Abstract—WiFi localization problem is basically a multi-sensor data fusion. This paper investigates the use of Bayesian and non-Bayesian Dempster Shafer (DS) data fusion in the context of WiFi-based indoor positioning via fingerprinting. Two novel DS mass choices are discussed. The positioning results are based on real-field measurement data from nine distinct multi-floor buildings in two countries. It is shown that a proper mass choice is crucial in DS processing and that, in spite of taking into account the data uncertainty, the DS data fusion is not offering significant advantage in terms of positioning performance over the Bayesian data fusion.

Keywords—3D localization; Bayesian data fusion; Dempster Shafer (DS) theory; Received Signal Strength (RSS)

I. INTRODUCTION AND MOTIVATION

Received Signal Strength (RSS)-based indoor localization has been gaining more and more attention in the research and commercial arenas in the last few years. WiFi-based indoor positioning is one of the most promising technologies, because of three main factors: 1) WiFi-s are widely deployed in indoor scenarios, 2) a mobile device in an indoor scenario can usually hear a large number of WiFi transmitters or Access Points (APs), 3) every mobile device is able to measure the RSS from various APs in range without the need of additional hardware. Moreover, it can offer satisfactory coverage because of the typical high density of APs in various buildings of interest, and it is a preferred candidate for future mass-market global indoor localization solutions [8][11].

RSS-based fingerprinting typically consists of two stages: a training stage (done offline, either in a manual or in a crowdsourced mode) and an estimation stage [9][11]. In the training stage, RSS values from all APs in a building are collected at various grid points and they are stored in a database, forming the so-called fingerprints. The database is stored on the server of the Location Service Provider (LSP). In the estimation phase, the mobile sends its request to be positioned to the LSP server. In mobile-centric approaches, which are the focus of this paper, the mobile receives a part of the training database (e.g., corresponding to a certain area where the mobile was identified), and it makes its own measurements to the APs in its range. Based on the heard APs and on the comparison with the information stored in the database, the mobile then estimates its position. Fingerprinting is indeed one of the most widespread location solutions [6][13][14][15]. We focus on the fingerprinting methods in here, because with the advent of crowdsourced data gathering and simultaneous localization and mapping (SLAM) algorithms [18], WiFi fingerprinting is becoming more and more a realistic solution to indoor positioning.
positioning accuracy over the Bayesian combining. We also 
show that such accuracy improvement is also highly dependent 
on the mass choice, and that, with an inadequate mass choice, 
one can lose in the accuracy compared to Bayesian data fusion.

II. FINGERPRINTING ALGORITHMS

Fingerprinting algorithms are typically based on two stages: 
training off-line stage and estimation phase [8][9][10][11]. In the 
training phase, the LSP collects on a server the measurements 
in the buildings of interest, in the form of 
\((x_i, y_i, z_i, T_{i, ap})\) where \(x_i, y_i, z_i\) are the 3D coordinates of 
the measurement point \(i\) with respect to a local reference 
coordinate system, \(i = 1, \ldots, N_{grid}\), \(N_{grid}\) is the total number of 
measurements or grid points, and \(T_{i, ap}\) is the RSS in the 
training phase, in logarithmic scale (dBm), coming from the 
ap-th AP, in the \(i\)-th grid point. Here, \(ap = 1, \ldots, N_{ap}\), with \(N_{ap}\) being the number of heard APs in the training phase 
at the \(i\)-th point. We denote by \(N_{ap}\) the total number of APs in 
the building of interest (each AP means an individual MAC 
address, but several APs can be at the same physical location). 
During the measurement process, several RSS measurements 
are taken in the same grid point, at different time instants, and 
with different orientations of the mobile device. In the 
database, we store only an average value over all times and 
orientations (e.g., the mean over all the RSS values in linear 
scale in the same point). To reduce the size of the server 
database, the grid points can be saved only with a certain grid 
step in \(x\) and \(y\) directions. For example, a 5 m grid is an 
adequate trade-off between the number of stored parameters 
and expected positioning accuracy based on our studies. An 
example of the histogram of the residual RSS with respect to 
the mobile is situated in the \(i\)-th grid point is:

\[
O_{ap} = \frac{1}{\sqrt{2\pi}\sigma_{ap}} \exp \left( -\frac{(O_{ap} - T_{i, ap})^2}{2\sigma_{ap}^2} \right)
\]

where \(\sigma_{ap}^2 = \sigma_{ap, meas}^2 + \sigma_{ap, shad}^2\) is a noise variance that lumps 
together the measurement \(\sigma_{ap, meas}\) (see Fig. 1) and shadowing 
errors \(\sigma_{ap, shad}\) effects.

Fig. 1. Example of the distribution of RSS residuals in forming the training 
database. Statistics over all AP in the building.

The probability above is in fact the likelihood 
\(p(\text{location} | \text{observation})\) that the mobile is located at grid 
point \(i\), being given the observation data from \(ap\)-th AP. If we apply 
the Bayesian rule, we can write:

\[
p(\text{location} | \text{observation}) = \frac{p(\text{observation} | \text{location})p(\text{location})}{p(\text{observation})}
\]

\(p(\text{location})\) is the prior or the initial belief in the mobile 
location. In the absence of prior knowledge, maximizing 
\(p(\text{location} | \text{observation})\) is the same as maximizing 
\(p_{i, ap} = p(\text{observation} | \text{location})\).

In the Bayesian linear combining [7], the evidence from 
various sources is combined linearly. The probability \(p_{i}\) that 
the mobile is situated in the \(i\)-th grid point is:

\[
p_i = \sum_{ap=1}^{N_{ap}} p_{i, ap} = \sum_{ap=1}^{N_{ap}} \frac{1}{\sqrt{2\pi}\sigma_{ap}} \exp \left( -\frac{(O_{ap} - T_{i, ap})^2}{2\sigma_{ap}^2} \right)
\]

and the \((x, y, z)\) coordinates of the mobile are estimated as 
those corresponding to the grid point \(i\) which maximizes \(p_{i}\) 
over the whole search space. The search space is defined as a parallelepiped with edges 
between some minimum and maximum values in each 
direction (i.e., \(x, y, \) and \(z\), respectively). The \(x\)-\(y\) minimum and 
maximum edges are based on the minima and maxima of the 
measured coordinate from the training phase in each building 
plus a small margin (here \(+/-10\) m) to account for the 
variations over time, mobile orientation, and presence of body 
losses. The \(z\) edges are based on the minimum and maximum 
floor heights. In the Bayesian log-likelihood combining [2][7], the evidence 
is combined via the product of individual (linear) probabilities 
or sum of log likelihoods:
Again, the mobile location is taken at the grid point \( i \) which maximizes \( \log p_i \). The log-likelihood combining is the most widely used combining method in WiFi-based positioning via fingerprinting [6][10]. The cost function, \( J_i \), to be maximized (after the fingerprint index \( i \)) in the Bayesian approaches is

$$ J_i = \begin{cases} \log(p_i), & \text{if linear Bayesian combining} \\ \log(p_i), & \text{if log Bayesain combining} \end{cases} $$

The mobile estimated \( \hat{\xi} \) coordinate is found in the \( i \)-th grid point which maximizes the cost function, or as an average of the positions of \( N_{nn} \) ‘nearest neighbors’, which have the \( N_{nn} \) highest values of the cost function:

$$ \hat{\xi} = \frac{1}{N_{nn}} \sum_{i \in \Omega} \xi_i, \quad \Omega = \{ i | J_i \geq J_k, k \notin \Omega \} \tag{5} $$

III. DEMPSTER-SHAFER MODELING WITH BINARY FRAME OF DISCERNMENT

In DS theory, the decision space is divided into mutually exclusive propositions or the so-called DS states, and all the DS states form a frame of discernment. The frame of discernment \( \Theta \) in DS modelling consists of all hypotheses for which the information sources can provide evidence. This set \( \Theta \) is finite and it consists of mutually exclusive propositions that span the hypotheses or decision space. If we denote by \( A \) the \( i \)-th DS state (see Section III.A for examples of choosing \( A \)), then \( \Theta = \{ A_i \}_{i=1,...,N} \), with \( N \) being the number of mutually exclusive states. If \( N = 2 \), then we have a binary frame of discernment. The power set is the set \( 2^\Theta \) of all possible combinations within the frame of discernment, including the empty set \( \emptyset \). For example, for a binary frame of discernment, the power set has 4 elements:

$$ 2^\Theta = \{ A, A_1, (A, A_1) \} \triangleq \{ B \}_{i=1,...,4}, $$

for a tertiary frame of discernment, the power set has 8 elements:

$$ 2^\Theta = \{ A, A_1, A_2, (A, A_1, A_2), (A, A_1), (A, A_2), (A_1, A_2) \} \triangleq \{ B \}_{i=1,...,8}, $$

etc. The power set basically includes evidences that support sub-sets of the frame of the discernment. Each element inside a power set is assigned a positive and sub-unitary mass \( m_i(B) \in [0;1] \), corresponding to the \( s \)-th source of evidence. These masses carry in fact the information about the degree of belief in certain evidence from the power set and they have to fulfill the following conditions:

$$ \sum_{B \in 2^\Theta} m_i(B) = 1, m_i(\emptyset) = 0, 0 \leq m_i(B) \leq 1 \tag{6} $$

In order to combine the evidence coming from two sources and supporting a conclusion \( C \), Dempster and Shafer suggested to following combination rule [1][4]:

$$ m(C) = \frac{1}{\sum_{A \in 2^\Theta} \sum_{B \in 2^\Theta} m_i(A) m_j(B) - m_i(A) m_j(B)} \tag{7} $$

where \( m_i(A) \) is the mass associated to source 1 and supporting the conclusion \( A_i \), \( m_j(A) \) is the mass associated to source 2 and supporting the conclusion \( A_j \), and \( A_i \) and \( A_j \) both support the conclusion \( C \). For combining more than two sources of evidence, an iterative process can be used. DS theory seems therefore well suited to WiFi localization problem because it does not require prior probabilities and it allows for some ‘nebulous’ or uncertain states, for example \( \{ A_i, A_j \} \), which means that any of the states \( A_i \) and \( A_j \) may be true. DS theory requires that the mass functions are to be assigned in a meaningful way to various sources of evidence and this choice of the mass functions is difficult and a crucial step in DS analysis.

The DS mass choice in the context of WiFi fingerprinting is not well documented in the existing literature. In the next subsections, we will detail the investigated mass choices.

A. Dempster-Shafer model for WiFi-based positioning

In WiFi localization, the most straightforward approach is to divide the hypotheses space into two mutually exclusive hypotheses: either the mobile is at a certain location (\( I \) state) or it is not at that location (\( N \) state). This means that the frame of discernment is:

$$ \Theta = \{ I, N \} \tag{8} $$

More complex models, such as a model where each state corresponds to the mobile location at a certain floor \( f: \Theta = \{ I_f, N_f \}_{f=1,...,N_{floor}} \) can also be envisaged, but they are out the scope of our paper. In our case, the power set becomes:\n
$$ 2^\Theta = \{ I, N, (I, N), \emptyset \}. $$

The state \( U = \{ I, N \} \) is the uncertain or \( U \) state, which tells us that we might not have sufficient evidence to decide whether the mobile is or not at a certain location.

B. Dempster-Shafer masses

If, based on the evidence coming from the \( ap \)-th AP, we divide first the hypotheses space into two regions: the certain region (where the mobile is either in or not in a certain grid point) and the uncertain region (where we cannot say whether the mobile is or not in a certain grid point), then we can allocate an uncertainty factor \( u_{ap} \) to the mass of the uncertain \( U \) state in the \( i \)-th grid point. It follows that the mass of the ‘certain’ state, that is the sum of \( I \) and \( N \) states, will be \( 1-u_{ap} \). The distinction between \( I \) and \( N \) states can be further done based on the probability \( p_{i,ap} \) of the mobile to be at a certain grid point based on the evidence of the \( ap \)-th
AP (equation 1). The $p_{i,ap}$ probability has been defined in eq. (1).

The DS masses can be thus defined as follows:

$$
\begin{align*}
    m_{i,ap}(I) &= (1 - u_{i,ap}) p_{i,ap} \\
    m_{i,ap}(N) &= (1 - u_{i,ap})(1 - p_{i,ap}) \\
    m_{i,ap}(U) &= u_{i,ap} \\
    m(\emptyset) &= 0
\end{align*}
$$

Clearly, if $u_{i,ap}$ is chosen between 0 and 1, then the masses above are positive and less than 1, because the probabilities $p_{i,ap}$ given by (1) are sub-unitary. With the above definition, we also have that the condition (8) on masses is satisfied:

$$
\sum_{\varnothing \subseteq Z} m_{i,ap}(Z) = 1
$$

The remaining major issue is how to choose the uncertainty factor $u_{i,ap}$, a fact that will be discussed in Section III.C.

C. Dempster-Shafer uncertainty factor choices

The data sources of the pieces of independence (in our case the APs) can be either independent of each other or correlated with each other. In the first case, the certainty associated with a piece of evidence should be dependent only on the source of that evidence, not on other pieces of evidence. The AP evidence is likely to be a combination of correlated and uncorrelated information.

For example, some APs can transmit from exactly the same location (the cases with multiple BSSID support in WiFi), and thus the channel fluctuations experienced between the transmitter and the mobile will likely be correlated in this case. However, if the APs are distant from one another, the RSSs are likely to be uncorrelated. These assumptions about the correlated profiles when we have same positions APs and uncorrelated profiles when we have different positions APs have also been verified on the measurement data. A snapshot with the power maps of four APs, two correlated (upper part) and two uncorrelated (lower part), is shown in Fig. 2 together with the 2D correlation coefficients, namely 0.97 for the correlated case and 0.03 for the uncorrelated case. The choice of the uncertainty factor in [5] and [12] is based on the assumption that all APs are correlated between them, thus the uncertainty of one AP also depends on the evidence coming from other APs.

In [5], the authors considered that the uncertainty has to be proportional to the fraction of the $ap$-th AP heard by the mobile, reasoning that “the stronger the RSS is, the bigger belief we give to the evidence” and interpreting the U state as the “uncertainty about the evidence”:

$$
u_{i,ap}^{unc} = \frac{1}{N_i} \sum_{ap=1}^{N_i} \frac{1}{10^{h_{ON}}} \text{ if } N_i = 1
$$

$$
= \frac{1}{N_i} \text{ if } N_i > 1
$$

The criticism to the above choice of the uncertainty factor is that it decreases the masses associated to I state when RSS increases, which is counter-intuitive. In [12] the uncertainty factor was defined by associating a lower uncertainty to stronger heard APs, such that a stronger heard RSS will point out towards a higher mass for the I state:

$$
u_{i,ap}^{unc} = 1 - \frac{h_{old}}{10^{h_{old}}} \text{ if } N_i = 1
$$

$$
= \frac{1}{N_i} \text{ if } N_i > 1
$$

The main problems with the choice of the uncertainty factors in [5],[12] (i.e., eqs. (10) and (11)) are that the masses associated to a certain source of evidence are also dependent on the evidence coming from the other sources, thus they implicitly assume that all APs are correlated. However, as illustrated in Fig. 2 and tested with the available measurement data, the power maps corresponding to APs at different locations have very low correlations between them, thus the assumption of fully correlated APs is not a valid assumption.

Moreover, the uncertainties in (10) and (11) may incorrectly give un-balanced weights to the different heard APs based on the instantaneous RSS values which are highly affected by shadowing and measurement mapping, as shown in the histograms of Fig. 1. It follows that the uncertainty factors per AP should be independent on the evidence coming from the other APs. At the same time, the uncertainty factor should be positive and sub-unitary, and it has to be defined solely based on the information arriving at the mobile from a single source, i.e., the MAC address, the instantaneous RSS from the selected AP, and, if known, the variance of the RSS residual values per AP. With this reasoning in mind and using the
prior information we have from the training phase about the standard deviation $\sigma_{ap}$ of the shadowing and measurement noises for each AP and about the mean residual value $\mu_{ap}$ of the mapped RSS into a single grid point, we also define the uncertainty factor of (13):

$$\mu_{map, ap} = \frac{1}{2\sigma_{ap}} \exp \left( - \frac{\mu_{ap}}{2\sigma_{ap}} \right) = \frac{1}{\sqrt{2\pi}\sigma_{ap}}$$

(13)

As the investigations of the measurement data showed that $\mu_{ap}$ values were very small for all the buildings (of the order of $10^{-6}$ or below), we can use the approximation of the right-hand side of (13) for the uncertainties, based only on the standard deviation of the shadowing and measurement noises.

D. Dempster-Shafer decision process

In order to find the mobile location, the cost function to be maximized in DS data fusion is equal to the belief function $J_i = m_i(I)$ defined in (13). Each new AP will contribute to the joint masses via the following iterative process [2]:

For $ap = 2, ..., N_p$

$$J_i - m_i(I) = m_{\{m_u\}}(N_u) + m_{\{m_u\}}(N_u) + m_{\{m_u\}}(I)$$

$$J_i - m_i(I) = m_{\{m_u\}}(N_u) + m_{\{m_u\}}(N_u) + m_{\{m_u\}}(I)$$

$$J_i - m_i(I) = m_{\{m_u\}}(N_u) + m_{\{m_u\}}(N_u) + m_{\{m_u\}}(I)$$

$$J_i - m_i(I) = m_{\{m_u\}}(N_u) + m_{\{m_u\}}(N_u) + m_{\{m_u\}}(I)$$

$$J_i - m_i(I) = m_{\{m_u\}}(N_u) + m_{\{m_u\}}(N_u) + m_{\{m_u\}}(I)$$

(14)

The steps of the DS-based fingerprinting are thus:

1. Obtain the fingerprint database on the mobile from the LSP or from user measurement database.
2. Define the search space as a parallelepiped containing the whole building (see the discussion in Section 2).
3. Define the probability based on eq. (1), with $\sigma_{ap}$ equal to an estimated shadowing variance (measurement plus shadowing parts) value based on the training data, see e.g., [3] and our examples in Fig. 1.
4. Define the uncertainty factors based on eqs. (10), (11), (12), or (13) according to the used algorithm.
5. Define the masses based on eq. (8), with uncertainty factors from Step 5.
6. Combine the evidence coming from the heard AP, according to the iterative process of eq. (14).
7. Find the position which maximizes the belief function $J_i = m_i(I)$, with $m_i(I)$ given in eq. (14) at the end of all iterations.
8. Optionally, also an average over the $N_{\text{NN}}$ nearest neighbors. The nearest neighbors are the first $N_{\text{NN}}$ which maximize the cost function from Step 7.

IV. MEASUREMENT-BASED RESULTS

The data gathering was done manually either with a Windows tablet (Acer Iconia Tab W500 tablet PC with Windows 7 OS) or with a Nexus tablet (Asus Nexus 7 with Android 4.3.1 OS).

Both tablets included the detailed building maps, based on HERE Maps. The type of device used in each building is shown in Table 1. The field data was gathered in 9 different buildings in Tampere (FIN) and Berlin (DE), for three types of buildings: university (Uni), shopping centers (Center) and office buildings (Office). The position inside the building was chosen manually, based on the maps at the measured floors. We expect horizontal errors below 0.5 m in the data collection stage (and no errors in the vertical direction). The number of floors ($N_f$), the number of collected initially collected fingerprints ($N_{\text{fp}}$), the number of heard APs ($N_{\text{AP}}$), and the average standard deviation of measurement and shadowing errors are shown in Table IV-1.

<table>
<thead>
<tr>
<th>Building</th>
<th>Mean. device</th>
<th>$N_f$</th>
<th>$N_{\text{fp}}$</th>
<th>Average meas. std</th>
<th>Average shadowing std</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uni 1, FIN</td>
<td>Windows</td>
<td>6</td>
<td>1591</td>
<td>309</td>
<td>4.08</td>
</tr>
<tr>
<td>Uni 2, FIN</td>
<td>Nexus</td>
<td>4</td>
<td>607</td>
<td>393</td>
<td>3.17</td>
</tr>
<tr>
<td>Center 1, FIN</td>
<td>Windows</td>
<td>3</td>
<td>594</td>
<td>354</td>
<td>3.29</td>
</tr>
<tr>
<td>Center 2, FIN</td>
<td>Nexus</td>
<td>6</td>
<td>282</td>
<td>69</td>
<td>2.47</td>
</tr>
<tr>
<td>Office 1, FIN</td>
<td>Nexus</td>
<td>7</td>
<td>859</td>
<td>995</td>
<td>3.49</td>
</tr>
<tr>
<td>Office 2, FIN</td>
<td>Nexus</td>
<td>4</td>
<td>2239</td>
<td>162</td>
<td>3.03</td>
</tr>
<tr>
<td>Office 3, FIN</td>
<td>Nexus</td>
<td>10</td>
<td>24045</td>
<td>727</td>
<td>6.96</td>
</tr>
<tr>
<td>Center 5, DE</td>
<td>Nexus</td>
<td>4</td>
<td>2552</td>
<td>631</td>
<td>2.93</td>
</tr>
<tr>
<td>Center 6, DE</td>
<td>Nexus</td>
<td>7</td>
<td>16298</td>
<td>878</td>
<td>2.95</td>
</tr>
</tbody>
</table>

Table IV-1. Characteristics of the measured buildings

One AP refers to one MAC address. It is possible (and very likely) that several MAC addresses are coming from the same physical location of a WiFi emitter (e.g., multiple BSSID), thus some measurements will be highly correlated between them, coming from the same physical channel. In our data analysis, we employed all the available measurements.

The measurement based results are shown in Table IV-2. The best values among all tested approaches are shown in bold-faced letters. Clearly, there are no significant differences between the Bayesian and three of the DS approaches. The DS approaches with masses given in eqs. (10-12) are offering better floor detection probability than the Bayesian approaches in the majority of cases, but the average improvement is below 2%. The newly proposed DS weights are slightly better on average than the old DS approach of [12] and much better than the old DS approach of [5]. The mass choice of [5] is clearly sub-optimal, and this case is the only one which gives consistently worse results than the others. This points out to the fact that the mass choice in DS approaches is a crucial step for a good functioning of the algorithm.

<table>
<thead>
<tr>
<th>Building</th>
<th>Perf. criter.</th>
<th>Bayesian, lin, eq. (7)</th>
<th>Bayesian, log, eq. (7)</th>
<th>Old DS masses</th>
<th>New ct DS masses</th>
<th>New var DS masses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uni 1, FIN</td>
<td>Mean pos. error [m]</td>
<td>3.4</td>
<td>6.0</td>
<td>20.5</td>
<td>8.9</td>
<td>5.3</td>
</tr>
<tr>
<td>Uni 1, FIN</td>
<td>$\sigma_{ap}$</td>
<td>92.45</td>
<td>86.73</td>
<td>45.92</td>
<td>88.36</td>
<td>95.71</td>
</tr>
<tr>
<td>Office 1, FIN</td>
<td>Mean pos. error [m]</td>
<td>7.4</td>
<td>7.3</td>
<td>20.3</td>
<td>9.0</td>
<td>7.3</td>
</tr>
</tbody>
</table>

Table IV-2. Mean 3D positioning distance error [m] and the floor detection probability $P_d$ [%].
Our analysis thus shows the following important points regarding DS analysis:

- **DS analysis** does not always incorporate the uncertainty due to measurements and shadowing errors.
- The choice of masses in **DS analysis** is crucial, as it can influence the results significantly.
- The benefit of **DS approaches** over Bayesian approaches, in terms of position error, is mostly seen in buildings without many open spaces between floors (e.g., typical office buildings).
- The floor detection probabilities of all studied approaches in Table IV-2, except for **DS algorithms** of [12], is very good (typically close to 85-90%), even in multi-floor buildings with open-spaces, such as the shopping centres.

**V. CONCLUSIONS**

In this paper, we have investigated the potential of DS data fusion on the context of WiFi indoor positioning via fingerprinting and we compared various data fusion approaches, based on Bayesian and non-Bayesian DS theory. The DS framework relaxes the Bayesian assumption of mutually exclusive hypotheses, allowing thus for more flexibility through the introduction of an uncertain state. We have introduced two new ways of defining the DS masses, based on variable and constant uncertainty factors and prior information collected from the training data, in order to fit better modeling the data quality and into quantifying the data conflict measures based on DS joint masses.

**ACKNOWLEDGEMENTS**

The authors express their warm thanks to the Academy of Finland (project 250266) for its financial support and to HERE for providing the measurement equipment. Parts of our measurement data for indoor localization is available at [http://www.cs.tut.fi/tlt/pos/Software.htm](http://www.cs.tut.fi/tlt/pos/Software.htm).

**REFERENCES**


