2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)

SomBe: Self-Organizing Map for Unstructured and Non-Coordinated iBeacon Constellations

Duc V. Le*, Wouter A.P. van Kleunen†, Thuong Nguyen‡, Nirvana Meratnia*, and Paul J. M. Havinga*
*Department of Computer Science, University of Twente, Enschede, The Netherlands
Email: {v.d.le, n.meratnia, p.j.m.havinga}@utwente.nl
†Locus Positioning, Enschede, The Netherlands
Email: wouter@locuspositioning.com
‡The Australian e-Health Research Centre, CSIRO, Herston, Queensland, Australia
Email: thuong.nguyen@csiro.au

***Abstract*—Bluetooth Low Energy (BLE) devices such as iBeacons have been popularly deployed for Location Based Services (LBS), including indoor infrastructure monitoring, positioning, and navigation. In these applications, the positions of iBeacons are assumed to be known. However, the location information is often unavailable or inaccurate as most iBeacons were deployed by different external parties. In addition, manual localizing the already-deployed iBeacons is costly and even impractical, especially in large-scale and complex indoor environments. Therefore, we propose a novel method, namely SomeBe, which can localize deployed iBeacons with a minimal effort and invasiveness to existing infrastructures. Specifically, our approach uses cooperative multilateration based on Received Signal Strength (RSS) of available smartphones and WiFi access points (APs) in the environment. Both Bluetooth signal strengths (between smartphones and iBeacons) and WiFi signal strengths (between smartphones and APs) are jointly employed in a single optimization cost function to surpass the local minima. Requiring that the positions of the APs are known only, the proposed cost function can also localize the iBeacons without knowing the positions of smartphones. To improve the localization accuracy, we employ a clustering method based on the RSS values for the coarse estimation of iBeacons' positions. SomBe also can be used to simplify iBeacon deployment as it can localize the iBeacons with a minimal effort. The performance evaluation results of our testbed experiments as well as realistic simulations show that SomBe outperforms non-cooperative approaches with 85% better in terms of accuracy.**

***Index Terms*—iBeacon localization; WiFi localization; RSS-based localization; location-based service; self-organizing iBeacons**

I. Introduction

In the last few years, iBeacons have become very popular thanks to their low cost and low power consumption. They were widely deployed for infrastructure monitoring and indoor Location-Based Service (LBS) in many public places such as airports, universities, hospitals, shopping malls, museums, stadiums, and warehouses. However, most iBeacons were deployed by different third parties or shop owners. Thus, the position of the iBeacons is typically unavailable or inaccurate. Since these places are large and often complex, it is laborious to recover the coordinates of all existing iBeacons manually. In some cases such as in the airports, localizing the deployed iBeacons is even impractical due to the structural complexity. Moreover, iBeacons may also be moved afterward due to building renovations, infrastructure replacement, and maintenance. Such location shifting makes it either infeasible or costly to keep position information of iBeacons up-to-date. Thus, there is a need for indoor positioning methods that allow the iBeacons to self-localize.

Existing approaches to iBeacon localization typically employ fixed Bluetooth sniffers with known positions to scan Received Signal Strengths (RSSs) of signals emitted from the iBeacons [1]. Since the Bluetooth range is relatively short (around $10-30$ m for indoor environments) and highly sensitive to obstacles, these localization approaches require an enormous number of Bluetooth sniffers to cover the entire place. Deploying a large number of Bluetooth sniffers is also costly and even impractical in certain situations. Some areas such as the airports even do not allow the deployment of Bluetooth sniffers due to privacy and security concerns. Since these areas are often large, it is also infeasible to scan these regions entirely by moving a Bluetooth sniffer around.

To address these challenges, we propose a novel cooperative smartphone-based approach, called Self-Organizing Map for Unstructured and Non-Coordinated iBeacons Constellations (SomBe), to quickly and accurately find the positions of the iBeacons and to help update their coordinates easily. SomBe aims at localizing deployed iBeacons using smartphones of the crowds with unknown locations and WiFi Access Points (APs) with known locations. Besides provided by the building owners, the APs locations may be freely obtained from the WiGLE (Wireless Network Mapping) [2], which contains the location information of the APs in many buildings.

A straightforward approach to this problem would first estimate the positions of the smartphones based on the known-position APs and then use the smartphones as the reference nodes to localize the iBeacons. While being able to estimate the positions of iBeacons, this approach faces a fundamental difficulty in obtaining the global optimum.

Therefore, we localize the iBeacons and smartphones through a cooperative optimization of a bridge multilateration cost function, called Bridge Sum of Squared Error (BISSE), which connects the known location of the WiFi APs to the unknown location of the Bluetooth iBeacons. To optimize the bridge cost function, we employ the Levenberg-Marquardt optimization algorithm as it was proven to be the best for non-linear least squares in [3], [4].

In a real-world deployment, smartphones running the SomBe application will opportunistically scan for surrounding APs and iBeacons whenever their WiFi and Bluetooth are enabled. Given scanned RSS values, the positions of iBeacons and smartphones are coarsely estimated based on a clustering method. Finally, a fine-grain step is employed to optimize the iBeacon positioning.

We evaluate our SomBe approach on testbed experiments, which resemble the indoor environments of the Schiphol airport. To obtain the testbed data, we deployed 36 iBeacons in a complex building that covers an area of 38×50 m. The building consists of multiple rooms and sections in different shapes and sizes. There are also numerous metal obstacles in the building, and the walls are made of various materials. This building can be considered as a typical public environment. We demonstrate our SomBe method through discovering the 36 iBeacons in this scenario. The experimental results show that SomBe can estimate quite accurately iBeacons even in such a complex structure, whereas the baseline approach mostly fails.

To the best of our knowledge, SomBe is the first solution for localizing non-coordinated iBeacons by using only smartphones in the crowds without requiring ground truth of smartphone locations, fingerprinting, or Bluetooth sniffers. SomBe is also the first to use a bridge cost function to localize uncoordinated iBeacons based on uncoordinated smartphones. Our specific contributions in this paper are:

(i) The SomBe method to construct a self-organizing map of unstructured and non-coordinate iBeacons using available APs and smartphones, without requiring the ground truth of smartphone positions, calibration, and environmental parameters.
(ii) Bridging WiFi and Bluetooth communications for indoor localization without deploying additional infrastructure such as WiFi and Bluetooth sniffers or using extra sensors such as inertial sensors.
(iii) The clustering-based technique to effectively estimate coarse positions such that minimizing the local minimum problem and improve localization accuracy.

The rest of this paper is organized as follows. Section II reviews the related techniques of WiFi- and iBeacon-based localization, especially for indoor environments. Section III provides the preliminary background on RSS least-squares modeling. Section IV presents SomBe, followed by performance evaluation and important observations presented in Section V. Section VI discusses advantages of SomBe as well as its potential to improve accuracy in other test beds. Finally, we conclude our paper in Section VII.

II. Related Work

Thanks to the low cost and low power consumption of iBeacons, using iBeacons as a means of localization has gained popularity. Many existing solutions including [5]–[7] use iBeacons with smartphones to provide location information. These studies assume that the iBeacons are stationary at a known location to function as anchors for localization. To the best of our knowledge, only a few studies addressed the inverse problem, which is positioning iBeacons using smartphones. For example, Koühne and Sieck [8] report on the use of smartphones to search for iBeacons. An proximity detection application was installed on the owner's smartphone to detect if the iBeacon is nearby.

Other studies such as [1] deployed Bluetooth sniffers at certain known locations to scan non-coordinated iBeacons. For a large area, this approach would be costly since it demands numerous Bluetooth sniffers. This approach also uses a fingerprint-based method, which relies on a map of error functions given availability of RSS measurements. Since fingerprinting depends on the environment, this approach is unable to cope with environmental changes.

Another approach to find deployed iBeacons is using the position of smartphones. In practice, the accurate location of the smartphones in indoor environments is hard to obtain. The most popular approach to precisely localize indoor smartphones based on WiFi signal strengths is fingerprinting [9]–[11]. However, building fingerprint databases is a laborious task as it requires the fingerprints of numerous calibrated positions. The built fingerprint database stays valid only for a short period as the environment may change due to objects and human mobility. Recent approaches [5], [6] propose to use dead-reckoning to reduce the burden of fingerprinting by using inertial sensors alongside WiFi RSSI. Although using inertial sensors could localize smartphones quite accurately, their power consumption is high, and it is vulnerable to privacy attacks by revealing user activity information.

Alternative approaches that are most related to our approach use the characteristic model of radio frequency propagation to avoid the laborious fingerprinting [12]–[17]. Example includes the Log-Normal Shadowing Model (LNSM) propagation model [18]. The LNSM defines the received signal strength as a function of the distance and two environmental parameters, i.e., the transmission power of the reference transmitter and the path-loss exponent. These parameters together with unknown coordinates can be estimated using a least-squares fitting technique [12]–[14]. However, these works do not address bridging communication links and typically yields inaccurately estimated locations in harsh environments due to the local minimum.

Using smartphones to sniff both WiFi APs and Bluetooth iBeacons is not new itself as being proposed in [19], [20]. However, these works aim at enhancing smartphones localization by using both WiFi and Bluetooth links. The iBeacons used in such works are deployed at fixed and known location. Conversely, in this paper the iBeacons are uncoordinated and we use smartphones as a bridge to localize the iBeacons. To the best of our knowledge, none of the previous works has addressed the use of smartphones as a bridge to localize uncoordinated iBeacons cooperatively.

III. Preliminaries

To eliminate the laborious-fingerprinting problem, we estimate the unknown position of a blind node based on a least-squares problem given pairwise distances between the blind node and its anchors. The problem can be solved with an optimization algorithm such as Levenberg-Marquardt [3], [4]. The most common technique to calculate pairwise distances is the LNSM, which allows us to compute the distance between the transmitter and the receiver (in our case the distance between smartphones and iBeacons or APs), based on RSS collected by the receiver (smartphones).

A. Log-Normal shadowing modeling

Given the RSS measurement in decibel milliwatts (dBm) at a distance of d from the transmitter to the receiver, the LNSM models the decay of the RSS over distance as follows.

$$\begin{aligned} P_d &\sim \mathcal{N}(\bar{P}_d, \sigma^2), \\ \bar{P}_d &= P_{d_0} - 10\beta \log_{10}(\frac{d}{d_0}), \end{aligned} \tag{1}$$

where P_d is the estimated power measured at a distance of d in dBm, P_{d_0} and β are the parameters representing the transmission power of the transmitter in dBm at a distance of d_0 and the path-loss exponent, respectively. The reference distance d_0 is typically set to 1 m for computation convenience. The LNSM also models the RSS measurements as a normal distribution with a mean of $\bar{P}_d$ and a standard deviation of σ.

To calculate the distance with RSS measurements, parameters P_{d_0} and β need to be known beforehand. Since (1) is simple and still valid in many indoor environments [18], it is commonly used in many works including [12]–[14], [21].

B. Self-calibrated Least Squares Estimator

Self-calibrated Least Squares Estimator (SLSE) can find the position of blind nodes such as smartphones and iBeacons without calibrating the environmental parameters or fingerprinting. The bottom line of Self-calibrating Least Squares Estimator (SLSE) is to determine the unknown location of the blind node as well as the unknown-environmental parameters by optimizing a cost function, which is the Sum of Squared Error (SSE) between real RSS measurement $\tilde{P}_d$ and their corresponding estimated model P_d. Mathematically, the cost function is typically described as follows.

$$F = SSE = \sum_{i=1}^{N} (\tilde{P}_{d_i} - \bar{P}_{d_i})^2, \tag{2}$$

where $d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2}$ is the Euclidean distance between the blind node and the anchor i in 2-dimensional space. The position of anchor node i is denoted by $(x_i, y_i)^T$ and the RSS measurement between the blind node and the anchor node i is denoted by $\tilde{P}_{d_i}$. P_{d_i} is modeled by Equation 1. Minimization of this sum of squares will return estimated position of the blind node as well as the environmental parameters. Since the model in (1) is simple to compute and calibrate, it is the most widely used model [12]–[14], [18], [21].

In other words, the SLSE problem is

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} (\tilde{P}_{d_i} - P_{d_0} + 10\beta \log_{10}(\frac{d_i}{d_0}))^2, \tag{3}$$

where $\theta = (x, y, P_{d_0}, \beta)^T$ is the set of unknown parameters.

In this paper, we employ SLSE on the standard approach, which first localizes the smartphones, then localizes iBeacons. Unless stated otherwise, throughout the rest of this paper, we consider approaches with SLSE as baselines, to compare with our cooperative approach, SomBe.

IV. SomBe: Self-Organizing Map for Unstructured and Non-Coordinated iBeacon Constellation

In this section, we first present an overview of our approach and real-world challenges. Then we present our SomBe approach that is based on a combined least-squares cost function for both WiFi and Bluetooth communications. Finally, we present two major phases of SomBe to estimate the location of non-coordinated iBeacons.

A. System Overview

As illustrated in Fig. 1, SomBe is designed for sensing systems that consist of iBeacons, WiFi APs, and smartphones. The iBeacons are attached to infrastructure for construction monitoring and function as reference nodes (anchors) for Bluetooth-based localization services. The iBeacons are nomadic and intermittently relocated after a short period. The APs are stationary for a long period. The smartphones carried by users are mobile. We assume that there are no Bluetooth base stations (sniffers) with known location information in the areas of interest. Smartphones opportunistically scan for radio signals emitted from surrounding iBeacons and APs. The collected RSS measurements are then sent to a central server. The most computationally intensive part of our approach is the optimization phase, which is done on the server side. The estimated coordinates of iBeacons, as well as the coordinates of the smartphones, will be sent back as a self-organizing map to the users helping them to find non-coordinated iBeacons, which attached to the infrastructure. Applications of using the self-organizing map provided by SomBe include indoor navigation and objects finding services for smartphone users in large and complex buildings.

B. Problem Statement

We consider a network that consists of M iBeacons with unknown positions and N stationary APs with known positions. We aim to localize M non-coordinated deployed iBeacons, given K RSS observations measured opportunistically by the smartphones in the region.

We assume that iBeacons were deployed randomly in the area, and there is no reference information (e.g., trajectories, movement patterns, or ground truth positions of the smartphones) to help predict the locations of the smartphones. Let

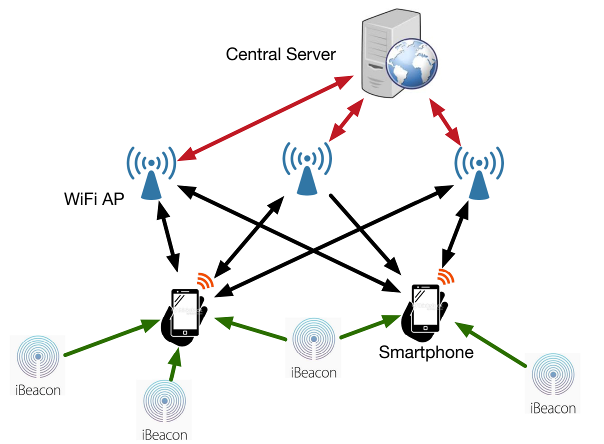


Fig. 1. An overview architecture of system applied the SomBe approach.

$B = \{(x_i, y_i)^T,\ i = 1, .., M\}$ denote the set of position vectors of non-coordinated iBeacons, i.e., the blind nodes. Let $A = \{(x_i, y_i)^T,\ i = M{+}1, .., M{+}N\}$ denote the set of position vectors of APs, i.e., the anchor nodes. Let $\tilde{P} = \{\tilde{P}_{i,j},\ i = M{+}N{+}1, .., M{+}N{+}K,\ j = 1, .., (M{+}N)\}$ denote the set of K observations of of RSS measurements collected by available smartphones in the area of interest, where $\tilde{P}_{ij}$ denotes the RSS of the measured power of observation i, transmitted from node j. Let $S = \{(x_i, y_i)^T,\ i = M + N + 1, .., M + N + K\}$ denote the set of position vectors where the corresponding RSS observations are measured. Since the propagation is symmetric, we assume that $\tilde{P}_{i,j} = \tilde{P}_{j,i}$. Note that the vector of the RSS measurements combine both Bluetooth and WiFi measurements. Note that the number of observations can be different from the number of the smartphones. The same smartphone may scan for the iBeacons and APs at multiple positions.

Applying the LNSM model described in (1), we model $P_{i,j}$ for shadowing or Line of Sight (LOS) as:

$$\begin{aligned} P_{i,j} &\sim \mathcal{N}(\bar{P}_{i,j}, \sigma^2_{i,j}), \\ \bar{P}_{i,j} &= P_{j,d_0} - 10\beta_j \log_{10}(\frac{d_{i,j}}{d_0}), \end{aligned} \tag{4}$$

where $d_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$.

Given the measurement set P and known positions of anchor set A, our objective is to estimate the unknown positions of blind nodes B without knowing the positions of observations.

Conventional approaches typically use SLSE to localize iBeacons in 2 steps [13]. First, the location of smartphones is estimated by the APs with positions known beforehand.

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=M+N+1}^{M+N+K} \sum_{j=M+1}^{M+N} (\tilde{P}_{i,j} - \bar{P}_{i,j})^2 \tag{5}$$

where $\hat{\theta}$ is the estimated unknown parameter matrix including the estimated locations of smartphones.

Second, the coordinates of iBeacons are estimated through the estimated locations of smartphones.

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=M+N+1}^{M+N+K} \sum_{j=1}^{M} (\tilde{P}_{i,j} - \bar{P}_{i,j})^2 \tag{6}$$

where $\hat{\theta}$ is the estimated unknown parameter matrix including the estimated locations of iBeacons. For the ease of discussion, in this paper we call the estimator in [13] SLSE. Although the conventional SLSE approach can localize the iBeacons in theory, the localization accuracy would be abysmal in practice since the errors are accumulated.

In our SomBe approach, we combine the measurements of WiFi and iBeacon channels into a single cooperative cost function:

$$\sum_{t=M+N+1}^{M+N+K} \left(\sum_{i=1}^{M} (\tilde{P}_{t,i} - \bar{P}_{t,i})^2 + \sum_{j=M+1}^{M+N} (\tilde{P}_{t,j} - \bar{P}_{t,j})^2 \right). \tag{7}$$

The unknown parameter matrix $\hat{\theta}$ includes the estimated positions of non-coordinated iBeacons, measuring positions, and environmental parameters.

$$\theta_{2M+N+K,2} = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_M & y_M \\ x_{M+N+1} & y_{M+N+1} \\ x_{M+N+2} & y_{M+N+2} \\ \vdots & \vdots \\ x_{M+N+K} & y_{M+N+K} \\ P_{1,0} & \beta_1 \\ P_{2,0} & \beta_2 \\ \vdots & \vdots \\ P_{M+N,0} & \beta_{M+N} \end{pmatrix}, \tag{8}$$

where $\{(x_i, y_i)^T,\ i = 1, .., M\}$ are the coordinate vectors of the iBeacons, $\{(x_i, y_i)^T,\ i = M + N + 1, .., M + N + K\}$ are the coordinate vectors of the observations, $\{P_{i,0},\ i = 1, .., M\}$ are the reference power of the iBeacons, $\{P_{i,0},\ i = M + 1, .., M + N\}$ are the reference power of the APs, $\{\beta_{i,0},\ i = 1, .., M\}$ are the path-loss exponents of the iBeacons, $\{\beta_{i,0},\ i = M + 1, .., M + N\}$ are the path-loss exponents of the APs. All these unknown parameters are estimated by applying the Levenberg-Marquardt optimization on (9), given K of the RSS observations.

C. Model optimization

To the best of our knowledge, the initial estimates of the environmental parameters have little impact on the estimation accuracy. Therefore, to obtain the initial values of environmental parameters, we just measure them at an AP and an iBeacon at different distances (e.g., at 1 m and 5 m). The most challenging part of the initial value estimation is the initial position estimates, which also have a significant impact on the estimation accuracy. Conventionally the initial estimates of θ are typically set random, zero, or central locations of the map when optimizing (5), (6), and (7). However, the accuracy of localization depends on the accuracy of the initial estimation, especially in a large complex space with None-line of Sight (NLOS). Therefore, we improve the accuracy of localization with a better initial estimation of the unknown parameters.

In a nutshell, SomBe performs the localization through two phases, which can be done either on the smartphones or a central server.

- Phase 1: Cluster-based initial values estimation.
- Phase 2: Bridge Sum of Squared Error optimization.

1) Cluster-based Initial Values Estimation: We first estimate the initial values of the observation positions, which are the positions of the smartphones where they measure the RSS of radio frequency from surrounding APs. This estimation is done by applying the Levenberg-Marquardt optimization on the error function in (5). To obtain the initial values of an observation position for the optimization in (5), we take the coordinate of the closest AP based on the strongest WiFi RSS measurement because the AP with the strongest RSS is the closest one to the observation position, except outliers due to fading channels. In case the smartphone cannot receive WiFi signal from any AP at a particular location, we use the central-map coordinates as the conventional approach. However, this problem rarely occurs in indoor environments because the APs are deliberately deployed to maximize the coverage.

The initial-value estimation of the iBeacon positions is more challenging than that of observations since the iBeacons can only communicate with the smartphones, of which locations are unknown. We overcome this issue by clustering the observations into clusters that have the centroids represented by the non-coordinated iBeacons as depicted in Fig. 2 – the number of groups equal to the number of non-coordinated iBeacons. Since the locations of observations are unknown, we cluster the observations directly based on their corresponding RSS values instead of their positions to avoid the cumulative errors. According to the LNSM propagation model, the higher RSS value, the closer distance between the observation and the iBeacon. The initial-location values of the iBeacons are computed as the mean locations of the observations in the corresponding clusters. In case a cluster does not have any member, we assign the central-map coordinates to the iBeacon.

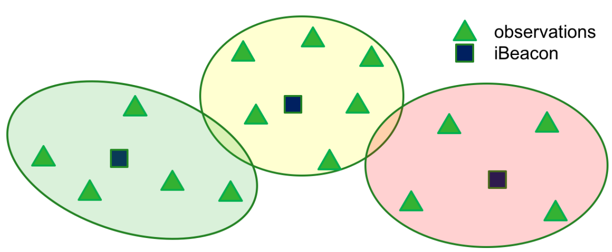


Fig. 2. Illustration of cluster-based initial values estimation.

For implementation, each observation has a RSS vector that consists of c RSS values from M surrounding iBeacons, $c \leqslant M$. We assign the observation to the group of iBeacon that has the strongest RSS. This process is repeated for all observations. By doing so, we will obtain M clusters of the RSS observations.

2) Bridge Sum of Squared Error (BISSE) Optimization: Given the initial parameter estimates obtained in the first phase and the RSS observations by the smartphones, SomBe uses the Levenberg-Marquardt optimization to estimate the optimal values of unknown parameters including the coordinates of iBeacons and observations.

Specifically, SomBe applies the Levenberg-Marquardt optimization in a cooperative manner. All unknown positions of iBeacons and observations are optimized simultaneously. Equation (7) can be represented as a combination of WiFi and iBeacon channels by:

$$F = \sum_{t=M+N+1}^{K} (f_t^{iBeacon} + f_t^{WiFi}). \tag{9}$$

Given K RSS observations, the error function $F = (f_1, f_2, .., f_K)^T$ is a vector of K error functions,

$$f_t = \sum_{i=1}^{M} (\tilde{P}_{t,i}^{iBeacon} - \bar{P}_{t,i}^{iBeacon})^2 + \sum_{j=1}^{N} (\tilde{P}_{t,j}^{WiFi} - \bar{P}_{t,j}^{WiFi})^2, \tag{10}$$

ALGORITHM 1: Iterative BISSE Optimization

INPUT:
$\{\tilde{P}_{i,j}\}$, initial coordinates $\theta^{(0)}$, reference coordinates $\boldsymbol{A}$
damping λ, λ_{up}, λ_{down}, accuracy ϵ, maximum iteration k_{max}
OUTPUT:
$\hat{\theta}$ minimizing $F = f(\theta)$ expressed by (9)
INITIALIZE:
$k := 0$; $\hat{\theta} = \theta^{(0)}$;
$f(\theta^{(k)}) := f(\hat{\theta})$;
while $(\frac{1}{2}||f(\theta^{(k)})||^2 > \epsilon) \& (k < k_{max})$ **do**
  $g(\theta) := J(\theta)^T f(\theta)$;
  $h := -(J(\theta)^T J(\theta) + \lambda I)^{-1} g(\theta)$;
  $\theta^{(k+1)} := \theta^{(k)} + h$;
  if $\frac{1}{2}||f(\theta^{(k+1)})||^2 < \frac{1}{2}||f(\theta^{(k)})||^2$ **then**
    $k := k + 1$;
    $\lambda := \lambda / \lambda_{down}$;
  end
  else
    $\lambda := \lambda \times \lambda_{up}$;
  end
end
$\hat{\theta} := \theta^{(k)}$

The pseudocode of the BISSE optimization based on the Levenberg-Marquardt algorithm for SomBe is summarized

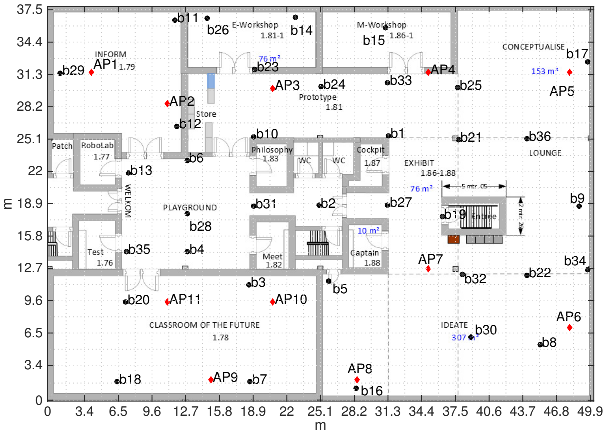


Fig. 3. Deployment area and placement of 36 iBeacons and 11 APs. Most iBeacons and APs have to be placed between the roof and celling due to constructive constraints.

in Algorithm 1. The BISSE optimization starts with initial guess $\theta^{(0)}$ which is estimated in the first phase. The estimated coordinates $\hat{\theta}$ is adjusted by the step h only for downhill steps. The damping factor λ is dynamically adapted according to the reduction of F. This adaption will bring the optimization closer to either the Gauss-Newton algorithm or the gradient descent direction to avoid the local minimum. The iterative loop stops when the residual $\frac{1}{2}||f(\theta^{(k)})||^2$ is smaller than a predefined ϵ or it reaches the maximum number of iterations k_{max}. More details of the Levenberg-Marquardt algorithm are presented in the Appendix.

V. Test Bed Experiments

In this section, we first describe our experimental setup to evaluate SomBe in a real indoor environment. Then we will present performance evaluation and discussion of results.

A. Experimental setup

Fig. 3 illustrates the placement of iBeacons and APs in our testbed environment, which is a relatively large and complex laboratory. The area size is approximately 38×50 m and has various rooms and sections that are separated by walls of different materials such as glasses, concrete, woods, plastics, steel. Due to the complex architecture and the restriction of the area, the iBeacons were deployed at specific places on the ceiling, along with the beams at about $4 - 6$ m high (see Fig. 4). The transmitting power of iBeacons was set to -59 dBm to reduce the power consumption. There exist 11 Cisco APs in the middle of the area for the best coverage. Such AP deployment is problematic for WiFi-based localization accuracy. For accurate localization, The APs should also be deployed near the edges of the area. The actual location of the iBeacons and the APs were recorded during the deployment with an error of about ± 0.25 m due to the complexity of the building.

For WiFi and iBeacon scanning, we develop a smartphone application that can scan and record RSS emitted from both iBeacons and APs simultaneously. We set the time intervals between scans to 1 second. This short sampling makes it possible for the smartphones to scan for RSS opportunistically for both cases, when the phone users are not moving or even when they are strolling (with a speed of less than 1 m/s). This 1-second assumption is based on the empirical device-discovery latency with our old-version smartphones, which is also consistent with the experimental results in [22]. To be able to scan while moving quickly, the smartphones must have a smaller discovery latency. Nevertheless, this is an engineering issue, which is beyond the scope of this paper.

To investigate the performance of SomBe, we compare it with four alternative approaches.

1) *SLSE*: The self-calibrated least squares approach [13] using the Levenberg-Marquardt method with initial estimates as the central positions of the area. This approach is considered as the baseline approach.
2) *SLSE-Cluster*: The self-calibrated least squares approach [13] using the Levenberg-Marquardt method with initial positions estimated by our cluster-based estimation. This approach is for investigating the effect of our proposed cluster-based estimation.
3) *SOMBE-Center*: SomBe using the Levenberg-Marquardt method with initial estimates as the central positions of the area. This approach is for investigating the effect of our proposed Bridge Sum of Squares Error (BISSE) function.
4) *Error Bound*: SomBe using the Levenberg-Marquardt method with actual location of observations (smartphones). This approach is considered as the lower bound of SomBe, when the position of observations is assumed to be known exactly.

We implement SomBe and compared algorithms in Matlab using the *fsolve* function with the Levenberg-Marquardt optimization. The Levenberg-Marquardt optimization is also implemented in [13] and widely regarded as the most robust modified Gauss-Newton algorithm. The maximum number of iterations k_{max} is set to 10^6. Other parameters of the Levenberg-Marquardt algorithm are kept by default, for example, the initial damping factor $\lambda = 0.01$.

B. Experimental results

Fig. 5(a), (b), and (c) illustrate the errors of estimated iBeacon coordinates of the Error Bound, SomBe, and SLSE approaches are used, respectively. The results show that SomBe provide much better estimated positions of iBeacons (3.78 m) and very close to the Error Bound (3.38 m), while SLSE in [13] performs unacceptable in terms of accuracy (21.84 m). Such results are consistent with our hypothesis – the BISSE cost function and cluster-based initialization would reduce the local minimum problem. The fact that most iBeacons are placed along the beams and the walls results in shadowing and makes the localization challenging. We also observe that all approaches have poor performance for the iBeacons that are close to the borders of the map. Both SomBe and the most existing localization algorithms have trouble localizing nodes near the edges of the areas of interest. This artifact is because

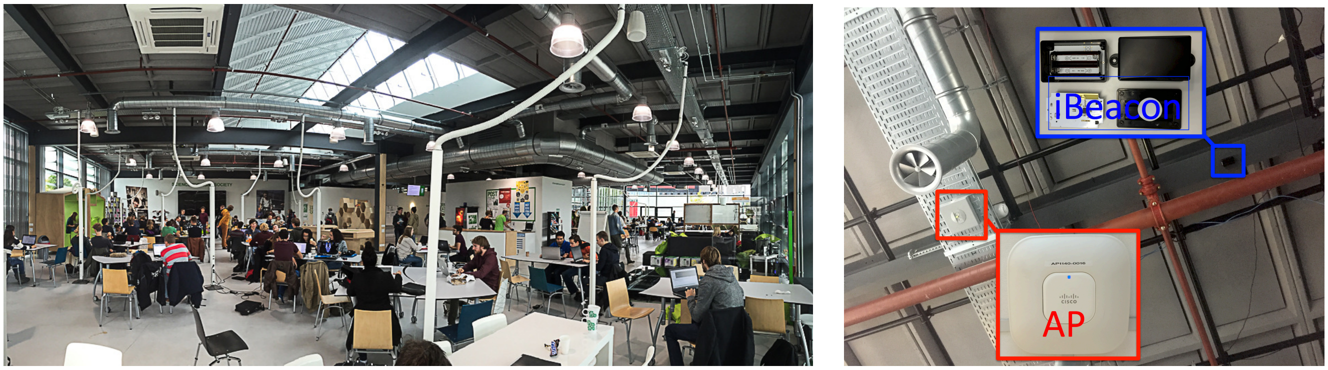
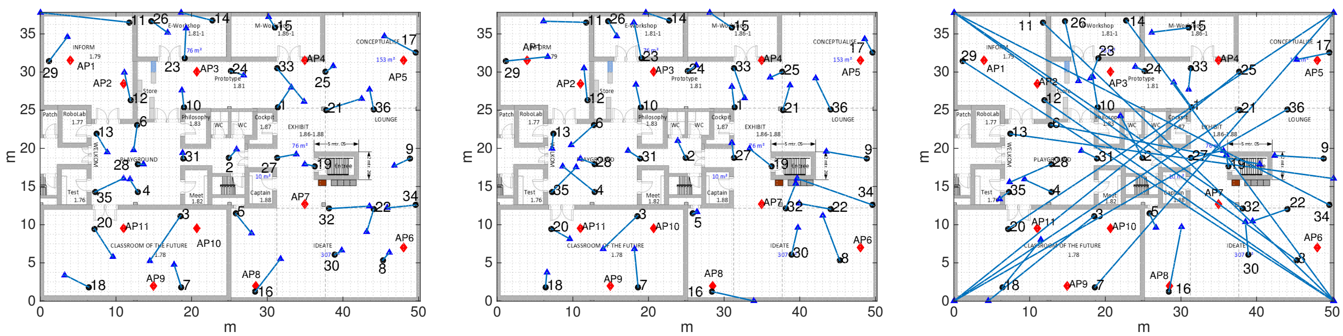


Fig. 4. The $IDEATE$ sub area of our laboratory (see Fig. 3). The laboratory has various rooms and sections that are separated by walls made of diverse materials such as glasses, concrete, woods, plastics, steel. Most iBeacons and APs can only placed near the walls and beams, right under the roofs of which heights are about 4 m.



(a) Error Bound: MAE = 3.38 m (b) SomBe: MAE = 3.78 m (c) SLSE: MAE = 21.84 m

Fig. 5. Estimated iBeacon positions marked as filled triangulars: (a) The lower error bound of localization when knowing exact location of smartphones; (b) Localization results of our SomBe approach; (b) Localization results of the base line approach SLSE [13] with the central map for all initial values.

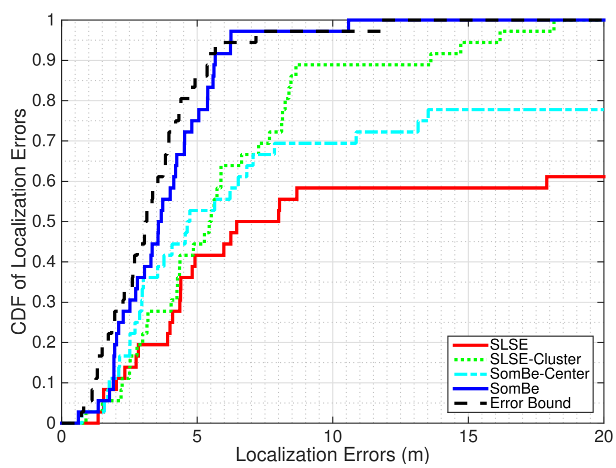


Fig. 6. Cumulative Distribution Function (CDF) of iBeacons localization errors when applying different approaches.

the border nodes have much fewer observations and they are all taken from the same side of the node. In short, since there is less information about the nodes near the borders, these border nodes are often un-localizable. Even when their locations can be estimated, the errors tend to be large. Therefore, an indoor localization method should be well aware of the problem of unreliability of the positions near to the edge of the map.

To investigate the reliability of the localization approaches, we compute the Cumulative Distribution Function (CDF) of the estimation errors. Fig. 6 plots the CDF of location errors of iBeacons when applying different localization approaches. The graphs show that the probability of accurately estimating the position of iBeacons by SomBe is much higher than the probability of the SLSE approach. For example, approximately 90% of the non-coordinated iBeacons were localized with an error lesser than 5.1 m by SomBe, whereas that probability for the SLSE approach is only 40%. The large localization error is mainly due to the iBeacons near to the edges of the map, which is an NP-hard problem for any multilateration algorithm. This significant improvement is a result of our both main contributions, the BISSE function and cluster-based initialization. Fig. 6 also shows that only SomBe could closely follow the Error Bound, of which all observations have known location information and the actual coordinates of iBeacons are used as initial positions for the Levenberg-Marquardt optimization.

Since the number of observations might have an impact on the localization accuracy, we vary the number of observations by randomly sampling the observation set, without replacement. Fig. 7 shows the bar plots with standard deviations of the average errors when varying the number of observations. We observe that the localization accuracy of the SomBe increases

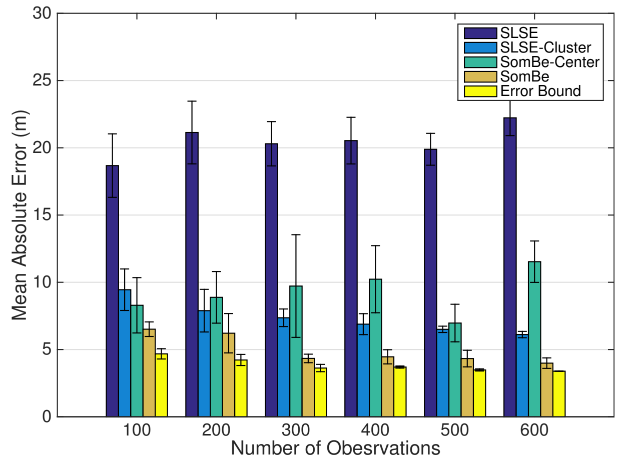


Fig. 7. iBeacons localization errors of different approaches when varying the number of observations.

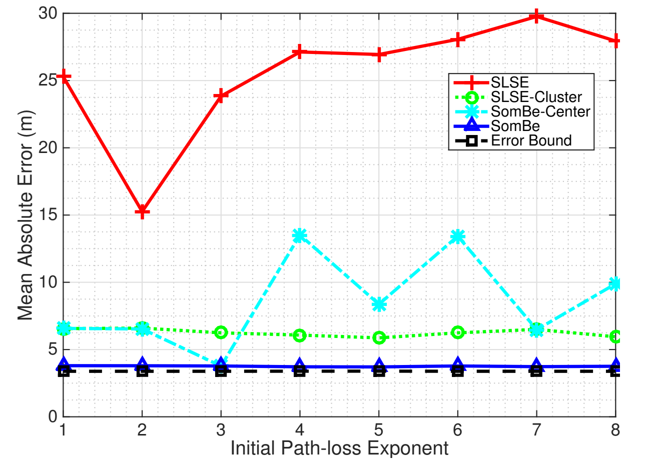


Fig. 8. iBeacons localization errors when varying only the initial value of the path-loss exponent β.

with the number of observations. Whereas, increasing the number of observations does not improve the localization accuracy of the baseline approach, SLSE. Besides, the small standard deviations of localization errors of SomBe show that the estimated positions are reliable. Fig. 7 also shows that the performance of SomBe is getting closer to that of the Error Bound when the number of observations increases. We also observe that the BISSE cost function (SomBe-Center) and cluster-based initialization (SLSE-Cluster) improve the performance of SomBe about 72% and 45% over SLSE, respectively, when applying them separately. We also noticed that it requires at least 300 observations such that all 36-iBeacon clusters have at least one observation for each iBeacon to compute the initial location. If not, the initial location of the iBeacon is set to the center of the map, which will result in a lower localization accuracy. In practice, such required 300 observations can be accumulated over time and reused for the next localization update. The system will start providing a good localization service as long as it gathers enough observations.

Since SomBe initializes the default values of the environmental parameters by simply measuring RSS at 1 m and 5 m within a duration of 5 minutes and in a standard environment, it is interesting to investigate the effects of this setting on the estimation accuracy. The reference power P_{d_0} is valid for most environments since there is rarely any obstacle within a radius of 1 m of the transmitters. However, the path-loss exponent β would vary a lot depending on each used channel as the distance from the smartphones to the iBeacons and the APs placed on the ceiling are quite far. Thus we set initial values of P_{d_0} to the measured values and set -33 dB for APs and -77 dB for iBeacons, and vary the initial values of β. The results in Fig. 8 show that the error of estimated coordinates obtained by SomBe are reliable. We also check the average of estimated β by SomBe. It is approximately 2.3 for all different initial setting of β, and close to our empirical measurement, 2.5. Conversely, the baseline SLSE approach is unstable when varying β. By looking at the performance of SomeBe-Center and SLSE, we conclude that the instability is mainly due to the poor initialization for the optimization.

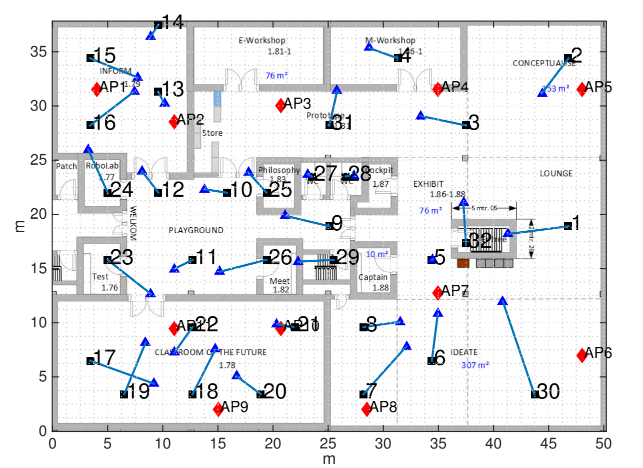


Fig. 9. Observation/smartphone localization errors at 30 random locations, of which have MAE = 3.5 m.

Last but not least, SomBe also returns the estimated locations of observations after optimizing (9). Fig. 9 shows a Mean Absolute Error (MAE) of 3.5 m between the actual locations and estimated locations of a set of 30 observations that were taken at random locations. These actual locations are recorded just for evaluation purposes. SomBe does not require any ground truth of observation locations. Overall, the estimated locations of all 642 observations have a MAE of 4.9 m. The estimated locations also have poor results when the locations are near to the edges of the map. When we optimizing only WiFi information in (9), as done by [23], these 30 estimated observations have a MAE of 9.8 m. Therefore, our SomBe localization do improve the localization accuracy of 64% mainly due to the collaborative AP-iBeacon bridging.

VI. Discussion

Our SomBe approach advances other approaches such as fingerprinting and dead-reckoning in several ways. First, our SomBe approach is self-organized and more suitable to update the map of non-coordinated iBeacons frequently since it does not require position information of the observations, where the smartphones take the RSS measurements. Conversely, fingerprinting approaches indeed require numerous labeled positions of observations to build fingerprint databases, which is laborious and even impractical in some large and complex spaces. Second, the hybrid approaches such as [5], [6] require both inertial data and APs' positions to estimate the current position of smartphones. In such way, it consumes more battery power and the activities of the users might be leaked through activity recognition using accelerometer data. Conversely, our SomBe approach uses available WiFi and Bluetooth scanning, which are commonly already turned on in indoor environments. Third, our SomBe approach does not require any extra infrastructure to scan for the smartphones or iBeacons. Our approach only uses the APs available in the interesting areas, of which locations are online available for numerous buildings [2]. It also does not require that the APs are at distinct landmarks like dead-reckoning to cancel the drifts of smartphone's trajectories. Thus our approach is more scalable and practical. Finally, our SomBe approach demands less effort to implement and to conduct the experiments since it does not require the additional infrastructure and fingerprinting; hence, it is easier for researchers in the community to apply to their work.

Besides the above advantages, our approach still has some limitations. First, the estimated positions in our testbed may be not accurate enough for demanding applications. The reason is that in this work we target harsh environments such as in a large airport, where the iBeacons were strictly deployed at specific locations such as the beams under the roofs. However, the localization accuracy could be improved when having a lot more measurements. The localization accuracy also can be enhanced with more reliable measurements by using modern smartphones with better Bluetooth chips and operating systems. Second, the assumption that smartphone is stationary for at least 1 second while taking the observations is inapplicable for the realistic scenario when the smartphones are in used. This 1-second interval was set based on the scanning capability of the smartphones in our experiments. More modern smartphones in future would reduce the discover latency to better cope with a real movement in a crowd.

Nevertheless, the testbed results still show that our SomBe approach is applicable in a wide range of applications. Examples include applications that provide a digital context beacon to a physical space (LBS), such as retail stores in a duty-free area of an airport. Not only can our SomBe approach localize the deployed iBeacons, but also it can help deploying the non-coordinated iBeacons quickly in a large scale. Users with a Bluetooth-enabled smartphone can find a retail store or receive promotion information of ones nearby. Another potential application is facility monitoring and tracking. Together with the available APs, it is feasible to monitor quite accurately the mobile things such as trolleys and freights that are attached iBeacons in large and complex logistics stores. Employees just need to carry smartphones running SomBe while walking in the area. Since the RSS measurements can be measured within a second, it is feasible to track both the location of employees and freights periodically, such as every 5 minutes.

Finally, our SomBe approach complexity is similar to the SLSE approach because the cluster-based initial guess is lightweight. It requires only K loops to cluster the RSS observations to the regarding iBeacon's cluster. The optimization take around one hundred milliseconds at the server in our experiment, which is typical personal computer.

VII. Conclusion

In this paper, we presented the SomBe approach, a self-organizing map of unstructured and non-coordinated iBeacons constellation, using the available Wireless Access Points (APs). Without requiring any Bluetooth sniffer with known location as in other existing works, SomBe aims at bridging the gap in RSS measurements between the APs and the iBeacons using the available smartphones, connecting the known locations of WiFi APs to the unknown locations of Bluetooth iBeacons. Given the RSS measurements, SomBe simultaneously localizes the iBeacons and the smartphones by the optimization of a bridge multilateration cost function, using the Levenberg-Marquardt optimization algorithm. The optimization also estimates the unknown environmental parameters. By doing so, SomBe can estimate the non-coordinated iBeacons without requiring ground-truth recording, fingerprinting, environmental parameters, or stationary iBeacon sniffers. To improve the accuracy of location estimation, we employed a cluster-based method to estimate the coarse position of the iBeacons. We validated SomBe with real-world experiments in a large complex building as well as with simulations. The results show that SomBe can localize the deployed iBeacons considerably accurately, outperforming the baseline approaches and close to the lower error bound. SomBe can be used as a minimally intrusive method to localize the already-deployed iBeacons or as an easy way to deploy new iBeacons on a large scale.

Acknowledgment

This work is partly supported by the COUNTDOWN project (Grant No. 509-21342) in the EFRO OP-OOST programme.

References

[1] İ. Okşar, "A Bluetooth signal strength based indoor localization method," in *IWSSIP 2014 Proceedings*. IEEE, 2014, pp. 251–254.
[2] "Wigle.net." [Online]. Available: https://wigle.net/
[3] D. V. Le, J. W. Kamminga, H. Scholten, and P. J. Havinga, "Nondeterministic sound source localization with smartphones in crowdsensing," in *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. IEEE, 2016, pp. 1–7.
[4] J. J. Moré, "The levenberg-marquardt algorithm: implementation and theory," in *Numerical analysis*. Springer, 1978, pp. 105–116.
[5] Y. Guo, Y. Sun, Y. Li, T.-Y. Wu, and M. S. Obaidat, "Accurate indoor localization with crowd sensing," in *Communications (ICC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–7.

[6] Z. Chen, Q. Zhu, H. Jiang, and Y. C. Soh, "Indoor localization using smartphone sensors and iBeacons," in *Industrial Electronics and Applications (ICIEA), 2015 IEEE 10th Conference on*. IEEE, 2015, pp. 1723–1728.

[7] G. Conte, M. De Marchi, A. A. Nacci, V. Rana, and D. Sciuto, "Bluesentinel: a first approach using ibeacon for an energy efficient occupancy detection system." in *BuildSys@ SenSys*, 2014, pp. 11–19.

[8] M. Koühne and J. Sieck, "Location-based services with ibeacon technology," in *Artificial Intelligence, Modelling and Simulation (AIMS), 2014 2nd International Conference on*. IEEE, 2014, pp. 315–321.

[9] H. Jing, J. Pinchin, C. Hill, and T. Moore, "Wi-Fi fingerprinting based on collaborative confidence level training," *Pervasive and Mobile Computing*, 2015.

[10] W. Meng, W. Xiao, W. Ni, and L. Xie, "Secure and robust Wi-Fi fingerprinting indoor localization." in *IPIN*, 2011, pp. 1–7.

[11] K. Kaemarungsi and P. Krishnamurthy, "Modeling of indoor positioning systems based on location fingerprinting," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 2. IEEE, 2004, pp. 1012–1022.

[12] N. Patwari, R. J. O'Dea, and Y. Wang, "Relative location in wireless networks," in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol. 2. IEEE, 2001, pp. 1149–1153.

[13] X. Li, "RSS-based location estimation with unknown pathloss model," *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3626–3633, 2006.

[14] F. Gustafsson and F. Gunnarsson, "Localization based on observations linear in log range," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 10 252–10 257, 2008.

[15] V.-D. Le, V.-H. Dang, S. Lee, and S.-H. Lee, "Distributed localization in wireless sensor networks based on force-vectors," in *Intelligent Sensors, Sensor Networks and Information Processing, 2008. ISSNIP 2008. International Conference on*. IEEE, 2008, pp. 31–36.

[16] D. J. Bijwaard, W. A. van Kleunen, P. J. Havinga, L. Kleiboer, and M. J. Bijl, "Industry: Using dynamic WSNs in smart logistics for fruits and pharmacy," in *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*. ACM, 2011, pp. 218–231.

[17] V.-H. Dang, V.-D. Le, Y.-K. Lee, and S. Lee, "Distributed push–pull estimation for node localization in wireless sensor networks," *Journal of Parallel and Distributed Computing*, vol. 71, no. 3, pp. 471–484, 2011.

[18] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal processing magazine*, vol. 22, no. 4, pp. 54–69, 2005.

[19] K. Antevski, A. E. Redondi, and R. Pitic, "A hybrid ble and wi-fi localization system for the creation of study groups in smart libraries," in *Wireless and Mobile Networking Conference (WMNC), 2016 9th IFIP*. IEEE, 2016, pp. 41–48.

[20] A. Baniukevic, C. S. Jensen, and H. Lu, "Hybrid indoor positioning with Wi-Fi and Bluetooth: Architecture and performance," in *Mobile Data Management (MDM), 2013 IEEE 14th International Conference on*, vol. 1. IEEE, 2013, pp. 207–216.

[21] S. Mazuelas, A. Bahillo, R. M. Lorenzo, P. Fernandez, F. A. Lago, E. Garcia, J. Blas, and E. J. Abril, "Robust indoor positioning provided by real-time RSSI values in unmodified WLAN networks," *IEEE Journal of selected topics in signal processing*, vol. 3, no. 5, pp. 821–831, 2009.

[22] K. Cho, W. Park, M. Hong, G. Park, W. Cho, J. Seo, and K. Han, "Analysis of latency performance of Bluetooth low energy (BLE) networks," *Sensors*, vol. 15, no. 1, pp. 59–78, 2014.

[23] M. Robinson and I. Psaromiligkos, "Received signal strength based location estimation of a wireless lan client," in *IEEE Wireless Communications and Networking Conference, 2005*, vol. 4. IEEE, 2005, pp. 2350–2354.

APPENDIX

The Levenberg-Marquardt algorithm consists of an iterative least-square minimization of a cost function based on a modification of the Gauss-Newton method. The aim is to find the optimal $\hat{\theta}$ such that the scalar error is minimal.

$$\hat{\theta} = \arg\min_{\theta} F(\theta) = \arg\min_{\theta} \frac{1}{2}||f(\theta)||^2. \tag{11}$$

To estimate θ, we can use some increments for each individual parameter $\Delta\theta_j$ to estimate the Jacobian $J(\theta)$:

$$J_{ij}(\theta) \simeq \frac{f_i(\theta + \Delta\theta) - f_i(\theta - \Delta\theta)}{2\Delta\theta_j} \tag{12}$$

The Levenberg-Marquardt algorithm consists of an iterative least-square minimization of a cost function based on a modification of the Gauss-Newton method. The aim is to find the optimal $\hat{\theta}$ such that the scalar error is minimal.

$$\hat{\theta} = \arg\min_{\theta} F(\theta) = \arg\min_{\theta} \frac{1}{2}||f(\theta)||^2. \tag{13}$$

To estimate θ, we can use some increments for each individual parameter $\Delta\theta_j$ to estimate the Jacobian $J(\theta)$:

$$J_{ij}(\theta) \simeq \frac{f_i(\theta + \Delta\theta) - f_i(\theta - \Delta\theta)}{2\Delta\theta_j} \tag{14}$$

Therefore, the actual error function is approximated by its linearization as follows:

$$F(\theta + h) \simeq F(\theta) + hg(\theta) + \frac{1}{2}h^T H(\theta)h, \tag{15}$$

where $g(\theta) = J(\theta)^T f(\theta)$ is the gradient, $H(\theta) = J(\theta)^T J(\theta)$ is the Hessian of the error function, and h is the iteration step. The step h is defined by the following equation:

$$(H(\theta) + \lambda I)h = -g(\theta), \tag{16}$$

where λ is the damping factor which is adapted dynamically according to a heuristic rule. If the reduction of F is rapid, a smaller value of λ can be used. It will bring the algorithm closer to the Gauss-Newton algorithm. If an iteration gives insufficient reduction in the residual, a large value of λ should be used. It will give a step closer to the gradient descent direction.

The first initial guess $\hat{\theta}^{(0)}$ must be provided for the algorithm to start optimizing the error function. Then each iteration the Levenberg-Marquardt optimization algorithm performs:

$$\hat{\theta}^{(t+1)} = \hat{\theta}^{(t)} + h. \tag{17}$$

If λ is very small, it means that it is a good step in the final stage of the minimization when $\hat{\theta}$ is close to θ. The condition to stop the minimization also can be based on $F(\hat{\theta}) = \epsilon$ condition, which $\epsilon = 0$ or is very small. Then we get $\hat{\theta}$ which is close to θ – unknown positions are estimated together with environmental parameters.