

Localization of Sensors in Indoor Wireless Networks: An Observation Model Using WiFi RSS

Daniel Alshamaa, Farah Mourad-Chehade, Paul Honeine

▶ **To cite this version:**

Daniel Alshamaa, Farah Mourad-Chehade, Paul Honeine. Localization of Sensors in Indoor Wireless Networks: An Observation Model Using WiFi RSS. Proc. 9th IFIP International Conference on New Technologies, Mobility and Security - Workshop on Wireless Sensor Networks: Architectures, Deployments, and Trends, 2018, Paris, France. 10.1109/NTMS.2018.8328699 . hal-01965906

HAL Id: hal-01965906
https://hal.science/hal-01965906

Submitted on 27 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Localization of Sensors in Indoor Wireless Networks: An Observation Model Using WiFi RSS

Daniel AlShamaa, Farah Mourad-Chehade
Institut Charles Delaunay, ROSAS, LM2S
Université de Technologie de Troyes
UMR 6281, CNRS, Troyes, France
daniel.alshamaa@utt.fr, farah.chehade@utt.fr

Paul Honeine
LITIS lab
Université de Rouen
Rouen, France
paul.honeine@univ-rouen.fr

***Abstract*—Indoor localization has become an important issue for wireless sensor networks. This paper presents a zoning-based localization technique that works efficiently in indoor environments. The targeted area is composed of several zones, the objective being to determine the zone of the sensor using an observation model. The observation model is constructed based on fingerprints collected as WiFi signals strengths received from surrounding Access Points. The method creates a belief functions framework that uses all available information to assign evidence to each zone. A hierarchical clustering technique is then applied to create a two-level hierarchy composed of clusters and of original zones in each cluster. At each level of the hierarchy, an Access Point selection approach is proposed to choose the best subset of Access Points in terms of discriminative capacity and redundancy. Real experiments demonstrate the effectiveness of this approach and its competence compared to state-of-the-art methods.**

***Index Terms*—Access point selection, belief functions, hierarchical clustering, localization, observation model, WiFi signals.**

I. Introduction

For several years, wireless sensor networks have aroused great interest in research communities. Such networks are being widely used in various fields to perform several tasks ranging from medical and military applications, to monitoring homes, hospitals and forests [1]. Localization is a key aspect of such networks, since the knowledge of the sensor's location is critical to process the information it acquired, to actuate responses to the environment, or to set an alarm in an emerging situation [2].

The most widely adopted approach in indoor localization is wireless fingerprinting [3]. Wireless fingerprinting leverages the available wireless transceivers along with already deployed networking infrastructure. The advantage of such solution is reducing time and cost of installing any additional hardware. This approach requires an offline training phase and an online localization phase. Many existing works have tackled the problem as exact positioning. The disadvantage of such approach is their need to a database with exact locations, whose construction is time-costly and complex. In addition, there exists several applications where the zone of the mobile sensor is of interest and not its exact position. This issue is important for the health-care domain for instance, where Alzheimer's patients might be lost in their nursing home [4], in museums for supporting guides and emergency management [5], for large malls to facilitate shopping [6], etc and where locating people in a specific zone of such environments is completely sufficient. Some researchers have studied the zoning problem without computing exact locations [7], [8], [9]. In [7], the signal strengths are collected while the user is moving, and the zone is determined when a sharp change in the strength indicator is detected. In [8], the design of overlapping zone partitioning is implemented. In [9], the authors tackle the zoning problem as a multi-class classification problem and use Support Vector Machines (SVM) to solve it. Even though a lot of approaches have been proposed in the literature using WiFi, the indoor localization problem still remains unsolved. The research community has not converged to a single, widely accepted solution that can achieve the desired accuracy. In fact, it was found that that WiFi signal strengths are unstable and they vary widely even at the same position with time, temperature, moving objects, and other parameters.

In our previous work [10], we proposed a basic observation model that works well in case of small number of zones. When the number of zones increases, the basic model fails to achieve a high accuracy due to difficulty in discriminating between overlapping distributions representing the different zones. This paper presents an extended model as described in the following. At first, the method consists in constructing a fingerprinting database that associates to each zone a set of WiFi signals strengths (RSS) collected from Access Points (APs). This database is used with the belief functions theory to create supersets of zones and affiliate evidence to each one according to APs. A clustering technique is then applied, leading to a two-level hierarchy composed of clusters, and the original zones in each cluster. At each level, an Access Point selection technique is proposed to choose the subset of APs with maximum discrimination and minimum redundancy. All evidence are combined to determine a level of confidence of having the mobile sensor resides in each zone. By taking into account information uncertainty, the proposed method yields levels of confidence for zones of covering the new observation. Experiments are conducted in a Living Lab, and the results show the effectiveness of the proposed method.

The remainder of the paper is organized as follows. Section II presents the localization approach. Section III demonstrates the experimental results and the comparison to state-of-the-art algorithms. Section IV concludes the paper.

II. Localization approach

A. Formulation of the problem

The localization problem is tackled in the following manner. Let N_Z be the number of zones of the targeted area, denoted by $Z_k,\ k = 1, 2, \ldots, N_Z$ and N_{AP} be the number of detected APs, denoted by $AP_n,\ n = 1, 2 \ldots, N_{AP}$. The aim of the presented algorithm is to propose an observation model, using the WiFi RSS, to assign a confidence level $\mathcal{C}f(Z_k)$ for each zone Z_k, for any new observation $\boldsymbol{\rho}$. Here, $\boldsymbol{\rho}$ is a vector of size N_{AP} of RSS measurements collected by the mobile sensor from surrounding APs. The observation model uses RSS data received from surrounding WiFi APs to estimate the zone of the mobile sensor. In an offline phase, fingerprints are collected by measuring the RSS of all existing APs in random positions of each zone. Then, in the online phase, once a new measurement of RSS is received, the model is used to assign a certain confidence to each zone.

B. Mass association

The observation model is constituted of fitting the RSS observations into statistical distributions, and using the belief functions theory as a framework for mass association. To take into account information uncertainty, supersets of zones are considered and not only the singletons, which permits associating masses as per available evidence.

Let $\mathcal{Z} = \{Z_1, \ldots, Z_{N_Z}\}$ be the set of all possible zones and let $2^{\mathcal{Z}}$ be the set of all the supersets of $\mathcal{Z}$, i.e., $2^{\mathcal{Z}} = \{\emptyset, \{Z_1\}, \ldots, \mathcal{Z}\}$. The cardinal of $2^{\mathcal{Z}}$ is equal to $2^{|\mathcal{Z}|} = 2^{N_Z}$, where $|\mathcal{Z}|$ denotes the cardinal of $\mathcal{Z}$. One fundamental function of the BFT is the mass function, also called the basic belief assignment (BBA). A mass function $m(\cdot)$ is a mapping from $2^{\mathcal{Z}}$ to the interval $[0, 1]$, defined according to a set of sources $\{AP_1, \ldots, AP_{N_{AP}}\}$. It satisfies:

$$\sum_{A \in 2^{\mathcal{Z}}} m_{AP_n,t}(A) = 1. \tag{1}$$

The mass $m(A)$ given to $A \in 2^{\mathcal{Z}}$ stands for the proportion of evidence, brought by the set of sources, saying that the observed variable belongs to A.

In order to define the sources' BBAs, all observations belonging to a set $A \in 2^{\mathcal{Z}}$ are fitted to a multi-dimensional distribution Q_A. Then, having an observation $\boldsymbol{\rho}$ related to $\{AP_1, \ldots, AP_{N_{AP}}\}$, the mass $m(A)$ is calculated as follows,

$$m(A) = \frac{Q_A(\boldsymbol{\rho})}{\sum_{A' \in 2^{\mathcal{Z}}, A' \neq \emptyset} Q_{A'}(\boldsymbol{\rho})}, \quad A \in 2^{\mathcal{Z}}, A \neq \emptyset. \tag{2}$$

The quantity $m(A)$ represents the amount of evidence brought by the set of sources $\{AP_1, \ldots, AP_{N_{AP}}\}$ saying that the observation $\boldsymbol{\rho}$ belongs to the set A, A being a singleton, a pair, or more. By taking all the supersets of $\mathcal{Z}$ and not only the singletons, the proposed algorithm uses all available evidences, even if they are uncertain about a single element. Note that $m(A)$ is not the probability of having $\boldsymbol{\rho}$ in A, but only an interpretation of the information brought by the sources by means of observation $\boldsymbol{\rho}$, that is, $m(A)$ could be higher than $m(B)$ even if $A \subset B$.

An adequate notion of the BFT to attribute masses to singleton sets is the pignistic level [11]. It is defined as follows,

$$BetP(A) = \sum_{A \subseteq A'} \frac{m(A')}{|A'|}, \tag{3}$$

where A is a singleton of $2^{\mathcal{Z}}$. The pignistic level is equivalent to the probability of having the observation belonging to the considered set. One could also compute the pignistic level of higher-cardinal supersets. However, only the singleton sets are taken into consideration, as we are interested in determining a level of confidence for the original zones only.

C. Clustering

Clustering aims to organize a set of data into groups called clusters, according to some criteria [12]. Hierarchical clustering builds a hierarchy of clusters or dendrogram driving two strategies: agglomerative or divisive approaches. In the agglomerative or the *bottom up* approach, each observation starts as an independent cluster, and pairs of clusters are merged upon moving up in the hierarchy; whereas in the divisive or *top down* approach, all observations start as one single cluster, and are split upon moving down in the hierarchy [13]. Here, the agglomerative strategy is adopted since it is less complex than the divisive case [14]. To avoid having observations of the same zone in different clusters, the proposed method considers the zones as units. Indeed, a statistical distribution is assigned to each zone, by fitting its corresponding observations to one of the existing statistical distributions. Let $Q_1, \ldots, Q_{N_Z}$ be the fitted multi-dimensional distibutions, defined over a set of parameters, of the zones $Z_1, \ldots, Z_{N_Z}$ respectively. Each zone is then considered as an independent cluster at the beginning of the algorithm. To merge clusters according to a criterion, the agglomerative hierarchical clustering technique measures the dissimilarity between the clusters. Of these criteria are single-linkage, complete linkage, Ward's minimum variance, etc [15].

Since distributions are being clustered here, statistical measures could be applied like Kullback-Leibler divergence, Hellinger distance, total variation distance, etc [16]. The Kullback-Leibler divergence or relative entropy of two distributions Q_i and Q_j of input $\boldsymbol{\rho}$ is defined as

$$D_{KL}(Q_i||Q_j) = \int_{\boldsymbol{\rho}} \log\left(\frac{Q_i(\boldsymbol{\rho})}{Q_j(\boldsymbol{\rho})}\right) Q_i(\boldsymbol{\rho}) d\boldsymbol{\rho}. \tag{4}$$

The relative entropy is asymmetric, always positive and equal to zero when the two distributions are identical. The J-divergence [17] symmetrizes the Kullback-Leibler divergence as follows

$$D_J(Q_i||Q_j) = D_{KL}(Q_i||Q_j) + D_{KL}(Q_j||Q_i). \tag{5}$$

This divergence computes the level of discrepancy or lack of similarity between probability distributions. It is a measure of how different two probability distributions, over the same event space, are [18]. The proposed clustering method employs the J-divergence as the dissimilarity measure to construct the

dendrogram. At each iteration, it merges the two clusters whose distributions have the maximal divergence. Merging two clusters means here a merge of all the observations of the infant clusters and a computation of a new distribution according to the new set of observations. By maximizing the divergence, the infant clusters would be dissimilar, which helps in discriminating between zones of each cluster. The algorithm is iterated until all the zones are merged into one cluster.

After the dendrogram is created, it should be cut based on the desired number of clusters. However, since there is no prior knowledge regarding this parameter, it is calculated by solving an optimization problem that takes into account both inter- and intra- clusters scatters. Several indices have been proposed to solve this problem [19], [20], [21]. A method developed by Krzanowski and Lai [22] finds the optimal number of clusters N_C as follows,

$$N_C = \underset{l}{\operatorname{argmax}} \left| \frac{DIFF(l)}{DIFF(l+1)} \right| \tag{6}$$

such that

$$DIFF(l) = (l-1)^{\frac{2}{N_{AP}}} W(l-1) - (l)^{\frac{2}{N_{AP}}} W(l), \tag{7}$$

where $W(l)$, the within cluster sums of squares, is equal to:

$$W(l) = \sum_{j=1}^{l} \sum_{\substack{k \\ Z_k \in C_j}} \sum_{\eta} ||\boldsymbol{\rho}^{k,\eta} - \mu_j||^2, \tag{8}$$

where $\boldsymbol{\rho}^{k,\eta}$ represents an RSS observation taken in zone Z_k, l takes its first value as $N_Z - 1$ and decreases by 1 at each iteration, μ_j is the mean of the distribution of the cluster C_j, and $|| \cdot ||$ is the Euclidean norm operator.

The dendrogram is cut at a certain level where N_C clusters are obtained, denoted by C_j with $j \in \{1, \ldots, N_C\}$. All infant clusters of each selected cluster are merged yielding a set of zones for each cluster. The set I_j denotes the set of indices of the zones included in the cluster j, that is, $Z_i \in C_j, \quad \forall i \in I_j$.

D. Access Point Selection

The Access Points are the sources of information, and hence the choice of reliable ones is indispensable for the localization process. Practically, the same installed AP in a certain building transmits signals on different terminals, each to give access to a certain population (staff, residents, visitors, etc) and on different channel bands (2.4 GHZ, 5 GHZ, etc). The original emitted power on each network is controlled and changed by the IT services as needed. Though information carries some redundancy, we aim here at using all available evidence to reach the best decision. So, what is meant here by AP selection is the choice of networks and not only physical APs. The observations have N_{AP} components, each one being related to a certain AP, of the set $F = \{AP_n\}, n = 1, 2, \ldots, N_{AP}$. Having the two-level hierarchy, the AP selection is applied at the cluster level, and also at the zones level within each cluster, the aim being to select the most useful APs that are capable of discriminating zones of each cluster and between clusters.

The AP selection algorithm is applied equivalently at zones of each cluster and between clusters. For the sake of simplicity, unique notations for clusters and zones are considered in the following, that is, let z denote either a cluster or a zone within a cluster, and let N denote their numbers. A greedy AP selection method is adopted to maximize the discriminative capacity and minimize the dependency of the selected APs. APs cannot be treated independently, since one that might be useless by itself can provide a significant improvement in the performance when taken with others [23]. Let $F' \subseteq F$ denote one non-empty subset. All the observations at the APs of F' belonging to each entity z_j are thus taken, and they are fitted to a distribution denoted $Q_{F',j}$, $j \in \{1, \ldots, N\}$. The distribution $Q_{F',j}$ is either univariate or multivariate depending on the cardinal of F'.

On one hand, having erroneous APs harms the performance of loclization. But since the computation of the exact error rate is cumbersome especially for high dimensions, the discriminative capacity of APs which is inversely proportional to the error rate is used [24]. Indeed, the farther the distributions $Q_{F',z_1}, \ldots, Q_{F',z_N}$ are one from the other, the more discriminative the AP subset F' is, and thus the less error rate is obtained. The Kullback-Leibler divergence is used to measure such a quantity. The discriminative capacity of a subset of APs $F' \subseteq F$ is then defined as follows,

$$DisC(F') = \sum_{a=1}^{N} \sum_{b=1}^{N} D_{KL}(Q_{F',z_a} || Q_{F',z_b}), \tag{9}$$

$D_{KL}(Q_{F',z_a} || Q_{F',z_b})$ being the Kullback-Leibler divergence measured between the distributions of the observations belonging to zones z_a and z_b, while considering only the APs of F'. The error rate of subset F' is defined as follows [24],

$$\mathcal{E}(F') = 2^{-DisC(F')}. \tag{10}$$

On the other hand, dependency is an important factor in AP selection. Having two or more dependent APs leads to a higher dimensionality and redundant performance. Here, the coefficient of multiple correlation is used as it is a measure of how much an AP is dependent upon other ones. The square of the multiple correlation coefficient of AP_j in a set of APs F' with respect to the group of APs $F' \setminus \{AP_j\}$ is defined as follows,

$$R_j^2 = c_j^T R_{xx,j}^{-1} c_j, \tag{11}$$

where c_j is the column vector with entries $d_{AP_i AP_j}$ for $AP_i \in F' \setminus \{AP_j\}$, $d_{AP_i AP_j}$ being the correlation between APs AP_i and AP_j computed using their observations, c_j^T being the transpose of c_j, and $R_{xx,j}^{-1}$ the inverse of the matrix of entries $d_{AP_i AP_{i'}}$ for AP_i and $AP_{i'} \in F' \setminus \{AP_j\}$. The dependency between all the APs of a set F' is the average multiple correlation coefficient of all $AP_j \in F'$, namely

$$\mathcal{R}(F') = \sum_{j} \frac{R_j}{|F'|}, \tag{12}$$

where $|F'|$ is the cardinal of F'.

The objective of the AP selection technique is to find the subset $F_s \subseteq F$ such that both $\mathcal{E}(F_s)$ and $\mathcal{R}(F_s)$ are simultaneously minimized. This can be solved using any multi-objective optimization technique such as Pareto front [25]; yet these solutions require the knowledge of $\mathcal{E}$ and $\mathcal{R}$ for all subsets, which is computationally unfeasible for large number of APs. For this purpose, a greedy search algorithm with backward elimination strategy is applied to choose this subset. One starts with the whole set of APs and progressively eliminates an AP, whose elimination satisfies a function that considers the two objectives mentioned above. Let F_y be the AP subset chosen at iteration $y \geq 1$, with $F_0 = F$ and the cardinal $|F_y|$ of F_y equal to $N_{AP} - y$. At each iteration $y \geq 1$, all the subsets of F_{y-1} having $N_{AP} - y$ elements are considered. Let $F_y^{(\lambda)}$, $\lambda = 1, \ldots, N_{AP} - y + 1$, denote these subsets. We define the function $g_y\big(F_y^{(\lambda)}\big)$ as follows,

$$g_y(F_y^{(\lambda)}) = \alpha \frac{\mathcal{E}(F_{y-1}) - \mathcal{E}(F_y^{(\lambda)})}{\max(\mathcal{E}(F_{y-1}), \mathcal{E}(F_y^{(\lambda)}))} + (1-\alpha) \frac{\mathcal{R}(F_{y-1}) - \mathcal{R}(F_y^{(\lambda)})}{\max(\mathcal{R}(F_{y-1}), \mathcal{R}(F_y^{(\lambda)}))}, \tag{13}$$

where $\alpha \in [0, 1]$ is a tradeoff parameter chosen by the user to assign a weight for each objective. A positive value of $g_y\big(F_y^{(\lambda)}\big)$ means that the subset $F_y^{(\lambda)}$ is better than F_{y-1} in optimizing the objectives. The greater $g_y(\cdot)$ is, the better the subset is. This leads to a selected subset at iteration y, $F_y = \arg\max_\lambda g_y\big(F_y^{(\lambda)}\big)$. A negative value of g means that there is no significant improvement in the objectives for the considered parameters and hence iterations stop when all $g_k\big(F_y^{(\lambda)}\big)$, $\lambda = 1, \ldots, N_{AP} - y + 1$, are negative and one chooses the set of APs $F_s = F_{y-1}$. This algorithm is applied at the clusters level to yield F_s and at the zones level of each cluster C_j to yield $F_{s,j}, j \in \{1, \ldots, N_C\}$.

E. Confidence-based zone estimation

We consider the outputs of the two preceding phases. Let C_j, $j \in \{1, \ldots, N_C\}$ be the set of created clusters, and I_j be the set of indices of the zones included in cluster j, that is, $Z_k \in C_j$, $\forall k \in I_j$. Let F_s be the selected subset of available APs at the clusters level, and $F_{s,j}$ the selected subset at the zones level of cluster C_j. The computations of Eq. (3) are applied at the two levels of the hierarchy, leading respectively to pignistic levels of all clusters and zones within each cluster $BetP^C(\{C_j\})$, $j \in \{1, \ldots, N_C\}$, and $BetP^j(\{Z_k\})$, $k \in I_j$. To attribute a confidence level by the observation model to each zone, pignistic levels of zones and clusters are combined as follows,

$$\mathcal{C}f(Z_k) = BetP^C(\{C_j\}) \times BetP^j(\{Z_k\}), \tag{14}$$

with $k \in I_j, j \in \{1, \ldots, N_C\}$.

III. Experiments

To evaluate the performance of the proposed method, real experiments were conducted in a WLAN environment at the first floor of the Living Lab at the University of Technology of Troyes, France. The considered floor of approximated area of 500 m^2 is partitioned into nineteen zones. A personal computer, with a WiFi scanner software, can distinguish APs of the network throughout their MAC addresses. It measures then the RSS of their transmitted signals. Note that 38 AP networks could be detected at the considered area. Sets of 50 measurements were taken in each zone, of which 30 were randomly used to construct the databases, and the others were kept for test. The collected RSS of the database were statistically fitted according to a significance level of 0.02. Table I shows the results of applying the observation model to 380 test points and the influence of each phase over the percentage of accuracy and the processing time. An estimation is said to be correct if the algorithm assigns the highest confidence level to the right zone. As this table clearly shows, the basic observation model attains an accuracy of 82.84% over the training data and 81.63% over new data. This low accuracy percentage is due to the wide overlapping of the various functions representing the distributions of the data in the different zones. However, when the two-level hierarchical clustering was carried out, a great enhancement in the percentage of accuracy was noted (88.21% over training data and 86.26% over new data). This amelioration is at the expense of the processing time. It is clear that both the offline training time and the online test time were almost doubled. In addition, the AP selection phase had a significant impact on the overall process. An accuracy of 92.78% on training data and 90.42% over new data could be obtained with a slight gain in the online test time, yet with an increase in the offline training time.

In this section, the proposed method is compared to some of the well-known classification techniques such as k-nearest neighbors, naive Bayes, multinomial logistic regression, neural networks, and SVM. The parameters of these methods were determined and tuned using a ten-fold cross validation. For k-nearest neighbors, the optimal number of neighbors was found to be 19. For naive Bayes and multinomial logistic regression, the maximum likelihood estimate was used to evaluate the probability of each zone. As for neural networks, radial basis functions were used as activation functions for a one single hidden layer. A Gaussian kernel was used for SVM. Table II shows the percentage of accuracy and the processing time of the proposed technique compared to the aforementioned methods. The proposed method outperforms all the other ones in terms of localization accuracy, with a competitive processing time.

IV. Conclusion and Future Work

This paper presented a confidence-based localization technique for indoor environments. The method proposes an observation model, in a belief functions framework, that creates supersets of zones and affiliates masses to each one using the APs as sources of information. A two-level hierarchy is then created using a hierarchical clustering technique. An AP selection approach is also proposed to choose the best subset of APs in terms of discriminative capacity and redundancy. Real experiments in a Living Lab demonstrate the effectiveness of the proposed localization approach and its competence as compared to state-of-the-art techniques. Future work will focus

TABLE I: Influence of each phase of the observation model on the accuracy and the processing time.

Applying the observation model	Accuracy over training data / new data (%)	training time / test time (s)
without clustering and AP selection	82.84 / 81.63	45 / 0.1174
without AP selection	88.21 / 86.26	92 / 0.2369
as it is	92.78 / 90.42	126 / 0.2172

TABLE II: Performance of methods in terms of localization accuracy and processing time.

Method	Accuracy over training data / new data (%)	training time / test time (s)
K-nearest neighbors	82.53 / 80.15	14 / 0.1311
Naive Bayes	83.33 / 80.95	42 / 0.1042
Multi nomial logistic regression	84.12 / 85.33	76 / 0.1559
Neural networks	86.51 / 85.71	83 / 0.1883
SVM	88.10 / 87.30	96 / 0.1912
Proposed method	92.78 / 90.42	126 / 0.2172

on the usage of mobility of sensors in combination of the observation model to enhance the overall accuracy.

ACKNOWLEDGEMENT

The authors would like to thank the FEDER and the Grand Est Region in France for funding this work, project LIPAD, and the Living Lab at the University of Technology of Troyes for allowing them to realize the experiments.

REFERENCES

[1] S. K. Gharghan, R. Nordin, M. Ismail, and J. A. Ali, "Accurate wireless sensor localization technique based on hybrid pso-ann algorithm for indoor and outdoor track cycling," *IEEE Sensors Journal*, vol. 16, no. 2, pp. 529–541, 2016.

[2] S. A. Mitilineos, D. M. Kyriazanos, O. E. Segou, J. N. Goufas, and S. C. Thomopoulos, "Indoor localisation with wireless sensor networks," *Progress In Electromagnetics Research*, vol. 109, pp. 441–474, 2010.

[3] D. Lymberopoulos and J. Liu, "The microsoft indoor localization competition: Experiences and lessons learned," *IEEE Signal Processing Magazine*, vol. 34, pp. 125–140, 2017.

[4] Y.-C. Liu, Y.-K. Ou, S.-N. Lin, and C.-W. Fang, "A study of the indoor walking navigation system for patients with early-stage Alzheimers disease," in *International Conference on Computer, Networks and Communication Engineering (ICCNCE 2013)*, Atlantis Press, 2013.

[5] E. Smirnov, N. Shilov, and A. Kashevnik, "Ontology-based mobile smart museums service approach for small & medium museums," 2012.

[6] S. Wang, S. Fidler, and R. Urtasun, "Lost shopping! monocular localization in large indoor spaces," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2695–2703, 2015.

[7] D. L. Lee and Q. Chen, "A model-based WiFi localization method," in *Proceedings of the 2nd international conference on Scalable information systems*, p. 40, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2007.

[8] K. K. Khedo, D. Sathan, R. Elaheebocus, R. K. Subramanian, and S. Rughooputh, "Overlapping zone partitioning localisation technique for RFID," *International Journal of Ubiquitous Computing (IJU)*, vol. 1, no. 2, pp. 20–32, 2010.

[9] A. Chriki, H. Touati, and H. Snoussi, "Svm-based indoor localization in wireless sensor networks," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2017 13th International*, pp. 1144–1149, IEEE, 2017.

[10] D. AlShamaa, F. Mourad-Chehade, and P. Honeine, "Zoning-based localization in indoor sensor networks using belief functions theory," in *Signal Processing Advances in Wireless Communications (SPAWC), 2016 IEEE 17th International Workshop on*, pp. 1–5, IEEE, 2016.

[11] P. Smets, "Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem," *International Journal of approximate reasoning*, vol. 9, no. 1, pp. 1–35, 1993.

[12] S. Mehrotra and S. Kohli, "Data clustering and various clustering approaches," *Intelligent Multidimensional Data Clustering and Analysis*, p. 90, 2016.

[13] K. Sasirekha and P. Baby, "Agglomerative hierarchical clustering algorithm-a," *International Journal of Scientific and Research Publications*, p. 83, 2013.

[14] C. D. Manning, P. Raghavan, H. Schütze, *et al.*, *Introduction to information retrieval*, vol. 1. Cambridge university press Cambridge, 2008.

[15] F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: an overview," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 2, no. 1, pp. 86–97, 2012.

[16] M. Basseville, "Divergence measures for statistical data processingan annotated bibliography," *Signal Processing*, vol. 93, no. 4, pp. 621–633, 2013.

[17] F. Nielsen and R. Nock, "Generalizing skew jensen divergences and bregman divergences with comparative convexity," *IEEE Signal Processing Letters*, 2017.

[18] C. Röver and T. Friede, "Discrete approximation of a mixture distribution via restricted divergence," *Journal of Computational and Graphical Statistics*, vol. 26, no. 1, pp. 217–222, 2017.

[19] L. Rokach and O. Maimon, "Clustering methods," in *Data mining and knowledge discovery handbook*, pp. 321–352, Springer, 2005.

[20] M. A. Islam, B. Z. Alizadeh, E. R. van den Heuvel, R. Bruggeman, W. Cahn, L. de Haan, R. S. Kahn, C. Meijer, I. Myin-Germeys, J. van Os, *et al.*, "A comparison of indices for identifying the number of clusters in hierarchical clustering: A study on cognition in schizophrenia patients," *Communications in Statistics: Case Studies, Data Analysis and Applications*, vol. 1, no. 2, pp. 98–113, 2015.

[21] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 2, pp. 411–423, 2001.

[22] A. Fischer, "On the number of groups in clustering," *Statistics & Probability Letters*, vol. 81, no. 12, pp. 1771–1781, 2011.

[23] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182, 2003.

[24] O. S. Jahromi, *Multirate statistical signal processing*. Springer Science & Business Media, 2007.

[25] P. Ngatchou, A. Zarei, and A. El-Sharkawi, "Pareto multi objective optimization," in *Intelligent systems application to power systems, 2005. Proceedings of the 13th international conference on*, pp. 84–91, IEEE, 2005.