referenced papers involved:

- eliminating the fingerprints of APs with an estimated coverage radius larger than 300 meters;
- eliminating the fingerprints of APs suspected to be mobile;
- eliminating the fingerprints collected using inaccurate GPS positioning.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2019 doi:10.20944/preprints201911.0186.v1

Peer-reviewed version available at Sensors 2020, 20, 427; doi:10.3390/s20020427

5.1. Characterization of the positioning algorithms

Each method has its characteristics which will influence the resulting positioning error and also the way it benefits from the crowdsensed fingerprints. The implementation of the weighted centroid method requires an initial step preceding the actual positioning that requires the estimation of the position of all the AP's in the fingerprint database. Using crowdsensed data for AP position estimation can easily lead to errors in estimating the AP's exact location. This is to be expected as crowdsensed fingerprints are collected alongside roads or near buildings. The initial step mandates the study of the influence number of crowdsensed fingerprints for each AP.

The log-Gaussian method will aim at finding the most similar fingerprints with the measurement vector used for positioning by the help of the log-Gaussian distance. The position estimation is then done using the k-Nearest Neighbor (k-NN) approach on the k most similar fingerprints. For the evaluation of the method, the number of neighbors is set to $k = 3$. Varying the number of neighbors will influence the results of the position estimation algorithm. For example, an increase in the value of k may lead to an estimated position further away from the test position due to the inclusion of less similar fingerprints that were collected at a greater distance. This can be solved by applying a weight to the k-NN algorithm or by evaluating the influence of the crowdsensed data on the position estimation algorithm.

The particularities of each positioning method coupled with the less-known influence of the crowdsensed fingerprints prompted us to research the following aspects which influence the accuracy of crowdsensed fingerprinting positioning:

- the influence of mobile APs;
- the average number of fingerprints for each APs;
- the signal level of the fingerprints;
- the average signal level of the APs;
- the estimated coverage radius of the APs;
- the number of neighbors used for the k-NN estimation.

5.2. Influence of mobile APs

Mobile APs can lead to significant positioning errors. For example, an AP installed in a public transport vehicle could lead to positioning estimation results placed in the part of the town where the AP was scanned. This prompts us to apply a method for removing the fingerprints of mobile APs, while also investigating the effects of their presence on the overall positioning performance.

The identification of mobile APs requires the analysis of the whole fingerprint database, not only of the test area. Mobile APs are classified using a method previously published by the authors in reference [6], with the following results:

- 3506 APs representing 34.8 % of the total are scanned once and can't be classified as fixed or mobile. thus they are eliminated from the comparative data-set;
- 395 APs representing 3.9 % of the total are classified as mobile and are eliminated from the comparative data-set.

Table 2 displays the resulting average positioning error for both positioning methods when applied to the data-set with mobile APs removed. The weighted-centroid method displays a minor increase in the average positioning error, caused by the elimination of APs seen only once while the log-Gaussian method benefits from the removal of the mobile APs.

5.3. Influence of the average number of fingerprints collected for each AP

The average number of fingerprints/AP could influence both positioning methods when considering crowdsensing as the collection method. A low number of fingerprints can limit the information available in the AP position estimation step characteristic to the weighted-centroid

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2019 doi:10.20944/preprints201911.0186.v1

Peer-reviewed version available at Sensors 2020, 20, 427; doi:10.3390/s20020427

Table 2. 2D positioning error for the data-set without mobile APs

Data-set	Weighted-centroid mean error (m)	log-Gaussian mean error (m)
Full data-set	30,017	45,414
Comparative data-set	34,245 (-14%)	28,180 (+37%)

Table 3. 2D positioning error influenced by the average fingerprint to AP ratio

Data-set	Weighted-centroid mean error (m)	log-Gaussian mean error (m)
Full data-set	30,017	45,414
Reduced fingerprint data-set	26,916 (+13%)	92,198 (-103%)
Average fingerprint data-set	37.896 (-20%)	31.147 (+30%)
Dense fingerprint data-set	49.439 (-64%)	38.274 (+15%)
Very dense fingerprint data-set	76.055 (-153%)	65.929 (-45%)

method while a large number of fingerprints correlated to a large coverage area might reduce the unique character of each fingerprint leading to an increase in the positioning error for both methods.

To study the previous hypothesis the data-set was split into sub-sets using the average fingerprint number for each APs as the criteria as following:

- reduced fingerprint data-set containing APs averaging 1 to 5 fingerprints resulting in 13647 fingerprints of 6798 APs;
- average fingerprint data-set containing APs averaging 6 to 20 fingerprints resulting in 24890 fingerprints of 2318 APs;
- dense fingerprint data-set containing APs averaging 21 to 60 fingerprints resulting in 26957 fingerprints of 828 APs;
- very dense fingerprint data-set containing APs averaging more than 61 fingerprints resulting in 10234 fingerprints of 128 APs.

The resulting mean positioning error for each sub-data-set is shown in table 3.

Contrary to expectations the weighted centroid displays increased precision for the reduced fingerprint data-set. This is caused by a higher positioning accuracy in scenarios where the estimated position of the APs in the positioning measurement are placed closely. For the data-sets with a higher fingerprint to AP ratio, the precision is lower, mainly due to a decrease in the AP position estimation caused by the higher number of fingerprints. The log-Gaussian method offers worse result for the reduced and very-dense data-sets and significantly better accuracy for the intermediate values. A low number of fingerprints for APs means that less data is available for the statistic characterization of the data while a large number of fingerprints often leads to the phenomena of geographic diffusion of precision where similar signal levels are found in fingerprints placed far away from each other.

Very dense fingerprint to AP ratios lead to a decrease in positioning accuracy leading to the conclusion that collecting and storing large volumes of fingerprints is not always beneficial to the fingerprint positioning system.

5.4. Influence of fingerprint signal level

The signal level influence on positioning accuracy was studied by splitting the data-set into sub-sets using the value of the signal level in dBm as a criterion. The sub-set split aimed at keeping the number of fingerprints in each sub-set similar. The results of the analysis are displayed in table 4.

As shown in table 4, the log-Gaussian method accuracy is significantly lower than that of the full data-set. This is due to the usage of the log-Gaussian distance for determining the similarity between fingerprints and positioning measurements. The similarity function is less likely to achieve optimum performance when comparing data within a 5 dBm interval. The weighted-centroid method offers

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2019 doi:10.20944/preprints201911.0186.v1

Peer-reviewed version available at Sensors 2020, 20, 427; doi:10.3390/s20020427

Table 4. 2D positioning error influenced by fingerprint signal level

Data-set	Weighted-centroid mean error (m)	log-Gaussian mean error (m)
Full data-set	30,017	45,414
Sub-set under -90 dBm	32.394 (-8%)	118.566 (-161%)
Sub-set between -86 and -90 dBm	37.574 (-25%)	78.873 (-74%)
Sub-set between -81 and -85 dBm	31.318 (-4%)	71.872 (-58%)
Sub-set between -76 and -80 dBm	24.784 (+17%)	82.650 (-81%)
Sub-set between -71 and -75 dBm	17.922 (+40%)	122.574 (-169%)
Sub-set over -70 dBm	15.489 (+48%)	136.788 (-201%)

Table 5. 2D positioning error influenced by average AP signal level

Data-set	Weighted-centroid mean error (m)	log-Gaussian mean error (m)
Full data-set	30,017	45,414
Sub-set under -90 dBm	32.394 (-8%)	118.566 (-161%)
Sub-set between -86 and -90 dBm	36.922 (-23%)	46.869 (-3%)
Sub-set between -81 and -85 dBm	42.766 (-42%)	40.885 (+10%)
Sub-set between -76 and -80 dBm	46.991 (-56%)	68.903 (-51%)
Sub-set between -71 and -75 dBm	32.799 (-9%)	166.648 (-266%)

insights regarding the influence of the signal level on positioning accuracy as it displays a lower error than that of the full data-set for fingerprints with values above -80 dBm.

5.5. Influence of average AP signal level

The average AP signal level is calculated as the average level of all collected fingerprints of each AP. Again, the data-set was split into sub-sets using the average signal level as the criteria, aiming at keeping the number of fingerprints in each set similar. The results of the analysis are displayed in table 5.

As fewer APs are included in each sub-set due to the filtering, the weighted-centroid method will perform worse in each scenario. The log-Gaussian displays a minor accuracy improvement for APs with an average signal level close to the average signal level of all collected fingerprints.

5.6. Influence of estimated AP coverage radius

The AP coverage radius is estimated by the help of the crowdsensing process. The scanning process continued for five months inside the test area allowing us to assume that street-level coverage was captured with precision. The average coverage radius of APs in the test area is approximately 47 meters. The coverage area was determined by the help of the Convex Hull operation, specific to database management systems. The data-set was split into sub-sets using the estimated coverage radius as the criteria, the results of the accuracy analysis being displayed in table 6. The sub-set split aimed at keeping a similar number of fingerprints for the sub-sets. This was possible for APs with a coverage radius under 80 meters. APs with estimated coverage radius larger than 80 meters were less numerous in the test area.

The resulting sub-sets based on the coverage radius have the following dimensions:

- Null radius - 6506 fingerprints of 4988 APs;
- Radius between 0 and 30 meters - 17618 fingerprints of 2634 APs;
- Radius between 30 and 50 meters - 19603 fingerprints of 1328 APs;
- Radius between 50 and 80 meters - 18659 fingerprints of 774 APs;
- Radius between 80 and 110 meters - 8074 fingerprints of 250 APs;
- Radius between 110 and 170 meters - 3665 fingerprints of 81 APs;

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 November 2019 doi:10.20944/preprints201911.0186.v1

Peer-reviewed version available at Sensors 2020, 20, 427; doi:10.3390/s20020427

Table 6. 2D positioning error influenced by average AP radius

Data-set	Weighted-centroid mean error (m)	log-Gaussian mean error (m)
Full data-set	30,017	45,414
Null radius	11.033 (+63%)	278.320 (-512%)
Radius between 0 and 30 meters	24.385 (+18%)	21.428 (+47%)
Radius between 30 and 50 meters	33.744 (-12%)	28.127 (+38%)
Radius between 50 and 80 meters	47.480 (-58%)	38.080 (+16%)
Radius between 80 and 110 meters	69.585 (-131%)	57.914 (-27%)
Radius between 110 and 170 meters	94.099 (-213%)	75.373 (-65%)
Radius larger than 110 meters	111.839 (-272%)	97.662 (-115%)

- Radius larger than 110 meters - 4728 fingerprints of 98 APs;

The positioning performance of the weighted-centroid method benefits from the usage of APs with an estimated coverage radius under 30 meters, due to its initial step which requires the estimation of the AP position. A lower radius AP is less likely to be positioned erroneously.

The log-Gaussian method offers significantly greater accuracy for APs with the estimated radius under 50 meters. For both methods, an AP coverage radius larger than 80 meters will result in larger positioning errors related to the geographic dilution of precision phenomena.

The analysis allows us to gain insight that can help set an upper limit on the estimated coverage AP area which in turn can limit the volume and effort related to the crowdsensing process.

5.7. Influence of the number of neighbors specific to the k-NN method

The k-NN method is often used for position estimation by fingerprinting systems. The k-NN method is combined with the log-Gaussian distance to select the most similar neighbors in the fingerprint space. Previous papers have proposed different values for k, according to the environment that the fingerprinting positioning system was to cover. References [20,34] propose that $k = 3 or 4$ for indoor fingerprinting systems.

The analysis of the number of neighbors influence on the accuracy of the positioning is done by varying the number between 1 and 10 and by computing the average 2D positioning error in each scenario. The data-set used contains 15002 fingerprints of 1929 APs in the test area with an average estimated radius between 0 and 30 meters. The results are displayed in Fig. 7.

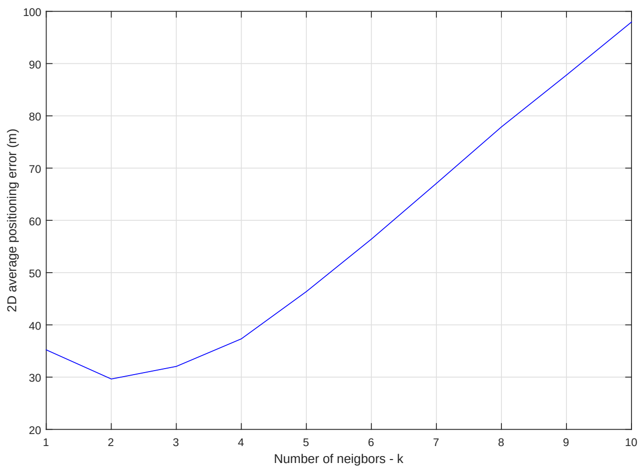

Figure 7. k-NN average 2D positioning error influenced by the number of neighbors

According to Fig. 7, the average positioning error is minimum for $k = 2$. The result is explained mainly by the data collection method. Collecting Wi-Fi measurements using crowdsensing means that most fingerprints will be captured alongside public roads or in the proximity of buildings, while most APs will be positioned indoor. Using three or more neighbors in the fingerprint space will likely lead to an estimation that is biased towards the two closest fingerprints and as Fig. 7 shows farther than the real position of the user. When $k = 2$ the positioning error is minimum due to the probability that the real position is alongside the road that the two fingerprints were collected.

6. Discussion

Our analysis shows that the positioning accuracy of outdoor fingerprinting systems can be significantly improved by filtering and optimizing the contents of the fingerprint database. The best results come close to an average 2D positioning error of 21 meters for the log-Gaussian distance method. These values do not come close to that achieved by previously published research on indoor systems. Nevertheless, our results show that urban fingerprinting positioning can reach GPS-like accuracy in similar use-cases. Reference [35], cited on the official *gps.gov* website states that GPS has an average positioning accuracy of 5 meters in rural areas and approximately 16.8 in urban environments.

The accuracy of crowdsensed fingerprinting positioning systems is shown to be dependent on the density of fingerprints quantified as the average number of fingerprints collected for each AP. As the average positioning error was computed for the whole test-area, including areas that were scanned less often and dense, it is safe to assume that for some of the most scanned roads inside the test-area the precision of the fingerprinting system will outperform that of the GPS.

7. Conclusion

Wi-Fi fingerprint positioning systems for urban environments have become an intense topic of research in the context of Internet-of-Things technology. These positioning systems have the capabilities of offering GPS-like positioning accuracy for energy-constrained devices equipped with a Wi-Fi or 5G radio network interface that is used both for communications and positioning.

To evaluate the performance of urban fingerprint positioning systems, we build our fingerprint database by the help of the crowdsensing technique for data gathering. Since the crowdsensing effort makes it hard to cover a whole city with a small number of contributors we choose to conduct our analysis on a dedicated test area specific to the Bucharest urban environment.

We analyze the crowdsensed fingerprint and we highlight important aspects related to the research problem including the presence of mobile APs, the influence of the average number of fingerprints collected per AP, the fingerprint signal level, the average street-level AP signal level, the influence of the estimated AP coverage radius and the performances of the k-NN method.

We conclude by proving that fingerprint positioning systems can achieve or even out-perform satellite positioning systems in an urban environment, mostly due to the density of Wi-Fi APs and the difficulties that other positioning systems have in the conditions.

Author Contributions: C.L.L. conceived the work, was responsible for data collection and validation, analyzed the data and wrote the paper. I.N. described the methodology, supervised the research and revised the paper. P.C. revised the work and was responsible for project administration and funding acquisition.

Funding: The work was funded by the Romanian Government UEFISCDI 9SOL project (PN-III-P2-2.1-SOL-2017-08-010).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Han, Q.; Liang, S.; Zhang, H. Mobile cloud sensing, big data, and 5G networks make an intelligent and smart world. *IEEE Network* **2015**, *29*, 40–45.

2. Choudhury, T.; Borriello, G.; Consolvo, S.; Haehnel, D.; Harrison, B.; Hemingway, B.; Hightower, J.; Koscher, K.; LaMarca, A.; Landay, J.A.; others. The mobile sensing platform: An embedded activity recognition system. *IEEE Pervasive Computing* **2008**, *7*, 32–41.
3. Pei, L.; Zhang, M.; Zou, D.; Chen, R.; Chen, Y. A survey of crowd sensing opportunistic signals for indoor localization. *Mobile Information Systems* **2016**, *2016*.
4. Zhou, X.; Chen, T.; Guo, D.; Teng, X.; Yuan, B. From one to crowd: A survey on crowdsourcing-based wireless indoor localization. *Frontiers of Computer Science* **2018**, *12*, 423–450.
5. Leca, C.L.; Ciotirnae, P.; Rincu, C.I.; Nicolaescu, I. Characteristics of crowdsourcing for outdoor radio fingerprinting positioning. Electronics, Computers and Artificial Intelligence (ECAI), 2017 9th International Conference on. IEEE, 2017, pp. 1–4. https://doi.org/10.1109/ECAI.2017.8166385, WOS:000425865900001.
6. Leca, C.L. Ubiquity of Wi-Fi: Crowdsensing Properties for Urban Fingerprint Positioning. *Advances in Electrical and Computer Engineering* **2017**, *17*, 131–136. https://doi.org/10.4316/AECE.2017.04016, WOS:000417674300016.
7. Leca, C.L. Overview of Romania 802.11 wireless networks security. Electronics, Computers and Artificial Intelligence (ECAI), 2017 9th International Conference on. IEEE, 2017, pp. 1–4. https://doi.org/10.1109/ECAI.2017.8166386, WOS:000425865900002.
8. Liu, J.; Shen, H.; Narman, H.S.; Chung, W.; Lin, Z. A survey of mobile crowdsensing techniques: A critical component for the internet of things. *ACM Transactions on Cyber-Physical Systems* **2018**, *2*, 18.
9. Lohan, E.S.; Torres-Sospedra, J.; Leppäkoski, H.; Richter, P.; Peng, Z.; Huerta, J. Wi-Fi Crowdsourced Fingerprinting Dataset for Indoor Positioning. *Data* **2017**, *2*, 32. https://doi.org/10.3390/data2040032, WOS:000424475500001.
10. Mendoza-Silva, G.M.; Richter, P.; Torres-Sospedra, J.; Lohan, E.S.; Huerta, J. Long-Term WiFi Fingerprinting Dataset for Research on Robust Indoor Positioning. *Data* **2018**, *3*, 3. https://doi.org/10.3390/data3010003, WOS:000428778200003.
11. Gentile, C.; Alsindi, N.; Raulefs, R.; Teolis, C. *Geolocation techniques: principles and applications*; Springer Science & Business Media, 2012.
12. Wang, J.; Luo, J.; Pan, S.J.; Sun, A. Learning-Based Outdoor Localization Exploiting Crowd-Labeled WiFi Hotspots. *IEEE Transactions on Mobile Computing* **2019**, *18*, 896–909. doi:10.1109/TMC.2018.2849416.
13. Basiri, A.; Lohan, E.S.; Moore, T.; Winstanley, A.; Peltola, P.; Hill, C.; Amirian, P.; e Silva, P.F. Indoor location based services challenges, requirements and usability of current solutions. *Computer Science Review* **2017**, *24*, 1–12.
14. Wu, C.; Yang, Z.; Liu, Y.; Xi, W. WILL: Wireless indoor localization without site survey. *IEEE Transactions on Parallel and Distributed Systems* **2012**, *24*, 839–848.
15. Liu, H.; Gan, Y.; Yang, J.; Sidhom, S.; Wang, Y.; Chen, Y.; Ye, F. Push the limit of WiFi based localization for smartphones. Proceedings of the 18th annual international conference on Mobile computing and networking. ACM, 2012, pp. 305–316.
16. Sun, W.; Liu, J.; Wu, C.; Yang, Z.; Zhang, X.; Liu, Y. MoLoc: On distinguishing fingerprint twins. 2013 IEEE 33rd international conference on distributed computing systems. IEEE, 2013, pp. 226–235.
17. Xiao, Z.; Wen, H.; Markham, A.; Trigoni, N.; Blunsom, P.; Frolik, J. Non-line-of-sight identification and mitigation using received signal strength. *IEEE Transactions on Wireless Communications* **2014**, *14*, 1689–1702.
18. Lui, G.; Gallagher, T.; Li, B.; Dempster, A.G.; Rizos, C. Differences in RSSI readings made by different Wi-Fi chipsets: A limitation of WLAN localization. 2011 International Conference on Localization and GNSS (ICL-GNSS). IEEE, 2011, pp. 53–57.
19. Sapiezynski, P.; Stopczynski, A.; Gatej, R.; Lehmann, S. Tracking human mobility using wifi signals. *PloS one* **2015**, *10*, e0130824. https://doi.org/10.1371/journal.pone.0130824, WOS:000358153000083.
20. Bahl, P.; Padmanabhan, V.N. RADAR: An in-building RF-based user location and tracking system. INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE. IEEE, 2000, Vol. 2, pp. 775–784. https://doi.org/10.1109/INFCOM.2000.832252.
21. Youssef, M.; Agrawala, A. The Horus WLAN location determination system. Proceedings of the 3rd international conference on Mobile systems, applications, and services. ACM, 2005, pp. 205–218. https://doi.org/10.1145/1067170.1067193, WOS:000230943400016.

22. Bolliger, P. Redpin-adaptive, zero-configuration indoor localization through user collaboration. Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments. ACM, 2008, pp. 55–60.
23. Park, J.g.; Charrow, B.; Curtis, D.; Battat, J.; Minkov, E.; Hicks, J.; Teller, S.; Ledlie, J. Growing an organic indoor location system. Proceedings of the 8th international conference on Mobile systems, applications, and services. ACM, 2010, pp. 271–284.
24. Ferris, B.; Fox, D.; Lawrence, N.D. WiFi-SLAM Using Gaussian Process Latent Variable Models. IJCAI, 2007, Vol. 7, pp. 2480–2485. https://www.semanticscholar.org/paper/WiFi-SLAM-Using-Gaussian-Process-Latent-Variable-Ferris-Fox/600376bb397c5f3d8d33856b3a6db314b6bda67d, WOS:000283721000398.
25. Rai, A.; Chintalapudi, K.K.; Padmanabhan, V.N.; Sen, R. Zee: Zero-effort crowdsourcing for indoor localization. Proceedings of the 18th annual international conference on Mobile computing and networking. ACM, 2012, pp. 293–304.
26. Wang, H.; Sen, S.; Elgohary, A.; Farid, M.; Youssef, M.; Choudhury, R.R. No need to war-drive: unsupervised indoor localization. Proceedings of the 10th international conference on Mobile systems, applications, and services. ACM, 2012, pp. 197–210. https://doi.org/10.1145/2307636.2307655.
27. Jung, S.; Lee, G.; Han, D. Methods and Tools to Construct a Global Indoor Positioning System. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **2018**, *48*, 906–919. doi:10.1109/TSMC.2016.2626797.
28. Laitinen, E.; Lohan, E.S. On the choice of access point selection criterion and other position estimation characteristics for WLAN-based indoor positioning. *Sensors* **2016**, *16*, 737. https://doi.org/10.3390/s16050737, WOS:000378510400107.
29. Wu, C.; Yang, Z.; Liu, Y. Smartphones based crowdsourcing for indoor localization. *IEEE Transactions on Mobile Computing* **2015**, *14*, 444–457. https://doi.org/10.1109/TMC.2014.2320254, WOS:000347098100016.
30. Li, B.; Quader, I.J.; Dempster, A.G. On outdoor positioning with Wi-Fi. *Positioning* **2008**, *1*, 0. https://doi.org/10.5081/jgps.7.1.18.
31. Quader, I.J.; Li, B.; Peng, W.; Dempster, A.G. Use of fingerprinting in Wi-Fi based outdoor positioning. *Proc IGNSS, Sydney* **2007**. https://pdfs.semanticscholar.org/c6fd/f92a834da0f3f580c454af7ed8953d331ed7.pdf.
32. Wang, J.; Tan, N.; Luo, J.; Pan, S.J. WOLoc: WiFi-only outdoor localization using crowdsensed hotspot labels. INFOCOM 2017-IEEE Conference on Computer Communications, IEEE. IEEE, 2017, pp. 1–9. https://doi.org/10.1109/INFOCOM.2017.8057096, WOS:000425232200153.
33. Sapiezynski, P.; Gatej, R.; Mislove, A.; Lehmann, S. Opportunities and challenges in crowdsourced wardriving. Proceedings of the 2015 ACM Conference on Internet Measurement Conference. ACM, 2015, pp. 267–273. https://doi.org/10.1145/2815675.2815711, WOS:000380613200022.
34. Torres-Sospedra, J.; Moreira, A. Analysis of Sources of Large Positioning Errors in Deterministic Fingerprinting. *Sensors* **2017**, *17*, 2736. https://doi.org/10.3390/s17122736, WOS:000423285800036.
35. Van Diggelen, F.; Enge, P. The worlds first gps mooc and worldwide laboratory using smartphones. Proceedings of the 28th international technical meeting of the satellite division of the institute of navigation (ION GNSS+ 2015), 2015, pp. 361–369.