Comparison of Anchor Selection Algorithms for Improvement of Position Estimation During the Wi-Fi Localization Process in Disaster Scenario

Oleksandr Artemenko*, Tobias Simon*, Andreas Mitschele-Thiel*, Dominik Schulz†, Muhammad Rheza Satria Ta‡

* Integrated Communication Systems Group, Ilmenau University of Technology, 98693 Ilmenau, Germany
Email: {Oleksandr.Artemenko, Tobias.Simon, Mitsch}@tu-ilmenau.de
† Communication Networks Research Laboratory, Ilmenau University of Technology, 98693 Ilmenau, Germany
Email: Dominik.Schulz@tu-ilmenau.de
‡ Swiss German University, Tangerang 15339, Indonesia
Email: rheza.satria.ta@gmail.com

Abstract—It is very common that during the localization process redundant reference data can be collected. This reference data is called “anchors”. For some position estimation algorithms, like trilateration or Min-Max, only three or four anchors for 2-d or 3-d localization respectively are required. The selection of certain anchors from collected amount can affect the localization accuracy significantly [1]. The main purpose of this paper is to propose and analyze several most promising Optimum Anchor Selection algorithms (OASIS) that can contribute to the improvement of the localization precision.

This paper deals with the position estimation of the “survived” Wi-Fi access points in the scope of the disaster scenario. All analyzed algorithms use trilateration for position calculation and run on the real data that was obtained during the measurement campaign in the university campus in Ilmenau (Germany) using flying multicopters. Additionally, the proposed algorithms will be compared with the results calculated with the multilateration method.

Therefore, the main purpose of this paper is to find the most efficient OASIS method in scope of localization accuracy and speed.

I. INTRODUCTION

There are numerous existing applications that require object localization and many others are constantly emerging. One among them - localization of Wi-Fi access points in the scope of disaster scenarios, which may facilitate people rescue - is being considered in this paper.

Much scientific research has been conducted in the area of localization techniques. However, none of them may be applied to our scenario with no or little modification. They are either fundamental theoretical works or very proprietary studies in the field of wireless networks (WNs), whose underlying assumptions and constraints are not applicable to our tasks. Additionally, there are only few investigations that consider the localization of indoor objects from outdoor environment. This scenario provides specific requirements on the localization procedure and can be described as follows.

An unmanned aerial vehicle (UAV) flies over an urban area, which suffers from disruption of the communications infrastructure as a result of a disastrous event, and samples 802.11 beacon frames. The objective is to localize the survived Wi-Fi infrastructure in the area of interest, analyze coverage gaps and eventually repair the network by placing additional relay nodes connected to UAVs. As a result, connectivity to the outer world may be provided for people, who could be blocked under collapsed constructions.

The received signal strength (RSS) measurements occur at the rate of 1 sample per second. Although this method has little reliability because of poorly predictable radio channel effects, it eliminates the need for extra circuitry complexity and results in lower hardware costs and energy consumption [2]. Moreover, it can be easily implemented with existing off-the-shelf wireless equipment.

This paper focuses explicitly on the selection of anchors which are likely to enhance localization performance. The RSS is used to estimate distances to access points. Resulting position is to be solved in real time using any of position estimation techniques like trilateration or multilateration concerning the database of anchors reaching hundreds of readings.

A. State of the Art

The study on anchor selection has shown that many researchers in the area of wireless networks address the issue of anchor placement and selection. However, their initial assumptions are not feasible under given circumstances. As an example of such, it can be mentioned anchor placement along a perimeter [3] which assumes a priori Wireless Sensor Network (WSN) deployment planning and localization information exchange in range-free algorithms [4].

Some approaches that utilize geometrical layout of anchor nodes are applicable to our study to some extent. The most notable example of the class is the dilution of precision (DOP), widely used in GPS nowadays [5]. It is a single dimensionless number that reflects how the selected satellite constellation affects the localization accuracy. More precisely, in GPS, it is defined as GDOP (geometrical DOP) and comprises a
combined effect of 3-d position DOP (PDOP) and time DOP (TDOP).

On one side, we deal in our case only with position estimation in 2-d plane and therefore the horizontal term of DOP is of the most significance. On the other side, existing solutions [6] cannot provide computational advantages if only one single term is taken into consideration. GDOP can be either calculated as a coherent parameter with closed-form formulae, or all of the components appear in the solution as in case of the conventional matrix inversion method. GDOP approximation with artificial neural networks require a costly training phase and are suitable only when used in close proximity to the geographical location where learning took place [6].

To calculate a GDOP value for a certain point it is yet necessary to know its position. In GPS applications, a rough initial position estimate produces in most cases reasonable results because of large distances between satellites and a user. Meanwhile within small Wi-Fi ranges and considerable measurements errors, this metric should be regarded as unsuitable.

Another simple variation on GDOP in terms of the angles between the anchors and the sensor node is given by [7]:

\[
GDOP(M, \phi) = \sqrt{\frac{M}{\sum_i \sum_{j>i} |\sin(\phi_{ij})|^2}},
\]

(1)

where \(M\) is the number of anchors involved in the localization process and \(\phi\) is the angle between each pair of anchors. But, according to the Wi-Fi localization scenario and considering the fact that all readings can appear only from one side of the node to be found, this expressions cannot be applied without corrections.

A simple convex-hull approach is also proven to contribute to localization accuracy [8]. However, there are only heuristic solutions of finding a subset of nodes which would form a convex-hull, e.g. the convex-hull detection method described in [9]. Besides, it suffers from the same limitation as the GDOP metric, i.e. necessity of the initial position estimate.

In addition to this, there exist some works that have been derived from solely geometrical perspective on the mutual position of the anchors. For instance, in [10] it is constituted that if anchors’ constellation forms an equilateral triangle localization accuracy is improved by 34.9% over the random anchor placement. Another work [11] suggests the optimum placement patterns with the number of references greater than three. The authors note that for a small number of anchors simple shapes like equilateral triangles and squares result in better localization performance. Furthermore, for a higher number of anchors such shapes should enclose one another. Extensions of shapes with equal sides like hexagons etc. would not yield optimal placement. The downside is still, that a search for a subset of anchors that form such shapes may be exhaustive and needs further examination and modification.

An intuitive approach was proposed in [12]: joint clustering technique. It chooses anchors with the strongest signal strength readings by grouping locations that observe the same set of strongest anchors into the same cluster. Another approach, information gain-based anchor selection method (InfoGain [13]), calculates the entropy of each anchor and selects the anchors with the strongest discriminative power. One drawback of these two algorithms is the need of choosing the efficient number of clusters \(k\) depending on environment conditions, which has strong affect on the accuracy of location estimation. Concerning our scenario, it is almost impossible to predict the situation after a disaster has happened. For this reason, we cannot rely on these methods without testing them.

B. Aim

Concerning the above motivation, the aim of this paper can be defined as follows. Different Optimum Anchor Selection algorithms (OASIS) are being introduced and tested on the real data that was obtained during the measurement campaign in the university campus using flying UAVs. The obtained results will be analyzed and compared with the well-known multilateration technique.

The choice of the most effective OASIS method in scope of localization accuracy and speed will be explained.

C. Paper Organization

The rest of this paper is organized as follows. In Sect. II, we introduce the OASIS methods chosen for our disaster scenario. Section III presents the experimental evaluation with the description of our working environment, used propagation model, obtained results and analysis of them. Finally, Sect. IV concludes this paper.

II. OASIS METHODS

In this section, we briefly describe all anchor selection methods that will be compared in our current research work. Some of the presented methods were found in the literature. In this case, they have been adopted to suit our scenario: Joint Clustering Technique [12] and Simple Convex Hull [8]. Some methods present our own intuitive proposals: Weighted Filter, Signal Strength (SS). The rest of the methods have been derived from analysis of the anchor placement algorithms found in [1]: Area, Perimeter, Density and AngleSS. Additionally, a brief description of Multilateration will be given.

A. Joint Clustering Technique

The main idea of this method is to choose \(k\) anchors whose observed signal strength are the top \(k\) strongest among a set of available anchors to perform localization [12].

B. Simple Convex Hull

In [8], a simple and an advanced convex hull techniques for selection of reference nodes in WSN were presented. We selected the simple hull to implement in our localization system because, according to the results obtained by the authors, simple hull shows better accuracy.

The simple convex hull considers only the distance metric to elect hull nodes. For \(N\) nodes \(n_1, ..., n_N\), the convex hull \(C\) can be given as the following expression [8]:
\( C \equiv \left\{ \sum_{j=1}^{N} \lambda_j n_j : \lambda_j \geq 0 \text{ for all } j \text{ and } \sum_{j=1}^{N} \lambda_j = 1 \right\} . \)

C. Weighted Filter

With this algorithm, we weight each new reading according to its signal strength level and collected previously history of received data. The readings are weighted corresponding to the defined threshold of -65 dBm signal level and the output ratio of 67% readings with strong signal (above -65 dBm) and 33% with weak signal (under -65 dBm). After the filter produced an amount of at least 3 sequences, data will be fed to trilateration. If more than 3 readings were selected, the Area algorithm will be applied on them.

D. Signal Strength

The signal strength-based method presents the most popular approach to filter the received data. With this algorithm, we select the three last strongest readings from the amount of obtained data.

E. Area

According to [14], the larger the area formed by four satellites is, the better (lower) the value of GDOP coefficient is. Reversely, the smaller area results in the worse (higher) value of GDOP. Applying this fact to our scenario and concerning the idea of simple shapes from [11], we calculate the area of all possible triangles created by triples of anchors and select the biggest one. The area of triangle can be seen as a measure for the linear independence of points in 2-d or 3-d space.

F. Perimeter

This method presents similar idea as “Area” but calculates the perimeter of triangles. It may also serve as a measure for linear independence. However, in certain scenarios, all three points of a triangle can be separated from each other but lie on the same straight line. In this case, the perimeter does not reflect the depth between the vectors.

G. Density

At the beginning, the algorithm goes through all possible combinations of anchors that form triangles (brute-force approach) and using the trilateration method calculates all possible locations of the node being localized. Subsequently, the area with the highest density and thus biggest probability will be determined. The middle point of this area yields the localization result.

H. Combination of Angle and Signal Strength

According to [11], localization error is the least when three reference nodes form an equilateral triangle. Using this fact, we propose an intelligent algorithm that combines two methods: the anchor selection based on angles (see (1)) with the selection based on signal strength.

The main idea of the algorithm is to divide the area around the node, which position is to be found, into three sections, each \( \frac{\pi}{3} \), and to select three anchors from each section with the strongest signal strength.

1. Multilateration

Each anchor selection scheme presented in this work is based on the trilateration algorithm. That is, positions are inferred from distance estimates given by three reference points in a two dimensional space. One may also use more than three reference points. In fact, suppose we have \( M \) reference points in an \( N \)-dimensional space. In this case, any subset of reference points with at least \( N + 1 \) distinct points can be used for localization. Consequently, it is possible to have anchor selection algorithms that use more than \( N+1 \) reference points. However, we use the minimum number of reference points (three reference points for two-dimensional space) for our evaluations.

In order to have a benchmark algorithm we also include a full multilateration scheme. This benchmark algorithm incorporates all reference points by minimizing the mean square distance error of the \( M \) reference points to the unknown target position. In the following, we will shortly present the applied multilateration scheme.

The position of reference node \( i \) is \( b_i \). The target node’s position \( p \) has to be estimated. This is done via the estimated distance \( d_i \) (based on measurements). Hence, in the ideal case we have

\[
\|b_i - p\|_2 = d_i^2 \tag{2}
\]

The left hand side can be rewritten as follows:

\[
\|b_i\|_2^2 + \|p\|_2^2 - 2b_i^T p = d_i^2
\]

Since there are \( M \) reference points, we again rewrite the above equation for every \( i = 1, \ldots, M \) to get the estimated position:

\[
p = 0.5B^+(a + \gamma) \tag{3}
\]

Here, we have \( \gamma := \|p\|_2^2 \) and \( [a_i] := \|b_i\|_2^2 - d_i^2 \) as well as the matrix of reference point positions \( B := [p_1, \ldots, p_M]^T \). The notation \( B^+ \) indicates the Moore-Penrose pseudoinverse. Prior to solving (3), one has to compute \( \gamma \) by solving the quadratic equation \( \gamma = p^T p = (0.5B^+(a + \gamma))^T (0.5B^+(a + \gamma)) \). In our case, we choose the \( \gamma \) which gives the least mean error in distance between each reference point and the target position (see (2)).

The described multilateration as well as all previously introduced algorithms were implemented and tested on the same hardware platform to enable a fair comparison of the obtained results.

III. Experimental Evaluation

In this section, we present the experimental testbed and evaluate the performance of the algorithms described in the previous section.

1A number of books multilateration is referred to as the hyperbolic localization using time difference of arrival (TDoA) measurements. In this work, we use the term as the natural extension of trilateration to more than \( N+1 \) reference points (e.g., more than three in 2-d).
A. Working Environment

We performed our experiment at the Ilmenau University of Technology, Germany. The basis for the evaluation of presented algorithms was measurement data collected with the UAV flying outside at the university campus.

Including the rotors, our 6-rotor multicopter has a diameter of about 1 meter and its flight time is about 20 minutes. The core of the UAV’s electronic platform is a gumstix computer module running a realtime Linux kernel. It is currently equipped with a Linksys WUSB54GC WLAN module for Wi-Fi localization as well as a control Wi-Fi interface for interfacing the flight control system using a notebook.

For GPS localization, we use a LEA-4H chip from Ublox using 2.5 x 2.5 cm patch antenna. The GPS-position itself is Kalman-filtered, using acceleration and attitude data from a CHR-6DM attitude and heading reference system.

After the data had been collected, it was proceeded through all algorithms described above. Therefore, each algorithm had the same input to enable a fair comparison.

B. Radio Propagation Model

Propagation models are used to approximate signal attenuation as a function of the distance between transmitters and receivers, being an important part of location estimation process. There exist plenty of models for indoor and for outdoor scenarios, but only few of them consider detecting the signal strength by an access point located inside a building relative to a device located outside.

For our disaster scenario, we have chosen the path loss model proposed in [15] that was designed to predict signal penetration into buildings from outside and consider the wall attenuation factor:

\[ P_t(d) = P_{t_0} - 10\alpha \log(d) - W + X\sigma, \quad (4) \]

where \( P_{t_0} \) is the signal strength 1 meter from the transmitter, \( \alpha \) is known as the path loss exponent, \( W \) is the wall attenuation factor, and \( X\sigma \) represents a Gaussian random variable with zero mean and standard deviation of \( \sigma \) dB [15]. The model was adopted according to the radio signal attenuation observed from real-world experiments. We applied following parameters according to the above equation:

\[ P_{t_0} = -40 \text{ dBm}, \quad \alpha = 3.32, \quad W = 4.8 \text{ dBm}, \quad X\sigma = 3.1 \text{ dBm}. \]

C. Results

To measure the performance of the proposed techniques, we use several metrics. We calculate the average localization error in meters to present the accuracy of selected methods. To present the robustness of algorithms with respect to the amount of collected data per node to be localized, we show the number of operations per location estimate defined as the total number of operations (multiplications) performed for a single location estimate. This is important in minimizing computation time as well as in minimizing the power consumption.

1) Accuracy: Figure 1 plots the cumulative distribution function (CDF) of the localization error for the OASIS methods (described in previous section) when varying the number of anchors from 3 to 272. In Fig. 2, the corresponding average localization error with respect to the whole measurement data is presented. According to these results, the key observations are the following:

- The smallest average localization error less than 2 m was obtained by Multilateration, Joint Clustering and Signal Strength-based methods with 1.6 m, 1.7 m and 1.8 m error respectively. Less good results are presented by AngleSS, Weighted Filter and Density methods with 2.1 m, 2.4 m and 2.7 m respectively. Area, Perimeter and Simple Convex Hull algorithms yield the biggest error with 3.9 m, 4.1 m and 8 m respectively.
- It is easy to notice, the algorithms that do not consider the strength of received signals (Area, Perimeter and Simple Convex Hull) show the worst results. These methods refer to the shapes formed by the anchors, which are situated furthest from the node to be found. These anchors present the most erroneous signal strength measurements.
- The multilateration scheme has been added as a benchmark for localization accuracy. We observed that multilateration has the least mean error and has a tendency for small localization errors. The applied multilateration is an unbiased estimator. Since it does not comprise
any weighting, its performance might be degraded in certain situations. In these cases, it can be outperformed by schemes like Joint Clustering, Signal Strength and AngleSS. However, it is still robust since it incorporates all measurement values.

- Additionally, for the proposed approaches, which all use signal-to-distance function, the quality of the RSS measurements is more important compared with the quality of selection and calculation. Data sequences that fit good to the used path loss model have mostly stronger signal strength and yield better localization performance. This fact explains good results obtained by Joint Clustering, AngleSS and Signal Strength-based methods.

- In Fig. 1, the localization error is presented from the other point of view. It is hard to ascertain, which algorithm shows the best overall results. For the location estimation precision of 2 m, the best probability is presented by AngleSS, Signal Strength and Joint Clustering with the 83rd, 83rd and 75th percentiles, respectively.

2) Number of Operations: As announced above, to compare the complexity of algorithms, we report the number of operations per single location estimate as well as the cumulative number of operations for all the described methods when increasing the number of anchors from 3 to 272.

For the most of the proposed OASIS methods, the complexity for calculation of one single estimate stays either almost constant over the whole operation time or increases very slowly with the growing amount of anchors being analyzed (Fig. 3). The Area and Perimeter algorithms show the number of operations increasing quadratic, because their main part includes brute-force approach for searching all possible combinations of anchors that form triangles. The worst results are presented by the Density algorithm: it uses brute-force technique as well, but each iteration here is much more complex.

The rest of the methods show linear increasing number of operations. Among of them, the multilateration is shifted up and presents more than 10 times slower computations because it involves solving an overdetermined system of equations.

In Fig. 4, the cumulative number of operations over the whole localization time against the amount of anchors is presented. These results show clearly that the algorithms with brute-force processing (i.e. Density, Area and Perimeter) are not useful for online localization process. Considering the fact that during the online location estimation there is new data coming each second, the localization algorithm cannot take longer than this time.

IV. Conclusion

In this paper, we presented different optimum anchor selection algorithms that improve the location estimation accuracy in comparison to some existing algorithms (e.g. Simple Convex Hull, Joint Clustering Technique). We showed that the choice of the certain OASIS method affects the localization accuracy significantly. The reported average localization error was from 1.6 m (Multilateration) to 8 m (Simple Convex Hull). According to the obtained error CDF, the best probability results to get the localization error under 2 m is presented by AngleSS, Signal Strength and Joint Clustering methods with the 83rd, 83rd and 75th percentiles, respectively.

Additionally, we compared the complexity of selected algorithms presenting the number of operations per single estimate. Although the Multilateration shows the best average localization error, it is much more complex than AngleSS, Signal Strength or Joint Clustering methods and presents more than 10 times slower computations due to the need of solving overdetermined system of equations.

In our scenario, we deal with the disaster situation. Therefore, the time constraints go through the whole research work as a very important requirement. Our scenario differs widely from those, which have been observed by investigation of anchor selection algorithms before this work. In such a way, we presented unique results in comparison of different anchor selection strategies to use under strict disaster conditions.

As a conclusion, we present in Fig. 5 the efficiency of OASIS methods as a CDF diagram. The efficiency factor represents the multiplication of the number of operations per location estimate and the localization error in meters. The smaller the number of needed operations and the localization error are, the more efficient a certain algorithm is. As expected, the best efficiency is provided by Signal Strength and
Joint Clustering methods, since they present good localization accuracy combined with the smallest computational effort. Furthermore, the combinations of algorithms can yield even better results. For instance, combining the Multilateration and Signal Strength-based methods, we obtained almost two mean localization error in comparison to the best results presented in this paper (see Fig. 2).

REFERENCES


