# Simultaneous Counting and Location of Persons Based on a Heterogeneous Sensor Setup

Gernot Ruscher and Thomas Kirste

Rostock University
Albert-Einstein-Straße 22, 18059 Rostock, Germany
firstname.lastname@uni-rostock.de

**Abstract.** Most passive location systems (e. g. pressure sensitive floor mats) inherently do not come with an identification mechanism, i. e. are ambiguous concerning the number of persons who cause readings: Two persons close beside each other could seem like one single person. In contrast, most active location systems operate with user-bound tags, and thus directly give the number of used tags. Possibly, both sensor types may be used in conjunction, but the number of persons and their respective locations need then to be simultaneously estimated.
This paper presents an approach that allows to fuse observations from both sensor types. We introduce a probabilistic model for representing the joint density of number and location of persons as well as the available sensor modalities. We show how this model can be used for estimating the posterior placement density given the sensor observations, applying a Gibbs sampling approach.

**Keywords:** probabilistic models, location estimation, sensor systems

## 1 Introduction and Motivation

In this paper, we analyse the chances and limits of a novel sensor fusion technique employing the exact probabilistic modelling of (i) two location sensing modalities, one being able to differentiate between persons and the other one with no such *labelling* capability, and (ii) the joint density of the number and respective locations of persons. We further model the causal interdependencies of these items and derive a sampling strategy in order to simultaneously estimate (ii) based on observations from (i). We present an initial evaluation and outline further research topics we identified based on this work. But prior to that, we give a brief introduction into background and motivation of this work, and depict potential application fields below.

Smart Environments [5, 6] are indoor environments intended to give assistance to people who perform specific everyday activities, with the latter being mostly predetermined by the properties and the equipment of the environment. The desired assistance is then given in form of automatic responses to the people's actions. The underlying techniques leading to such *proactive* assistance functionalities can conceptionally be divided into two main subprocesses: First, *Intention Analysis* determines the current *user goal* (probabilities). This happens based on (i) a *user* or *activity model* and (ii) on a *observation* or *sensor model*. By then, *Strategy Synthesis* can take place, which is supposed find a reasonable plan of actions for the (ad-hoc) device ensemble. Since Intention Analysis relies on robust and accurate sensor data, and location information is a fundamental subset of sensor data, this work aims at enhancing the quality and accuracy of existing location systems by sensor data fusion.

## 2 Technical Background

A great number of location sensing systems is available. The probably simplest ones are little expressive but simple binary sensors such as Passive Infrared sensors, light barriers, and pressure mats, which are merely able to detect the presence of an object which is heavy or big or warm enough to trigger the sensor. Passive location systems are inherently not able to distinguish between different objects, and thus do not provide any identification or labelling mechanism.

A rather sophisticated passive location system is given by the *SensFloor* system, which is basically a flooring of sensitive tiles, with the relatively high resolution of 32 sensor tiles per square metre. The manufacturers' description [12] reads as follows: "A grid of sensors underneath the flooring [that] detects local capacitive changes in the environment brought about by humans walking on the floor. By design, this method does not allow for an identification of individuals." The latter comes with the advantage of no requirement of additional technical devices on the part of the persons to be tracked. The sensor tile arrangement can be seen in Fig. 1. Every floor tile is able to detect if its area is being covered by an object. It will give signal *true* if an object is present and signal *false* otherwise. Through the built-in recalibration mechanism, sensor tiles are able to adjust themselves to fixed furniture parts as chairs and tables.

Within the *SmartFloor* project [9] pressure sensitive floor plates were utilised for person identification and tracking, based on information about the load, exerted by persons, which is quite more expressive data than e. g. the binary SensFloor signals. Similar projects or products utilising floor tile based sensors or sensor data are: *SmartCarpet* [11], *TileTrack* [14], and *Z-Tiles* [10]. For a good overview of location tracking approaches, [7] may be consulted.

By contrast, *active* location systems require additional technical devices on the part of the users which are to be located or tracked. Thus, they allow for a unique assignment of sensor data to persons. Most of them are highly specialised solutions based on UWB pulses such as the *Ubisense RTLS* [4] and the *XSens Motion Grid* [2, 8], or on ultra-sound pulses such as *Sonitor* [1]. Another example is a fixed RFID scanner which produces a reading everytime an RFID tag passes at short distance, with the tag making the passing person clearly identifiable. By all active systems known to the authors, tags are handled system-unique. Thus, they allow inherently to keep track of the tagged persons, e. g. even in situations when two or more objects pass each other at short distance, which is a characteristic problem during tracking tasks. To the contrary, passive systems

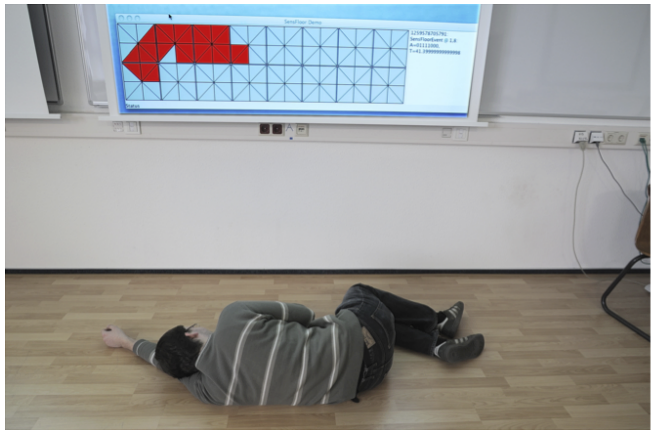

**Fig. 1.** *SensFloor* sample setup and sensor data, red areas indicate activated cells

cannot discriminate or identify persons, for what reason this tracking problem can become an intractable task to them, and the exclusive usage of passive systems would quickly suffer from this technical limitation. Taken this together with the quite strong assumption to have every user carry a tag, a combination of these two kinds of location systems, so that this assumption can be weakened, seems to make sense.

The authors of [13] have established a location system taxonomy which distinguishes between intrinsic and extrinsic traits, with the latter being traits that are not held by a humans inherently. They need to be "lent" by some technical device resembling the above-mentioned *tag*. But note, that the disambiguation among passive and active systems we require in the context of this paper differs a little from [13]: In our point of view, any sensor system that comprises some object ID mechanism and thus gives us a unique assignment of a location measurement to a specific person and hence the opportunity to keep track of people inherently, would be called *active* system, regardless of whether a technical device was involved. To give an example, all kinds of camera-based approaches making use of face recognition (with the face being an intrinsic trait) to identify people and hence being able to keep track of them, would need to be stated as *active* in this context.

For our experiments we utilised our SmartLab: An instrumented meeting room (cf. Fig. 4 of section 3) equipped with a number of remotely controllable devices. Besides for experiments, the SmartLab is frequently used as a common seminar room for lectures, presentations, meetings, and the like. Our laboratory is equipped with a couple of sensors, required to observe state changes in the room. Besides sensors capable of detecting whether the windows are closed or opened, or measuring the current temperature, or detecting persons that enter or leave the room, we have got a SensFloor installation covering a specific part of the room area as well as a Ubisense RTLS system running. Besides the sensors there is a number of remotely controllable devices which are required in typical meeting rooms: Dimmable lamps as well as movable projection screens and sun shades, controllable via EIB, a computer video and audio matrix switcher to connect brought-in devices with the installed projectors and an audio equipment, just to mention the most important. Furthermore, our SmartLab also features a powerful middleware [3] which on the one hand allows for ad-hoc control of all currently present hardware by using simple commands. On the other hand, apart from triggering device actions, our middleware enables every device to make its specific properties accessible to other components in the system, for instance through a tuple space. Thus, every component has the opportunity to watch the entire world state at a glance whenever needed.

# 3 Our Approach

## 3.1 Situation and Sensor Abstraction

We consider the area of our SmartLab as a finite set of locations $\mathcal{L}$, e. g. the cells of a grid. With $a(o, l)$ being the area covered by an object $o$ at location $l \in \mathcal{L}$, $c(a_1, a_2)$ describes the coverage of an area $a_1$ by another area $a_2$ as

$$c(a_1, a_2) \triangleq \frac{\sigma(a_1 \cap a_2)}{\sigma(a_1)}$$

with $\sigma(S)$ denoting the size of area $S$, e. g. the number of covered cells.

The measurement area of our *SensFloor* installation (cf. section 2) can be considered as a set of $n$ floor tiles $\tau_{1:n}$, with every $\tau_i$ covering a specific area $a_{\tau_i}$. Every such sensitive floor tile can be seen as a random variable $Z_i$, giving values of either 1 or 0, while a reading of 1 indicates an object covering that tile, and a reading of 0 means no coverage. In order to take erratic sensor data into account, we will introduce $e_{1|0}$ as the probability of getting a 1-reading even if the tile area isn't covered by any object, and $e_{0|1}$ denoting the probability of reading 0 even if the floor tile is entirely covered.

As an active location system based on tags we have presented *Ubisense RTLS*. We assume every tag $\gamma_i$ with $i \in [1 : m]$ to be fixed to the person $\omega_i$. Each of those tags $\gamma_i$ can be seen as a random variable $Y_i$ giving values $l_i \in \mathcal{L}$. The probability of receiving the sensor reading $y_i$, given object $\omega_i$ at location $l_i$ can be described e. g. as $p(y_i|l_i) = \mathcal{N}(y_i|\mu = l_i, \Sigma)$, a normal distribution centered at $l_i$ with $\Sigma$ denoting the measurement error of this sensor reading.

## 3.2 Probabilistic Model and Inference Strategy

Further random variables $X_i$ ($i \in [1 : m]$) bring the above random variables $Y_i$ and $Z_j$ into relation with the persons which are to be located. The structure of the random variables together with their conditional interdependencies are given by the graphical probabilistic model in Fig. 2 and defined as follows.

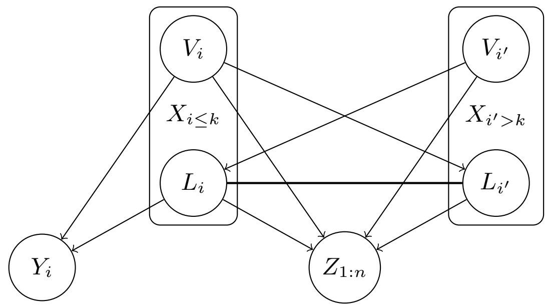


**Fig. 2.** Graphical model of the probabilistic structure governing our SmartLab.

We assume that locations of people cannot overlap. Thus, a maximum of $m$ persons $X_{1:m}$ are able to be present, with $m$ depending on the size of the room as well as its equipment with desks and other furniture elements and also on the spatial needs of each person. We assume that $k$ persons ($k \leq m$) are wearing tags. We call the $m - k$ persons who are not wearing a tag *anonymous*. We can hence assume a minimum of $k$ and a maximum of $m$ persons present in our room, and the following question was raised: How many anonymous persons are indeed present in our room and what are the most probable locations of all of them?

Our graphical model Fig. 2 describes every person $X_{1:m}$ as a joint conditional density, consisting of two random variables: *Visibility* $V_i$ and *Location* $L_i$. Thereby, $V_i$ is a flag that reads $V_i = 1$ if $X_i$ is visible and $V_i = 0$ otherwise, and was introduced for handling different numbers of persons: Only if a person is visible it can cause sensor data. The location variable $L_i$ resembling the person's whereabout delivers elements out of the set of possible locations $\mathcal{L}$.

Because each person $X_i$ with $i \leq k$ is wearing a tag it has got a direct impact onto the sensor data of this tag $Y_i$. Solely visible persons can be the cause for sensor data at all. The sensor tiles $\gamma_j$ are fixed to their locations within the room and can possibly be activated by every user within range, but again only if this user is set visible. This explains why every person's visibility $V_i$ is linked to its own tag variable $Y_i$ and to all of the sensor tile variables $Z_{1:n}$.

Furthermore, whenever $V_i = 1$ this means that person $X_i$ will occupy some free space which then cannot be used by other persons. That's why every $V_i$ is linked to every $L_{i'\neq i}$ . Finally note the bidirectional correlation between all of the $L$ variables. This is because user $X_i$ cannot share location with user $X_{i'}$ if $V_i = V_{i'} = 1$. Therefore, if we have several location variables $L_{1:m}$ all of them are correlated due to the assumption of non-overlapping personal occupation areas.

Based on this model, our goal is to compute visibility and location of all persons $X_{i:m}$ simultaneously, given $k$ tag readings and $n$ sensor tile readings:

$$p(x_{1:m}|y_{1:k}, z_{1:n}) = p(v_{1:m}, l_{1:m}|y_{1:k}, z_{1:n})$$

Because of the mutual interdependencies of the $L_i$ variables, this posterior density is possibly not manageable at all the analytical way, even when using grid-based approaches. Thus, we needed to resort to sampling methods. For this actual case, Gibbs sampling promised to be an appropriate inference method by sequentially checking if (i) making a single person $X_i$ visible and (ii) placing it at some available space would bring out a better explanation of the observed sensor data. Based on this, we can sample $v_i$ and $l_i$ proportional to the possible visibility and location values. The structure of a simplified Gibbs sampling process is shown exemplarily in Fig. 3.

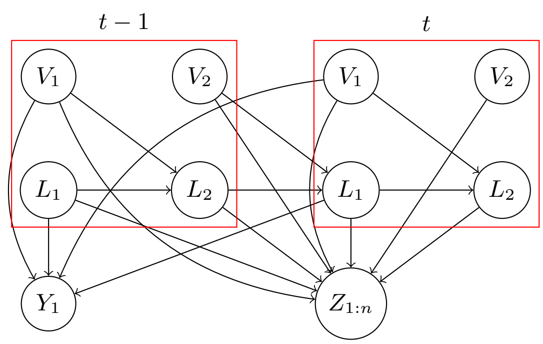


**Fig. 3.** Expanded Gibbs sampling structure of a simplified model considering $m = 2$ persons with $k = 1$ of them wearing a tag. Sample duration: 2 time slices. **Note:** $t-1$ and $t$ are sample indices, *not* timesteps in terms of a clock time.

Given a sample $x_{1:m}^{(t-1)}$, Gibbs sampling creates for every $i \in [1 : m]$ a new sample $x_{1:m}^{(t)}$ by sequentially drawing

$$x_i^{(t)} \sim p(x_i^{(t)} \mid x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n})$$

Since $x_i^{(t)} = (l_i^{(t)}, v_i^{(t)})$, we have

$$\begin{aligned} & p(l_i^{(t)}, v_i^{(t)} \mid x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n}) \\ & \quad = p(l_i^{(t)} \mid x_{1:i-1}^{(t)}, v_i^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n}) p(v_i^{(t)} \mid x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n}) \end{aligned}$$

and we can draw

$$v_i^{(t)} \sim p(v_i^{(t)} \mid x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n}) \tag{1}$$

$$l_i^{(t)} \sim p(l_i^{(t)} \mid x_{1:i-1}^{(t)}, v_i^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n}) \tag{2}$$

We end up having two densities, which we want to sequentially draw samples from. The following two sections cover their derivation.

### 3.3 Sampling a Person's Visibility

By marginalisation and the Bayes' Theorem:

$$
\begin{aligned}
&p(v_i^{(t)} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n}) \\
&= \int \frac{p(y_{1:k}, z_{1:n} \,|\, v_i^{(t)}, l_i^{(t)}, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}) p(v_i^{(t)}, l_i^{(t)} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)})}{p(y_{1:k}, z_{1:n} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)})} dl_i^{(t)}
\end{aligned}
$$

By defining $\alpha \triangleq 1/p(y_{1:k}, z_{1:n} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)})$:

$$
\begin{aligned}
&= \alpha \int p(y_{1:k}, z_{1:n} \,|\, x_{1:i}^{(t)}, x_{i+1:m}^{(t-1)}) p(v_i^{(t)} \,|\, l_i^{(t)}, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}) \\
&\qquad \times p(l_i^{(t)} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}) dl_i^{(t)}
\end{aligned}
$$

Concluding from Fig. 3 that $p(v_i^{(t)} \,|\, l_i^{(t)}, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}) = p(v_i^{(t)})$, and assuming $p(v_i^{(t)})$ to be a constant, by $\alpha' = \alpha p(v_i^{(t)})$:

$$
= \alpha' \int p(y_{1:k}, z_{1:n} \,|\, x_{1:i}^{(t)}, x_{i+1:m}^{(t-1)}) p(l_i^{(t)} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}) dl_i^{(t)} \tag{3}
$$

$$
\begin{aligned}
&= \alpha' \int p(y_i \,|\, x_i^{(t)}) p(y_{j \neq i} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}) p(z_{1:n} \,|\, x_{1:i}^{(t)}, x_{i+1:m}^{(t-1)}) \\
&\qquad \times p(l_i^{(t)} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}) dl_i^{(t)}
\end{aligned}
$$

By $\alpha'' = \alpha' p(y_{j \neq i} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)})$, since $p(y_{j \neq i} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)})$ is a constant:

$$
= \alpha'' \underbrace{\int \overbrace{p(y_i \,|\, l_i^{(t)}, v_i^{(t)}) p(z_{1:n} \,|\, x_{1:i}^{(t)}, x_{i+1:m}^{(t-1)}) p(l_i^{(t)} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)})}^{=:g(l_i^{(t)}, v_i^{(t)})} dl_i^{(t)}}_{=:f(v_i^{(t)})} \tag{4}
$$

To determine $\alpha''$, we note that $\sum_{v_i^{(t)} \in \{0,1\}} p(v_i^{(t)} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n}) = 1$. So $\alpha'' f(v_i^{(t)} = 1) + \alpha'' f(v_i^{(t)} = 0) = 1$ which gives $\alpha'' = 1/\big(f(1) + f(0)\big)$ and

$$
p(v_i^{(t)} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n}) = \frac{f(v_i^{(t)})}{f(1) + f(0)}
$$

We can hence sample $v_i^{(t)}$ by (1) computing $f(1)$ and $f(0)$, (2) drawing $u \sim \mathcal{U}(0,1)$, and (3) setting $v_i^{(t)} := 1$, if $u > \frac{f(0)}{f(1)+f(0)}$ and $v_i^{(t)} := 0$, otherwise.

With the following three paragraphs, we outline the main ideas and the derivations of the three factors of $g(l_i^{(t)}, v_i^{(t)})$ (term 4).

$\boldsymbol{p(y_i \,|\, l_i^{(t)}, v_i^{(t)})}$ This factor describes the sensor reading probability of tag $y_i$ given both visibility and location samples of person $\omega_i$ at sampling step $t$. For simplicity reasons, we assumed that anonymous persons wear *mock* tags, which always return 1:

$$p(y_i \,|\, l_i^{(t)}, v_i^{(t)}) = \begin{cases} 1 & \text{if } i > k \\ \mathcal{N}(y_i | \mu = l_i^{(t)}, \Sigma) & \text{otherwise} \end{cases}$$

Note that for tagged users $p(y_i \,|\, l_i^{(t)}, v_i^{(t)}) = 0$ if $v_i^{(t)} = 0$. That is why $v_i^{(t)} = 0$ is impossible, since in this case $f(v_i^{(t)}) = f(0) = 0$ so that $f(1)/\big(f(1) + f(0)\big) = 1$. Thus $p(V_i^{(t)} = 1) = 1$.

$\boldsymbol{p(z_{1:n} \,|\, x_{1:i}^{(t)}, x_{i+1:m}^{(t-1)})}$ The probability of the *SensFloor* readings $z_{1:n}$ given both the visibility and location of every person, can be simplified due to the independence of sensor tile readings given the causes: $p(z_{1:n} \,|\, x_{1:i}^{(t)}, x_{i+1:m}^{(t-1)}) = \prod_{j=1}^{n} p(z_j \,|\, x_{1:i}^{(t)}, x_{i+1:m}^{(t-1)})$. We assume for each tile a causal connection between its *coverage* $c$ and the probability $p(Z = 1 \,|\, c)$ of getting a 1-reading from that tile. Furthermore, we introduce two error probabilities, $e_{1|0}$ indicating the probability of a 1-reading even if $c = 0$ and $e_{0|1}$ the other way round. Based on this, we defined a linear *SensFloor* activation function: $p(Z = 1 \,|\, c) = e_{1|0} + c * (1 - e_{1|0} - e_{0|1})$. Other shapes like threshold or logistic functions, possibly tile-specific, could also be used if appropriate.

$\boldsymbol{p(l_i^{(t)} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)})}$ The probability of a location $l_i$ at sampling step $t$, given both the visibility and location of all other persons, encompasses several aspects:

- Once we have placed all persons except $\omega_i$, they cover specific parts of the area. The remaining areas can then be treated as probable locations.
- Non-accessible areas like desks and other furniture can be regarded through a predefined weight function which assigns lower weights to those areas.
- Person-specific location preferences (possibly resulting from the person's current activity) can also be covered, yet were initially assumed to be constant.
- The set of possible locations can be restricted to a subset if there are areas where persons cannot be the cause for sensor data at all.

With this, we are able to handle the complexity which arises from the pairwise interdependencies between the persons' locations by

$$p(l_i^{(t)} \,|\, x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}) = \begin{cases} \alpha\, p(l_i^{(t)} \,|\, \omega_i) & \text{if } l_i^{(t)} \in R(\omega_i; x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}) \\ 0 & \text{otherwise} \end{cases}$$

where $R(\omega_i; x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)})$ is a set of available locations (based on the spatial needs of person $\omega_i$ and all the other persons), weighted by a function $\alpha$.

### 3.4 Sampling a Person's Location

After having obtained a sample $v_i^{(t)} \sim p(v_i^{(t)} \mid x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n})$, we are now able to sample

$$\begin{aligned} l_i^{(t)} &\sim p(l_i^{(t)} \mid x_{1:i-1}^{(t)}, v_i^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n}) \\ &\propto p(y_{1:k}, z_{1:n} \mid x_{1:i-1}^{(t)}, v_i^{(t)}, l_i^{(t)}, x_{i+1:m}^{(t-1)}) p(l_i^{(t)} \mid x_{1:i-1}^{(t)}, v_i^{(t)}, x_{i+1:m}^{(t-1)}) \\ &= p(y_{1:k}, z_{1:n} \mid x_{1:i}^{(t)}, x_{i+1:m}^{(t-1)}) p(l_i^{(t)} \mid x_{1:i-1}^{(t)}, x_{i+1:m}^{(t-1)}) \end{aligned}$$

These two factors have already emerged in the previous section within term 3.

For a discrete set of locations $\mathcal{L}$, sampling $p(l_i^{(t)} \mid x_{1:i-1}^{(t)}, v_i^{(t)}, x_{i+1:m}^{(t-1)}, y_{1:k}, z_{1:n})$ amounts to (1) computing $w_i^{(t)}(l) := g(l, v_i^{(t)})$ for every $l \in \mathcal{L}$, and (2) drawing $l_i^{(t)}$ proportional to $w_i^{(t)}(l)$.

## 4 Initial Simulation Studies and Results

The SmartLab situation is depicted in Fig. 4 (right). The area (approx. $7\times 7$ metres in size) was represented by an $18 \times 18$ matrix. Coloured grid cells stand for covered areas while occupation ranges from red (fully covered) via blue (less covered) through to white (not covered). There is a U-shaped desk arrangement and a chair. The SensFloor is represented by 64 grey cells in the left part of the room. Little gray squares indicate activated tiles. The gray circular area represents a tag reading at position $(3, 6)$ and the 95% confidence region for the covariance governing the underlying distribution (here: $\Sigma = \left(\begin{smallmatrix} 0.6 & 0.0 \\ 0.0 & 0.6 \end{smallmatrix}\right)$).

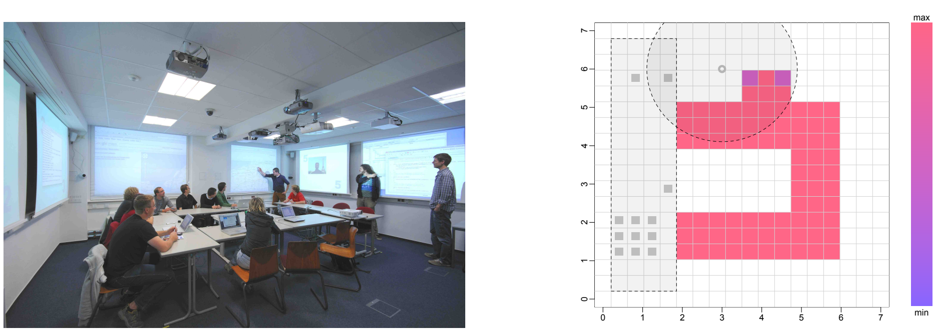


**Fig. 4.** left: our SmartLab, right: initial setting for our simulation studies

As there is just one tag reading, we know there is one tagged person. Following questions were raised: How many untagged persons are there? And what are the potential locations of all the persons?

Two simulation runs with $N = 1000$ drawn samples each have been performed, and a maximum number of $m = 11$ persons has been assumed. The first run (Fig. 5 (left)) employed parameters indicating relatively *unreliable* sensors ($e_{1|0} = 0.1$, $e_{0|1} = 0.1$, $\Sigma = \left(\begin{smallmatrix} .6 & 0 \\ 0 & .6 \end{smallmatrix}\right)$) and a *small* footprint. The second run (more *reliable* sensors $e_{1|0} = 0.01$, $e_{0|1} = 0.001$, $\Sigma = \left(\begin{smallmatrix} .2 & 0 \\ 0 & .2 \end{smallmatrix}\right)$, and *large* footprint).

Figure 5 shows the simulation results using those two different parameter sets. For each run a single sample is depicted upper left to give an idea. Find next to it the distribution of the number of visible people. The lower left picture shows the distribution of all the placements of the single tagged person, and in the lower right figure all possible locations of all anonymous persons are shown. Note that the anonymous persons are permutable, that is why their placements have all been put into a single figure. The first run resulted in a mean of 5.3 persons (sd=0.99, median=5), while the second provided a mean of 3.37 persons (sd=0.61, median=3).

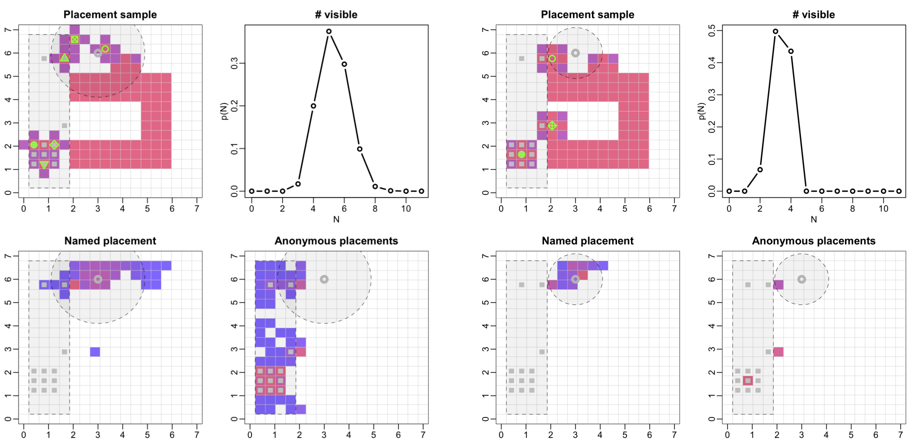


**Fig. 5.** left half: Simulation results using *unreliable* sensors and *small* person footprint
right half: Simulation results using *reliable* sensors and *large* person footprint

## 5 Conclusion and Future Work

The paper presents first a classification scheme for multimodal sensor equipments which distinguishes between active and passive location systems, the first of which operate in a tag-based way and allow unique assignment of measurements to users. Passive location systems without any identification mechanism cannot make any statement about which person has caused an observed sensor reading.

An overview was given of the specific setup of our experimental environment, the SmartLab. It is equipped with various sensor systems two of which are a *SensFloor* system covering an area of 4 square metres and a *Ubisense RTLS* cell. Within this work, the latter has been considered as a representative for active location systems while the SensFloor system has been demonstrated as a sample passive location system.

Given this multimodal sensor setup as well as an application scenario regarding multiple persons being partially equipped with active tags and a set of sensor readings originating from both location systems at a specific point in time, the question was raised: How many persons are inside that environment in total *and* what are their respective locations?

In order to answer the above question, we obtained a unified probabilistic modelling approach, which is fit to take simultaneously into account the number of persons *and* their respective locations with respect to sensor data from both active as well as passive location systems. Furthermore, the presented approach provides several modelling interfaces to deal with the following issues, which either already have been addressed or can readily be tackled:

– Application of detailed *error models* for both kinds of location systems to model their actual *reliability*, possibly even down to the level of individual tags or sensor tiles, when necessary (cf. section 3.3)
– Consideration of *non-accessible areas* due to e. g. walls, desks, chairs, and other furniture
– Distinction of differing personal occupation areas which cannot overlap
– Integration of personal or role-based location preferences possibly capable to regard current user or group activity

Note that the last point inidcates an interface to overlying components such as activity or intention recognition (cf. section 1). With this, certain knowledge about current user activites could possibly enhance location results.

We have shown results of initial simulation studies. They were delivering promising results, though they revealed several open questions: The actual parameters which fit for the existing location systems in our lab still need to be figured out. Once we have obtained the number of anonymous persons, probable locations for each of them need to be estimated. *Bump Hunting* algorithms should be suitable for this task, and we are working on this topic at the moment.

Besides this, within the presented model anonymous persons are only of interest regarding their number, and *not* with respect to specific attributes, what makes them pairwise permutable. We have not yet looked carefully enough at this circumstance and still need to learn if and how we can profit from that.

Note further that we have not elaborated within this work the integration of location development over time, the *motion* of persons. The proposed model and inference method so far are especially of interest when we are *beginning* inference, i. e. when we want to process the very first observation vector. Though we have not yet integrated motion-related issues, the proposed modelling approach can be extended in order to support person tracking tasks. At first sight this extension could comprise the integration of further random variables representing motion speed and direction, and would further require a mirroring of the whole model from Fig. 2 together with the formalisation of the respective relations between random variables of timesteps $t$ and $t+1$. The employment of an appropriate inference method (e. g. (Rao-Blackwellized) Particle Filter) could make the model suitable for online processing.

Another issue which we are currently elaborating is the exact determination of the paramaters of the proposed sampling algorithm: The error models of both presented location systems have not yet been investigated carefully enough. But initial measurements concerning Ubisense RTLS encourage us to carry out more detailed studies: They created doubt if normal distributions fit best to model location errors. In addition to this, *one* error model could be insufficient to describe the error distribution(s). However, the results will be helpful anyway, as we are examining simulation tasks of smart environments as well.

## References


1. Sonitor (FEB 2012), www.sonitor.com
2. XSens MVN MotionGrid (FEB 2012), www.xsens.com/en/general/motiongrid
3. Bader, S., Ruscher, G., Kirste, T.: A middleware for rapid prototyping smart environments: Experiences in research and teaching. In: Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing. pp. 355–356. ACM, Copenhagen, DK (SEP 2010)
4. Cadman, J.: Deploying commercial location-aware systems. In: Proceedings of the 2003 Workshop on Location-Aware Computing. pp. 4–6 (OCT 2003)
5. Chen, G., Kotz, D.: A Survey of Context-Aware Mobile Computing Research. Tech. rep., Hanover, NH, USA (2000)
6. Cook, D., Das, S.: Smart Environments. Wiley (2005)
7. Hightower, J., Berliniello, G.: Location systems for ubiquitous computing. IEEE Computer 34(8), 57–66 (2001)
8. Hol, J., Schön, T., Gustafsson, F.: Ultra-wideband calibration for indoor positioning. In: 2010 IEEE International Conference on Ultra-Wideband (ICUWB). vol. 2, pp. 1–4 (SEP 2010)
9. Orr, R.J., Abowd, G.D.: The smart floor: A mechanism for natural user identification and tracking. pp. 1–6. ACM Press (2000)
10. Richardson, B., Leydon, K., Fernstrom, M., Paradiso, J.A.: Z-tiles: building blocks for modular, pressure-sensing floorspaces. In: CHI '04 extended abstracts on Human factors in computing systems. pp. 1529–1532. CHI EA '04, ACM, New York, NY, USA (2004)
11. Savio, D., Ludwig, T.: Smart carpet: A footstep tracking interface. In: Proceedings of the 21st International Conference on Advanced Information Networking and Applications Workshops - Volume 02. pp. 754–760. AINAW '07, IEEE Computer Society, Washington, DC, USA (2007)
12. Steinhage, A., Lauterbach, C.: SensFloor® and NaviFloor®: Large-area sensor systems beneath your feet. In: Chong, N.Y., Mastrogiovanni, F. (eds.) Handbook of Research on Ambient Intelligence and Smart Environments: Trends and Perspectives, vol. 2, pp. 41–55 (SEP 2011)
13. Teixeira, T., Dublon, G., Savvides, A.: A survey of human sensing: Methods for detecting presence, count, location, track and identity, under review
14. Valtonen, M., Maentausta, J., Vanhala, J.: Tiletrack: Capacitive human tracking using floor tiles. In: Proceedings of the 2009 IEEE International Conference on Pervasive Computing and Communications. pp. 1–10. IEEE Computer Society, Washington, DC, USA (2009)