Mobility-based Tracking Using WiFi RSS in Indoor Wireless Sensor Networks
Daniel Alshamaa, Farah Chehade, Paul Honeine

To cite this version:

HAL Id: hal-01965903
https://hal.archives-ouvertes.fr/hal-01965903
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.
Mobility-based Tracking Using WiFi RSS in Indoor Wireless Sensor Networks

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—Tracking of mobile sensors is an important research issue in wireless sensor networks. This paper presents a zoning-based tracking technique that works efficiently in indoor environments. The targeted area is composed of several zones, the objective being to determine the zone of the mobile sensor in a real-time tracking process. The proposed method creates a belief functions framework that combines evidence using the sensors mobility and observations. To do this, a mobility model is proposed by using the previous state of the sensor and its assumed maximum speed. Also, an observation model is constructed based on fingerprints collected as WiFi signals strengths received from surrounding Access Points. Real experiments demonstrate the effectiveness of this approach and its competence compared to state-of-the-art methods.

Index Terms—Belief functions, evidence fusion, mobility, tracking, WiFi signals.

I. INTRODUCTION

Recent advances in computing and communication have enabled the proliferation of wireless sensor networks. These networks are tremendously being 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]. This paper addresses the zoning-based tracking problem in indoor environments, where the zone of the localized sensor in real-time is of interest. This issue is important for health-care applications for instance, where Alzheimer’s patients might be lost in their nursing home [3], in museums for supporting guides and emergency management [4], for large malls to facilitate shopping [5], etc where locating people in a specific zone of such environments is completely sufficient.

Many existing works have been proposed to tackle the localization problem. The GPS technology for example, is widely integrated in vehicle tracking systems [6]. However, it has limitations in indoor environments due to the large attenuation caused by buildings’ walls and ceilings [7]. To extend the capability of mobile localization applications in indoor environments, researchers worked on alternative solutions based on various types of signals like ultra wide band, WiFi, zigbee, bluetooth, etc [8], [9], [10]. One of the advantages of WiFi signals over the others is that one can rely only on the Access Points (APs) present inside the building, with no need to additional hardware. For that reason, several localization algorithms and connectivity-based methods that use received signal strengths (RSS) of WiFi signals were developed [11], [12], [13]. These methods depend on a pathloss model, which is not efficient especially in indoor environments. Alternatively, techniques that employ fingerprinting are widely implemented. They collect strengths of WiFi signals at exact reference positions in a database and then apply the k-nearest neighbors scheme [14], neural networks [15] or kernel-based learning [16], [17] to solve the localization problem.

This paper proposes a new tracking method for indoor environments. The proposed method uses belief functions to estimate the sensors zones by combining the evidence related to the sensors mobility and observations, as described in the following. The method makes use of the sensors mobility by assuming a maximum speed of movement of sensors in indoor environments. This allows a prediction of the next possible destinations of the mobile sensor, and hence leading to a mobility model. Here, the belief functions framework is used to propagate the previous step evidence till the current one. Moreover, the proposed method consists in constructing a fingerprinting database that associates to each zone a set of WiFi signals strengths collected from the APs. This database is used with the belief functions theory to create superset of zones and affiliate evidence to each one according to each AP. The APs, which are the sources of information, are discounted according to their error rate. Their evidence is then combined via the belief functions fusion rule. Mobility and observation evidence is then combined to determine a level of confidence of having the mobile sensor residing in each zone. By taking into account information uncertainty, the proposed method yields several possibilities of zones with different levels of confidence of covering the new observation. Real experiments are conducted in a Living Lab to track elderly people. Results show the effectiveness of the proposed method, compared to state-of-the-art algorithms.

The rest of the paper is organized as follows. Section II describes the problem. Section III presents the tracking approach. Section IV demonstrates the experimental results and the comparison to state-of-the-art algorithms. Section V concludes the paper.
The tracking problem consists of estimating the mobile sensor’s zone in real time using its mobility and the signals strengths or RSS it collects from the APs. It is tackled in the following manner. Let

- $N_Z$ be the number of zones of the targeted area, denoted by $Z_k$, $k = 1, 2, \ldots, N_Z$;
- $N_{AP}$ be the number of detected APs, denoted by $AP_n$, $n = 1, 2, \ldots, N_{AP}$;
- $\rho_i$ be the vector of size $N_{AP}$ of RSS measurements collected by the mobile sensor at the instant $t$ from surrounding Access Points;
- $v_{max}$ be the maximum speed of the mobile sensor in the indoor environment.

The aim of the proposed algorithm is to find a function $h : \mathbb{R}^{N_{AP}} \to [0, 1]^{N_Z}$ such that $h(\rho_i) = (C_f(Z_1), \ldots, C_f(Z_NZ))$, where $C_f(Z_k)$ is the level of confidence of having the mobile sensor of observation $\rho_i$ residing in the zone $Z_k$ at the instant $t$.

III. TRACKING APPROACH

In this section, the tracking approach is presented. It consists of assigning evidence through a mobility model and an observation model, and combining them in a belief functions framework, as shown in Fig. 1.

A. Mobility model

Mobility plays an important role in providing location-based services and tracking of mobile sensors. Here a mobility-based approach is implemented that assumes a maximum speed of the mobile sensors, and is used as a source of information to track in real time the movement of this sensor. We propose two models: the basic model is based on the original succession of zones; the advanced one is more accurate based on the transition between created sub-zones and necessitates a specific data acquisition phase. The two models are described in the following paragraphs.

1) Basic mobility model: Let $v_{max}$ be the maximum speed of the mobile sensor in the target area, $\Delta_{loc}$ the time interval in which the localization algorithm is executed, and $d_{min,ij}$ the minimal geographical distance between the two zones $Z_i$ and $Z_j$. To determine $d_{min,ij}$, a point is placed in zone $Z_i$ such that it is the closest possible to zone $Z_j$, and another one is placed in the zone $Z_j$ such that it is the closest possible to $Z_i$. The maximum distance that the sensor can travel is then deduced $d_{max} = v_{max} \times \Delta_{loc}$. Let $p_{ij}, i, j \in \{1, \ldots, N_Z\}$, denote the coefficient of transition from zone $Z_i$ to zone $Z_j$ within the localization period $\Delta_{loc}$. Then,

\[
\rho_{ij} = \begin{cases} 
0, & \text{if } d_{max} < d_{min,ij}; \\
1, & \text{if } d_{max} \geq d_{min,ij}.
\end{cases}
\]

Being in a zone $Z_i$ at instant $t - 1$, the sensor could be at instant $t$ at any zone $j \in \{1, \ldots, N_Z\}$ of the ones having $p_{ij} = 1$. These zones are called the following zones of $Z_i$. The confidence of $Z_i$ at time $t - 1$ is then propagated to time $t$ by distributing $C_f(Z_i)$ to $Z_j$. The mobility evidence given to a zone $Z_k$ at time $t$ is the aggregation of all evidence deduced from its preceding ones having $p_{ij} = 1$, $\forall i \in \{1, \ldots, N_Z\}$. This leads to a mobility mass at time $t$ computed in the following manner,

\[
m_{f_{t-1}}(Z_k) = \sum_{i=1}^{N_Z} p_{ik} \times \frac{C_f(Z_i)}{\sum_{j=1}^{N_Z} p_{ij}}.
\]

2) Advanced mobility model: In this paragraph, a more precise mobility model is presented. This model requires a specific data acquisition phase as described in the following. Each zone $Z_i$ is divided into $N_X$ sub-zones $X_{i,\ell}$ according to its architecture: $N_X - 1$ connection sub-zones and one main sub-zone $X_{i,N_X}$. Each connection sub-zone is a section area in front of a door connecting $Z_i$ with neighbor zones, as indicated in Fig. 2. Its dimensions are defined in a way to cover all possible positions in the zone $Z_i$ at which the sensor could cross the door to go to a neighboring zone within the localization period. The main sub-zone $X_{i,N_X}$ is the remaining section area in $Z_i$. A sensor being in the main sub-zone of zone $Z_i$ at time $t - 1$ would remain in the same zone at time $t$. In a data acquisition phase, RSS values are collected in each sub-zone and fitted to a multi-dimensional statistical distribution. Let $Q_{i,\ell}(\cdot)$ be the distribution representing the data of the connection sub-zone $X_{i,\ell}$, $\ell \in \{1, \ldots, N_X - 1\}$, and $Q_{i,N_X}(\cdot)$ be the distribution representing the main sub-zone $X_{i,N_X}$. Having the previous observation $\rho_{t-1}$, membership weights $q_{t-1}(X_{i,\ell})$, $i \in \{1, \ldots, N_Z\}, \ell \in \{1, \ldots, N_X\}$ could be computed to quantify the membership of the sensor to any sub-zone of each zone $Z_i$ at $t - 1$. This is performed by calculating the probability resulting from each fitted distribution with respect to the previous observation followed by a normalization phase,

\[
q_{t-1}(X_{i,\ell}) = \frac{Q_{i,\ell}(\rho_{t-1})}{\sum_{\ell=1}^{N_X} Q_{i,\ell}(\rho_{t-1})}.
\]

The confidence of each zone $Z_i$ at $t - 1$ is then converted to its sub-zones in the following manner,

\[
C_{f_{t-1}}(X_{i,\ell}) = Q_{i,\ell}(Z_i) \times q_{t-1}(X_{i,\ell}).
\]
Let $r_{i,t,j}$ be the coefficient of transition from the connection sub-zone $X_{i,t}, t \in \{1, \ldots, N_X\}$ of zone $Z_i, i \in \{1, \ldots, N_Z\}$ to original zone $Z_j, j \in \{1, \ldots, N_Z\}$, such that

$$r_{i,t,j} = \begin{cases} 0, & \text{if } d_{\text{max}} < d_{\text{min},i,t,j}; \\ 1, & \text{if } d_{\text{max}} \geq d_{\text{min},i,t,j}, \end{cases} \quad (5)$$

where $d_{\text{min},i,t,j}$ is the minimal distance between connection sub-zone $X_{i,t}$ and zone $Z_j$. The mass associated to each zone by the mobility model can be thus deduced,

$$m_{\text{MM},t}(Z_k) = \sum_{i=1}^{N_Z} \sum_{\ell=1}^{N_X} C_{i,t,k}(X_{i,t}) \cdot \frac{r_{i,t,j}}{\sum_{j=1}^{N_Z} r_{i,t,j}} \quad (6)$$

Here the mass of $Z_k$ at time $t$ is the aggregation from $t-1$ of the confidence of its main zone and a part of the confidence of the connection sub-zones of other zones able to lead to it.

### B. Observation model

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 mass to each zone.

#### 1) Mass association

The observation model consists of fitting the RSS observations into statistical distributions, and using the belief functions theory as a framework for mass association and evidence fusion. 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 $Z = \{Z_1, \ldots, Z_{N_Z}\}$ be the set of all possible zones and let $2^Z$ be the set of all the supersets of $Z$, i.e., $2^Z = \{\emptyset, \{Z_1\}, \ldots, Z\}$. The cardinal of $2^Z$ is equal to $|2^Z| = 2^{N_Z}$, where $|Z|$ denotes the cardinal of $Z$. One fundamental function of the BFT is the mass function, also called the basic belief assignment (BBA). A mass function $m_{AP_n,t}(\cdot)$ is a mapping from $2^Z$ to the interval $[0,1]$, defined according to a certain source $AP_n, n \in \{1, \ldots, N_{AP}\}$, it satisfies:

$$\sum_{A \in 2^Z} m_{AP_n,t}(A) = 1. \quad (7)$$

The mass $m_{AP_n,t}(A)$ given to $A \in 2^Z$ stands for the proportion of evidence, brought by the source $AP_n$ at instant $t$, saying that the observed variable belongs to $A$.

In order to define the APs BBAs, all observations related to each AP belonging to a superset $A \in 2^Z$ are fitted to a distribution $Q_{AP_n,A}$. Then, having an observation $\rho_{n,t}$ related to $AP_n, n \in \{1, \ldots, N_{AP}\}$, the mass $m_{AP_n,t}(A)$ is calculated as follows,

$$m_{AP_n,t}(A) = \frac{Q_{AP_n,A}(\rho_{n,t})}{\sum_{A' \in 2^Z, A' \neq A} Q_{AP_n,A'}(\rho_{n,t})}, \quad A \in 2^Z, A \neq \emptyset. \quad (8)$$

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

#### 2) Discounting operation

The detected APs are not completely reliable. Indeed, each AP could yield an erroneous interpretation of evidence for some observations. In order to correct this, one can discount the BBAs of Eq. (8) by taking into account the error rate of the AP. The discounted BBA $m_{AP_n,t}(\cdot)$ related to $AP_n$ having an error rate $\alpha_n$ is deduced from the BBA $m_{AP_n,t}(\cdot)$ as follows [18],

$$m_{AP_n,t}(A) = \begin{cases} (1 - \alpha_n)m_{AP_n,t}(A), & \text{if } A \in 2^Z, A \neq \emptyset; \\ \alpha_n + (1 - \alpha_n)m_{AP_n,t}(A), & \text{if } A = \emptyset. \end{cases} \quad (9)$$

By doing this, the amounts of evidence given to the supersets of $Z$ are reduced, and the remaining evidence is given to the whole set $Z$.

Now, to compute the error rate of a certain source $AP_n$, consider an observation $\rho_{n,t}$ being truly in $A$. The source $AP_n$ is assumed not reliable if, according to $\rho_{n,t}$, it associates more evidence to any set other than $A$, that is, the mass associated to $A$ is less than the mass of another subset of $2^Z$. Since the BBAs are defined using the statistical distributions related to each set, then an AP is erroneous for all observations of $A$ where $Q_{AP_n,A}(\rho_{n,t})$ is less than any $Q_{AP_n,A'}(\rho_{n,t})$, for any $A' \neq A$. Let $\epsilon_n(A)$ be the error rate related to the set $A$ with respect to $AP_n$. Then,

$$\epsilon_n(A) = \int_{D_{n,A}} Q_{AP_n,A}(\rho) d\rho, \quad (10)$$

such that $D_{n,A}$ is the domain of error of set $A$ according to $AP_n$, defined as follows,

$$D_{n,A} = \{\rho \mid Q_{AP_n,A}(\rho) \leq \max_{A' \in 2^Z, A' \neq A} (Q_{AP_n,A'}(\rho))\}. \quad (11)$$

The error rate $\epsilon_n$ of $AP_n$ is then the average error of all sets according to this AP, namely

$$\alpha_n = \frac{\sum_{A \in 2^Z} \epsilon_n(A)}{|2^Z|}. \quad (12)$$
3) Fusion of evidence: According to the information retrieved from the APs, the mass functions $m_{AP_{n,t}}(\cdot)$ are defined. Combining the evidence consists in aggregating the information coming from all the APs [19]. The mass functions can then be combined using the conjunctive rule of combination as follows,

$$m_{\gamma,t}(A) = \sum_{A^{(n)} \in 2^Z \cap \bigwedge_{n=1}^{N} A^{(n)} = A} \prod_{n=1}^{N} m_{AP_{n,t}}(A^{(n)}),$$  \hspace{1cm} (13)

for all the sets $A \in 2^Z$, with $A^{(n)}$ is the set $A$ with respect to the Access Point $AP_n$. This fusion rule leads to a more informative and specialized mass function [20]. The mass function is then normalized, leading to the Dempster rule of combination:

$$m_{\Phi,t}(A) = \frac{m_{\gamma,t}(A)}{\sum_{A' \in 2^Z} m_{\gamma,t}(A')}$$ \hspace{1cm} (14)

4) Pignistic transformation: An adequate notion of the BFT to attribute masses to singleton sets is the pignistic level [21]. It is defined as follows,

$$BetP_t(A) = \sum_{A' \in A} \frac{m_{\Phi,t}(A')}{|A'|}.$$ \hspace{1cm} (15)

where $A$ is a singleton of $2^Z$. 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. Hence, the mass associated to each zone by the basic observation model at each instant $t$ can be computed as follows,

$$m_{\Theta,t}(Z_k) = BetP_t(\{Z_k\}), k \in \{1, \ldots, N_Z\}.$$ \hspace{1cm} (16)

C. Confidence-based zone estimation

The associated masses by the mobility model and the observation model are combined by fusing the evidence of the two models to yield a confidence level of having the mobile sensor resides in each zone as follows,

$$C_f(Z_k) = m_{\Phi,t}(Z_k) = \frac{m_{\Theta,t}(Z_k) \times m_{M,t}(Z_k)}{\sum_{k=1}^{N_Z} m_{\Theta,t}(Z_k) \times m_{M,t}(Z_k)}.$$ \hspace{1cm} (17)

The zone having the highest confidence is then selected. It is worth noting that this method yields ranked results, allowing for a second zone choice if the first one was erroneous.

IV. EXPERIMENTS

To evaluate the performance of the proposed method, real experiments were conducted in a Living Lab addressing the localization of elderly people. As shown in Fig. 3, the considered floor of approximated area of 500 m$^2$ is partitioned into eighteen 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 APs networks could be detected at the considered area. Sets of 30 measurements were taken in each zone to create the training database. Other 10 trajectories, each of 50 observations, were considered to validate the proposed algorithm. The measures were taken in random positions and orientations of the personal computer. The collected RSS of the database were statistically fitted according to a significance level of 0.02. The two mobility models, each by its own, were used to affiliate certain evidence to each zone by considering a maximum speed of 1.5 m/s. The observation model was then constructed by using the fitted distributions in the belief functions framework. The method generated then a set of masses using the presented observation model. It is noted that by using the proposed method with the basic mobility model, an average accuracy of 88.67% could be attained. This important enhancement in the overall average accuracy is due to the presence of diametrically opposed erroneous zones with respect to Access Points. These types of errors could be easily recovered by the basic mobility model. Another amelioration in the overall accuracy to 91.82% has been noted upon combining the advanced mobility model with the observation model. The advantage of this proposed model is in the high accuracy achieved in assigning masses for sub-zones. This is due to the large number of APs selected inside each zone, which yields a more accurate decision when combined.

The proposed method is compared to well-known classification techniques such as k-nearest neighbors, naïve Bayes, multinomial logistic regression, neural networks, and support vector machines (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 used to estimate the class membership was found to be 19. For naïve Bayes and multinomial logistic regression, the maximum likelihood estimate was used to evaluate the probability of having the data instance belong to each class. 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 I 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 classification accuracy, with
TABLE I: Performance of methods on tracking in terms of accuracy and processing time.

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy (%)</th>
<th>Training time (s)</th>
<th>Execution time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-nearest neighbors</td>
<td>81.82</td>
<td>14</td>
<td>0.1311</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>78.76</td>
<td>42</td>
<td>0.1042</td>
</tr>
<tr>
<td>Multi nominal logistic regression</td>
<td>82.94</td>
<td>76</td>
<td>0.1559</td>
</tr>
<tr>
<td>Neural networks</td>
<td>83.82</td>
<td>83</td>
<td>0.1883</td>
</tr>
<tr>
<td>SVM</td>
<td>86.47</td>
<td>96</td>
<td>0.1912</td>
</tr>
<tr>
<td>Proposed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observation model</td>
<td>84.91</td>
<td>87</td>
<td>0.1774</td>
</tr>
<tr>
<td>Observation model + basic mobility model</td>
<td>88.67</td>
<td>102</td>
<td>0.2157</td>
</tr>
<tr>
<td>Observation model + advanced mobility model</td>
<td>91.82</td>
<td>129</td>
<td>0.2388</td>
</tr>
</tbody>
</table>

comparable processing time.

V. CONCLUSION AND FUTURE WORK

This paper presented a new confidence-based tracking technique for indoor environments. The method creates a belief functions framework that uses two models: a mobility model based on assuming a maximum speed of sensors in indoor environments, where two models are proposed to associate an evidence to each zone. The basic mobility model is based on the original succession of zones, while the advanced one creates sub-zones and requires an additional data acquisition phase; and an observation model, based on the belief functions theory, that creates supersets of zones and affiliates masses to each one using the APs as sources of information. The mobility and observation models are combined in the created belief functions framework to determine a level of confidence of having the mobile sensor resides in each zone. Real experiments in a Living Lab demonstrate the effectiveness of the proposed tracking algorithm and its competence as compared to state-of-the-art techniques. Future work will focus on an extended version of the observation model to enhance the overall accuracy. Moreover, tracking of sensors on multi-floor buildings will also be investigated.

ACKNOWLEDGEMENT

The authors would like to thank the ERDF 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 and supplying them with all required information.

REFERENCES