Multi-Target Data Association in Binary Sensor Networks for Ambulant Care Support

Sebastian Matthias Müller
OFFIS – Institute for Information Technology
Oldenburg, Germany
Email: sebastian.mueller@offis.de

Andreas Hein
School of Medicine and Health Sciences
University of Oldenburg
Oldenburg, Germany
Email: andreas.hein@uni-oldenburg.de

Abstract—**Numerous applications of home automation, security and ambulant medical care use binary sensors such as passive infrared motion sensors or light barriers to monitor activity in the house. While the data of individual sensors does not facilitate the recognition and separation of the presence of more than one person, there exist multi-target tracking algorithms that allow for at least a partial separation of activity in data from multiple persons. While many tracking algorithms demonstrate good performance across various sensing modalities and sensor setups, little research has been done to determine the impact of placement and varying density of sensors for tracking performance. This paper presents the results of two evaluations of a Bayesian multi-hypothesis multi-target tracking algorithm on data of residents monitored by a network of binary sensors. The algorithm's performance is evaluated across varying quantity and placement. It is shown that the approach outperforms other approaches in low-resolution setups using data collected in a home lab and further demonstrate its applicability in a field trial across three households. While tracking performance naturally decreases with the number of sensors, it also strongly varies by sensor positioning.**

Keywords–Multi-target tracking; Assisted living; Wireless sensor networks; home automation; smart homes; telemedicine.

I. Introduction

This article is an extended version of an article originally submitted to the International Conference on Ambient Computing, Applications, Services and Technologies (AMBIENT) 2015 [1].

The emergence of research on technical support systems for ambulant care and support for patients and elderly stem from numerous recent societal developments as well as changes in demographic structure. First, the coincidence of prolonged life expectancy [2], [3] and the atomization of households [4], [5] puts an increasing care demand into the hands of third parties, especially in Western countries with low birth rates: US census data shows that the average household size in the United States declined from 4.5 persons per household in 1900 to 2.6 in 2000 [6]. According to the German Federal Statistical Office, the number of single households will increase sixfold relative to the population between 2010 and 2030. At the same time, the ratio between care personnel supply and demand will cut in half [7]. Second, increasing life expectancy, in combination with improved medical care and "modern lifestyles and behavior" causes an increase in the proportion of population living with chronic diseases, thus furthering demand for ambulant care [8]. Third, there is a general trend towards outpatient care by hospitals. According to the Avalere Health analysis of American Hospital Association Annual Survey, the percentage of revenue for community hospitals in the United States has increased from 25% to approximately 44% in 20 years between 1992 and 2012 [9]. The average length of stay of a patient after surgery dropped from 7.0 in 1993 to 5.4 days in 2013.

These developments drive the research on technical support systems in home and care environments. Applications for such include automated assessments such as mobility measurements [10], activity monitoring such as the detection of Activies of Daily Living (ADLs) [11] or fall detection for automated emergency calls [12]. Existing literature shows the importance of ADLs such as bathing and eating and Instrumental Activities of Daily Living (IADLs) as indicators of physical and cognitive abilities of elderly individuals [13]. ADL performance has also been shown to improve technology design for patients suffering from Alzheimer disease [14]. Overall, ADLs and IADLs are considered the "gold standard" for measurement of functional ability. Thanks to the increasing availability of sensors to detect motion (passive infrared motion sensors) and activities (door contact sensors, acceleration sensors, RFID chips and sensors), there is a strong hope to eventually perform such tests automatically. At the same time, such sensors can also provide data for other useful services to increase safety and comfort of people in their home [10], [15], [16]. Most of these, however, have only been tested in one-person households.

To preserve a maximum of privacy and comfort while at the same time collecting data necessary for the application, many approaches include the use of ambient sensors such as motion sensors and light barriers. Since the data collected from these sensors does not carry identifying information, use of any such application in settings where more than one person – the patient – moves or resides becomes difficult. Complex sensors, such as cameras and microphones are usually considered invasive and are thus rarely accepted in living spaces. Body-worn sensors are often forgotten or ignored due to discomfort. Binary sensors such as light barriers and motion sensors are easy to retrofit, have relatively little power consumption and can be installed unobtrusively. A no-requirements sensor model also enables us to install more complex sensors (such as laser scanners or depth-finding cameras) as required. The necessary information can be extracted from their data by partitioning the sensors' range and converting activity in each partition to a binary signal.

To separate data from multiple persons moving in a space monitored by binary sensors, a multi-target tracking algorithm using Bayesian estimation and multi-hypothesis tracking is presented. This algorithm makes no assumptions on the selection and placement of sensors or sensing technology. Tracking takes place on a graph of the sensors and their spatial relation. It is thus not helpful in determining the precise location of a present person, but at (or below) room-level accuracy. More importantly, however, it can help determining when there is more than one person present, and helps to separate the activity data. This algorithm performs particularly well on low-resolution data, such as when only few binary sensors are used. To study its precision, the algorithm is tested across various sets of sensors, varying by placement and number. The data was recorded at a home lab and was labelled and verified using video recordings. A decreasing number of sensors will likely have an impact on the tracking accuracy, but is important in regard to energy consumption, costs and user acceptance. It is shown that data from two residents in an apartment can be separated with high (>90%) accuracy, and that the selection and placement of sensors can play a significant role in tracking accuracy. To test its applicability, as a second evaluation a field trial is conducted in which three two-person households are equipped with sensors and an implementation of the algorithm, running on a small-form-factor PC, determines the number of people present.

The remainder of this article is structured as follows: Section II summarizes related works on multi-target tracking and activity monitoring in the home using binary sensors. Section III describes the algorithm and the theoretical principles surrounding data association and multi-hypothesis tracking for single- and multi-target tracking. Section IV describes how the algorithm and its implementation were evaluated, including data preparation, the sensor placement concept and the field trial setup. The results of the evaluation are presented in Section V. Section VI summarizes the article and Section VII gives an outlook on future work.

II. Related Work

Prior work has shown that data collected from sensor networks allow for the deduction of information used in activity monitoring, care assessments and behavior modeling. Target tracking, in particular multi-target tracking, is a task often applied to visual data such as video feeds and images. The practical application of multi-target tracking in binary or low-resolution home sensor networks has been subject to little research.

A. Target Tracking in Home Sensor Networks

Target tracking in the home is a fundamental problem of ubiquitous computing, and proposed solutions span a variety of sensors, including cameras, laser scanners, RFID (Radio frequency identification) and infrared or ultrasound badges [17], [18], [19].

Wilson and Atkeson describe an algorithm for tracking of multiple persons and their activity status in a binary sensor network [17]. Similar to the proposed use of a weighted graph in this article, the authors use a transition matrix representing transition probabilities between sensors. By keeping track of the targets' identities, personal motion models emerge. The data association is achieved using a particle filter. During a five-day experiment in a house instrumented with 49 sensors (contact switches, motion sensors), data during two-person scenarios was correctly assigned 82.1% of the time. While the approach solves both the data association as well as the identification, it is based on individual motion models and thus relies on data that is commonly unavailable.

Krüger et al. use a particle filter and *action plans* to assign sensor events from motion sensors and light switches to tracks and simultaneously identify the target [20]. Action plans describe action sequences in terms of sensor data. These plans can be synthesized or learned from historic data. For the evaluation, an office corridor was equipped with six light switches and six motion sensors. The mean squared error across time and all targets is reported as approximately 0.26 for two-person scenarios. The work shows how – similar to trained motion models – previous knowledge of a person's plans can help tracking individuals in binary sensor networks.

Oh and Sastry perform tracking on data of binary sensor networks and passage connectivity graphs [21]. The graphs are calculated from transition probability matrices. A tracking algorithm, derived from the Viterbi algorithm, pruning strategies and multiple target tracking extensions are presented. No evaluation on real world data is conducted.

Marinakis et al. derive the topology of a sensor network in terms of transition times and probabilities from data of unspecified sensors [22]. The authors use Monte Carlo Expectation Maximization to assign activity to agents (people present) in order to build a graph of the sensor network. 95% of the topology of simulated node graphs is recovered correctly. The results for a trial using a network of cameras and photocell-based sensors are not reported.

B. Activity Monitoring in Home Sensor Networks

The aim of this work is to test the performance of a multi-target tracking algorithm on data collected by binary sensors such as light barriers and motion sensors. The motivation is to *a)* collect activity data of an ambulant care patient, which is optimally collected during times when the patient is alone at home (to make sure the data originates from the person in question only), and *b)* to activate security measures such as fall detection and automated emergency calls, which are only necessary when the person is alone temporarily.

Researchers at the Tokyo Medical and Dental University have collected data from homes instrumented with binary sensors (motion detectors and contact switches) for one year. Patterns of activity such as absence, use of stove and sleeping times were "clearly identifiable" [23]. Researchers at the University of Virginia used an array of motion detectors and contact switches to detect ADLs [24]. Sensor readings are clustered to groups based on room, duration, and time of day to show that many clusters correspond to ADLs.

Numerous studies show that data collected from sensor networks in living spaces allow for the deduction of information relevant in applications of activity monitoring, care assessments and behavior modeling. Logan et al. showed that ambient motion-based sensors provide the most useful information for detection and classification of daily in-home activities in a study compared to RFID, on-body and on-object sensors. In their study, infrared motion sensors yielded the best results overall, although classification performance on this data was better on activities that are strongly correlated with locations in the home, such as ”watching TV” and ”meal preparation” [25].

Data from binary sensors can also be used to calculate average room residence time and frequency: Assessment tests are partly realizable by using recordings from light barriers and reed contacts alone [15]. The authors argue that light barriers alone do not constitute sufficient evidence of a person entering a room, because people may change directions between rooms. It is suggested to combine light barriers with sensors covering larger areas. Room residence times are calculated by manually labeling the sensors constituting a room using a floor plan and knowledge of the sensors’ placement. In a similar study, the authors model user behavior of a resident from the probability of location at a certain time of day and the frequency of presence in a location in a defined period of time [26]. Models are created for rooms individually (bathroom, bedroom, living room, kitchen). Based on the number of anomalous behavior detected, the authors conclude that the models’ performance varies by room: Presence in the bathroom is best modeled duration-based, while the timeslot-based model yielded better results for the other rooms.

Frenken et al. [10] use ambient sensors in an attempt to automate measurement of mobility and gait velocity, as required in the Timed Up and Go assessment [27]. For this, five apartments are equipped with home automation sensors and one with an additional laser range scanner. It is shown that the data is suitable to compute gait velocity at home. While data from the laser range scanner is proven to be more precise than home automation sensor data, no statistical post-processing or filtering was performed on the latter.

Shin et al. use motion sensors to determine normal behavior and to detect unusual behavior [28]. The behavior is described by ”activity level, mobility level and non-response interval (NRI)”. Days are divided into timeslots for which the features are computed. The activity level measures a person’s motion at a certain timeslot by summing up the number of motion sensor events detected in its coverage area. The mobility level is computed by counting location changes during a certain timeslot as depicted by two different sensors triggering. The non-response interval is the duration between two consecutive movements and is calculated by summarizing these no-motion durations for each timeslot. Using an approach coined ”support vector data description”, normal behavior and anomalies are correctly classified with an accuracy of more than 90%.

Skubic et al. use motion and temperature sensors as well as a pneumatic bed sensor to determine alert parameters (bathroom activity, bed restlessness, kitchen activity, living room activity) to model normal behavior [29]. The parameters are computed for three time periods, a 24-hour day, daytime (8am to 8pm) and nighttime (midnight to 6am) to describe the behavior learned over the span of two weeks. Fuzzy pattern tree and support vector machine classification is used to determine whether the feature vector of the current day is abnormal. Both classifiers achieved similar results.

Floeck et al. use automation sensors to identify human behavior by means of inactivity profiles [30]. Inactivity is defined as the time between consecutive sensor events, so that inactivity ends with any detected sensor event. Plots of inactivity intervals over a learning period of 31 days are created for reference, whereby inactivities due to absences are not regarded. To classify the current day as normal or abnormal day, the day’s inactivity vector is compared with the corresponding timeslot in the reference vector. Dice’s coefficient is used to measure the similarity of these vectors: An inactivity value is identical to the corresponding reference value if it is in a predefined tolerance region of the reference value. A modified approach uses the maximum duration of inactivity over several days: A threshold is calculated based on recorded maximum durations plus an additional tolerance of 30 minutes. When abnormal behavior is detected, an alert is generated.

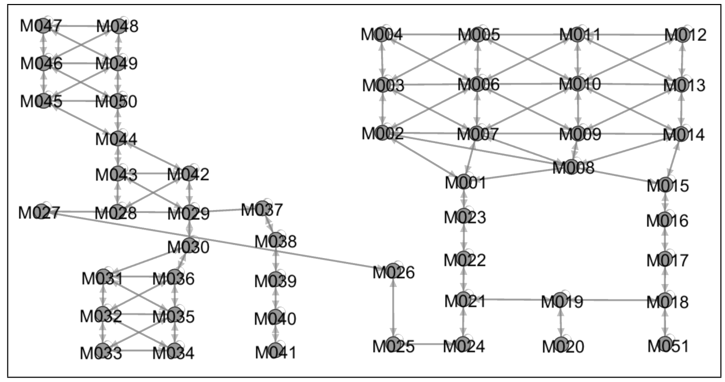

Figure 1. Graph of sensors (with their internal IDs and their spatial relations used in the evaluation (adapted from [31]).

As can be seen, most studies of activity monitoring are designed to be used in single-person households, although the foundations for data association and identification have been laid. Whatever the approach is, it is important that a tracking algorithm works on a multitude of sensor technologies, and without restricting the setup in terms of sensor density or placement. In the following section, such an approach, based on multi-hypothesis tracking, is presented. An evaluation of its performance using different numbers and placement of sensors follows.

III. Approach

To describe the tracking approach used in these studies, the idea of the sensor graph is introduced first. Then, Bayesian filters and how they can be used to help track individuals on this graph will be explained. Finally, the multi-hypothesis tracking approach and how it differs from previous implementations is described.

A. Sensor Graph

A graph of sensors $s_1, \ldots, s_N$ is defined as a weighted, directed graph $G = (V, L)$, where $V = \{1, \ldots, N\}$ is the set of nodes in the graph representing sensors, and L is the set of all edges (u, v) for which there is a direct passage from the sensing region of sensor u to the sensing region of v which does not intersect any other sensing regions. Informally, two sensors u, v are connected if it is possible for a person to traverse from the sensing region of u to the sensing region of v without activating any other sensor.

Each resident in the target space is represented by a discrete Bayesian filter on an unweighted, undirected graph

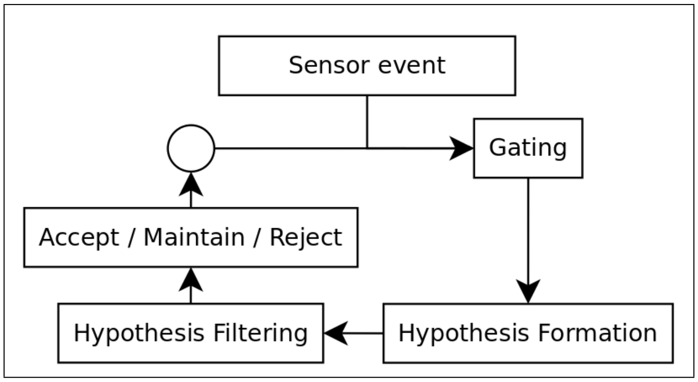

Figure 2. Hypothesis formation overview.

consisting of sensors as nodes and edges representing their spatial adjacency. For the evaluation data, this graph was published by Crandall et al. [31] (Figure 1). If the adjacency relations are not known, they can be approximated by a path planning algorithm using a floor plan, if available [32], or can be generated from prerecorded data [33].

B. Tracking of individuals

Bayesian filters estimate the state of a dynamic system from noisy data. The location of a person is described as a probability distribution over all nodes in the graph, rather than selecting the most likely sensor as their location, to represent the location of each individual. This is done because it may help to calculate a more precise location later on, especially when sensor regions overlap. More importantly still, it helps the tracker to recover more quickly when a noisy measurement is assigned to the individual's track. Lastly, the aim is to replace the manually constructed, unweighted graph with a weighted graph that is constructed from in-situ recorded data, which reflects individual behavior by using transition probabilities between sensors as edge weights (cf. [22]).

All sensors are subject to noise, and although only binary data is sought, the sensors may still fail to detect activity or detect irrelevant activity, such as from pets or through the window. Due to duty cycle regulations in the range of 868 Mhz, many motion and home automation sensors come with an additional source for noise: measurement delay. Many sensors do not measure or report measurements for a specified amount of time after triggering. This period can last from a few seconds to several minutes due to radio communication regulations. This results in sensors missing the presence or movement, thus breaking the continuity of measurements of a motion track.

At which point and how many filters are created – that is to say, how many individuals are assumed to be present – depends on the performance (*belief*) of the previously existing track: When new measurements cause the current data to be more likely when assigned to more or fewer tracks (= individuals) than before, a new filter is spawned or an existing filter is discarded. The data could be bundled into larger updates within reasonable time frames (cf. [34]), but in this case an update occurs for each new sensor event.

C. Multi-target tracking

Fundamentally, to distinguish activity of two or more persons in an area monitored by binary sensors, the activity must be spatially separable on the graph. Ideally, there is an inactive sensor between two persons. This concept has been thoroughly described by Oh and Sastry [21]. However, whether two neighboring sensor events are assigned to one or two targets is determined by the weights of the graph and the evaluation function.

1) Multi-Hypothesis Tracking: When new sensor data arrives, hypotheses are created by considering all possible assignments of the data to existing and new tracks ("hypotheses") until the filter's *window size* is reached. This is particularly useful in a low-resolution setting like this, where individuals may occlude each other in sensor readings for any period of time. The window size in multi-hypothesis tracking (MHT) describes the maximum number of events (or time steps) that are considered before choosing a likely hypothesis. Windowing is necessary to limit the number of possible hypotheses and to limit the information loss in case no acceptable hypothesis is found and the data is discarded. The influence of the window size on tracking accuracy has been shown previously [35]. For this evaluation, a window size of 10 events is used. The sensors used in the field trial collected an average of 25.5 events per hour, making the average windows size 23.5 minutes. The sensors used in the home lab have a delay of approximately two seconds, making the temporal extent of the windows much smaller.

The idea of multi-hypothesis tracking dates back to 1979, when Donald B. Reid published "An algorithm for tracking multiple targets" [36]. Reid's algorithm was developed to work on data from a continuous scale sensor (radar). Therefore, Reid speaks of associating measurements to clusters. In the work presented here, the target space is discrete (nodes on a graph), and targets and their locations are stored as a probability distribution over the space using Bayesian filters.

There are two significant differences between Reid's original work and the approach described here. First, in accordance with Reid's *type 2* sensor, the sensor model used in this work expects *positive reports* only, meaning that only sensor data reporting activity is considered. However, tracks are updated per hypothesis, rather than generated and filtered individually (*hypothesis-oriented MHT*). This means that hypotheses are not constructed from *compatible* tracks, but all possible combinations of updates of existing hypotheses. Second, the tracker is updated every time a sensor reports activity. Because of this, and the fact that the given state space is discrete, computational complexity is reduced. For a more detailed description of track- and hypothesis-oriented MHT, see Blackman [34].

For each triggered sensor, a new hypothesis based on all previously existing hypotheses is created, in which the triggered sensor is

- considered noise and discarded,
- used to update one of the existing filters, or
- assigned to a new filter.

2) Filtering: For an unspecified number of possible targets the number of possible hypotheses follows Bell's Number B_{n+1}. Due to the exponential growth ($> 4.74 \times 10^{13}$ for 20 events), a number of filters to optimize computation efficiency must be employed.

All hypotheses must pass a gating function before they are considered for evaluation (see Figure 2). In this case, this gating function is a simple comparison of the prior probability of each filter to a threshold value. Afterwards, hypotheses are filtered based on confidence, noise ratio and similarity. This procedure is performed until a single hypothesis remains or the window size is reached. In the former case, the hypothesis is accepted, the underlying Bayesian filters updated, and the window size reset. In the latter case, all hypotheses are evaluated. If no single, dominating hypothesis can be found, all hypotheses are discarded and the underlying filters reset.

The size of the window strongly influences the performance of the algorithm. A larger window size will result in a larger number of correct associations, but also in a larger number of discarded sensor events [35].

Figure 2 depicts the general multi-hypothesis tracking logic. For a more in-depth description of multi-hypothesis tracking, see Blackman [34] or Reid [36].

IV. Evaluation

In the following, the data that was collected for both the home lab tests and field trial in terms of how it was collected and what was evaluated is described. For the home lab tests, several subsets of sensors were evaluated to study the impact of sensor placement on tracking performance. For the field trial, sensors were installed based on these results.

A. Experiment I: Home Lab

1) Data: The data used for this evaluation was recorded at the Center for Advanced Studies in Adaptive Systems (CASAS) at the University of Washington [38]. It shows activity of two residents of a smart home environment, residing in a 4-room, 2-story apartment for approximately 8 months. For this evaluation, subsets of the data recorded by the 50 motion sensors mounted to the ceiling are used. The smart home is also equipped with contact sensors on doors and cabinets, temperature, water and electricity sensors. For the purpose of this trial, however, motion sensors offer the most precise and least noisy data.

Data for which at least both residents are present and active is used. Among those, time frames

- that last at least 20 minutes or contain at least 300 sensor events,
- in which both residents change rooms at least once, and
- in which neither resident is inactive for more than 20% of the time

are chosen. The result are twenty time frames, with 330 to 910 sensor events with durations between 24 and 530 minutes. After selection, each of the 13321 sensor events was labelled as originating from Resident 1, Resident 2 or a third person using the manually labelled events and the laboratory's floor plan.

2) Data Association: The algorithm can track any number of targets. However, the intended area of application – small households – allows one to use an evaluation function that is tailored towards few targets (1-3). For this evaluation, the algorithm was optimized to track two targets by using an evaluation function that favors one- and two-track hypotheses. Equation (1) describes the evaluation function, where h is the hypothesis in question, $conf(p_n)$ is the belief of the Bayesian filter at the most recent event location n, $\|p\|$ is the number of paths (= targets) in h, and m is the expected number of targets in the sensor space.

$$eval(h) = \frac{\sum_{i=1}^{n} conf(p_n)}{\frac{\|p\|^2+m}{m+1}} \tag{1}$$

3) Sensor Placement: To get a better understanding of how the number of sensors affects tracking accuracy, the algorithm is also applied to subsets of the original set of sensors in decreasing size (40, 30 and 20 sensors). Instead of choosing the sensors randomly, characteristics of sensors deemed possibly influential on tracking performance were chosen:

a) Number of neighboring sensors: Based on the assumption that sensors in doorways, which usually have few neighboring sensors, are critical in tracking room transitions, those in larger areas with many neighboring sensors are removed. The number of neighboring sensors can be calculated from the sensor graph.

b) Duration of stay: Given that tracking stationary targets is much simpler than moving targets, subsets of sensors that cover areas in which the average duration of stay is short are considered. The duration of stay can be calculated from the duration between consecutive sensor events in recorded data.

c) Activity: Considering the application of in-home activity monitoring, it is imperative that the placement of sensors for tracking accuracy improvement does not interfere with the necessity of covering those areas in which the majority of activity is taking place. Thus, sensors based on the amount of activity covered are selected and filtered. The amount of activity covered by a sensor is simply calculated by the number of times it is triggered.

These criteria were used to create subsets of data of varying size, selected by increasing, as well as decreasing order of the respective criterium (cf. Figure 6).

4) Sensor Clustering: The procedure of selecting subsets of sensors for tracking performance evaluation was also conducted for sets of 10 sensors. However, due to the selection criteria, most of the sets had removed whole rooms, and in one case all data from one individual. Thus, in order to evaluate tracking performance on 10 sensors, the sensors are clustered manually by rooms and spatial adjacency (see Figure 3), and resulting clusters are treated as individual sensors. This results in a more realistic scenario, in which motion sensors often cover different size areas up to whole rooms.

For this evaluation, data from all sensors is used, but the sensor IDs are replaced with IDs for their corresponding cluster. This way, use of all sensor events is made, but their spatial resolution is decreased.

B. Experiment II: Field Trial

To test the applicability of the approach in a more realistic setting, it is tested in an evaluation in three households equipped with retrofitted sensors. Since the identity of the residents cannot be reconstructed without video recordings or other identifying data, this study focuses on comparing the actual and the calculated number of people present. Thus, claims about the false associations between residents cannot be made. However, this information helps developing home security and care support services that rely on information on the number of people present.

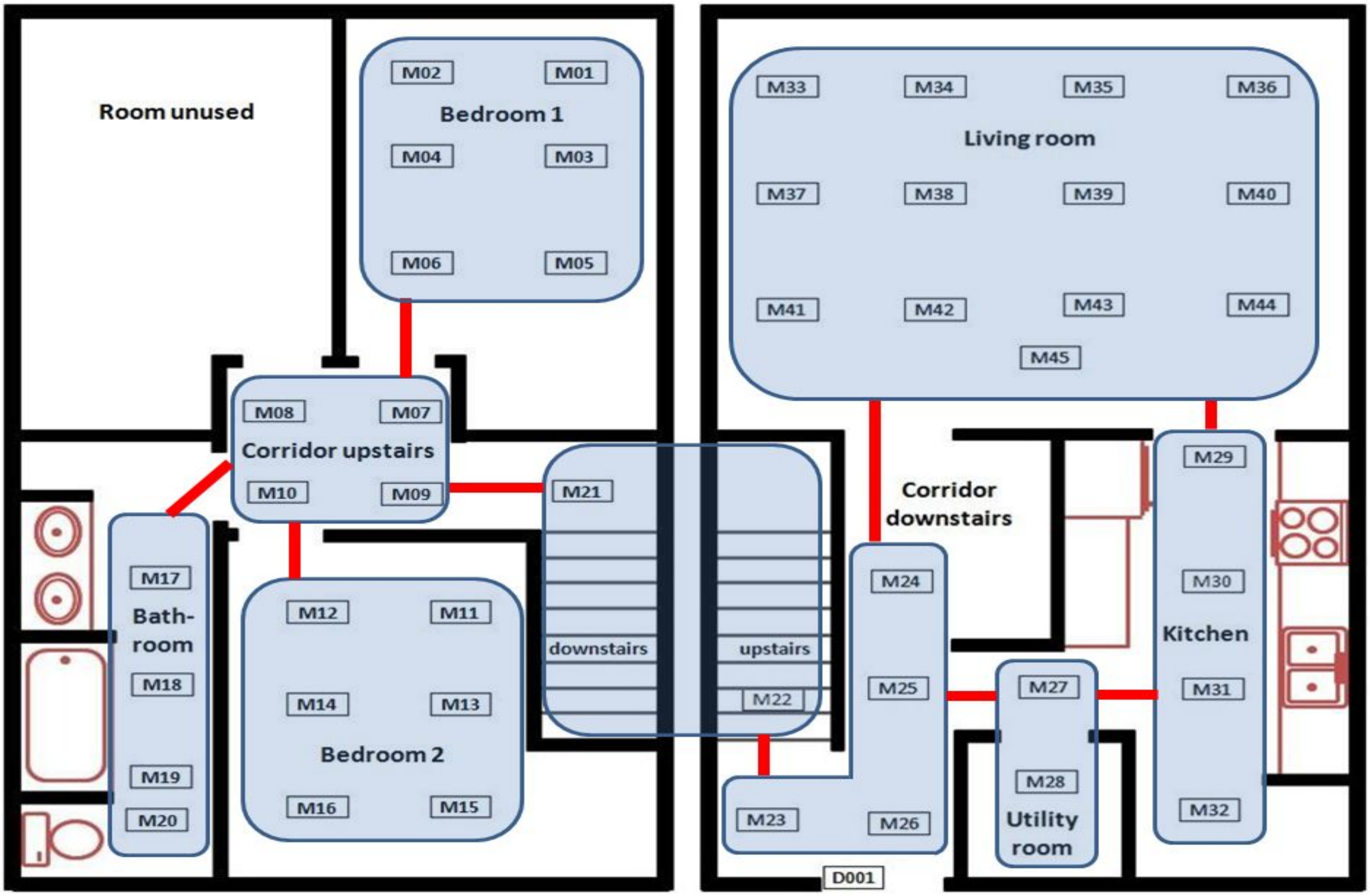

Figure 3. Layout of home lab, with placement of sensors and their corresponding clusters. Red lines between clusters lines indicate spatial adjacency as implemented in the sensor graph. After Crandall & Cook [37].

1) Setup: Each apartment was equipped with 13 to 15 PIR motion sensors, depending on the size of the apartment. Five of the sensors were installed to cover as large an area as possible, to enable an ”inactivity monitoring”. These sensors were also used in a separate study to detect unusual behavior (e.g., falls) in order to automatically invoke an emergency call. The remaining sensors were placed to cover room transitions (e.g., hallways and door frames) and smaller areas in which most activity takes place (such as the kitchen counter and the bathroom sink). The sensors further included two light barriers and two pressure mats near the front entrance. Their primary task was to enable detection of people entering and leaving the apartment. For the sake of this evaluation, however, they only provide further tracking data. A schema of the hardware setup is shown in Figure 4. A detailed description of the care related services implemented in the project has been written by Pauls et al. [39], [40] and Gerdes et al. [41].

The participants were recruited by the Johanniter-Unfall-Hilfe e.V. as part of the research project *Cicely*. Three participant couples were recruited. All lived in a two-room apartment, which was subsequently equipped with the sensors, a ”small-form-factor PC”, a radio receiver to collect the sensor data and a UMTS router to place emergency calls and to remotely monitor the correct functioning of the setup.

2) Data: During the evaluation using home lab data, the tracking results could be validated using video recordings from video cameras installed in the home lab. During the field trial, use of video cameras was considered impractical as well as a burden on the participants’ privacy. Instead, they were asked to fill out an ”activity log” (see Figure 5), which states sleep times and the number of people present for the whole duration of the trial. The sensors recorded data for a period of up to six weeks. With an average of approximately 610 sensor events per day, 55000 sensor events were recorded.

Not all data could be used for analysis, as participants missed to fill out the activity log for parts of the trial. Thus, this analysis focuses on the time frames in which the number of targets can be precisely reconstructed.

3) Data Analysis: During the trial, the algorithm ran as part of a software framework developed for management of Ambient Assisted Living (AAL) installations. Sensors were connected to this framework using two 868 Mhz radio receivers, which were in turn connected to the PC via USB. The tracking algorithm is implemented in Java, and was started as a service. Each time a sensor event was registered, the tracker was updated. At the end of the trial, the raw, recorded data as well as the tracking output was collected from the households and the estimated number of people present was extracted from the tracking results.

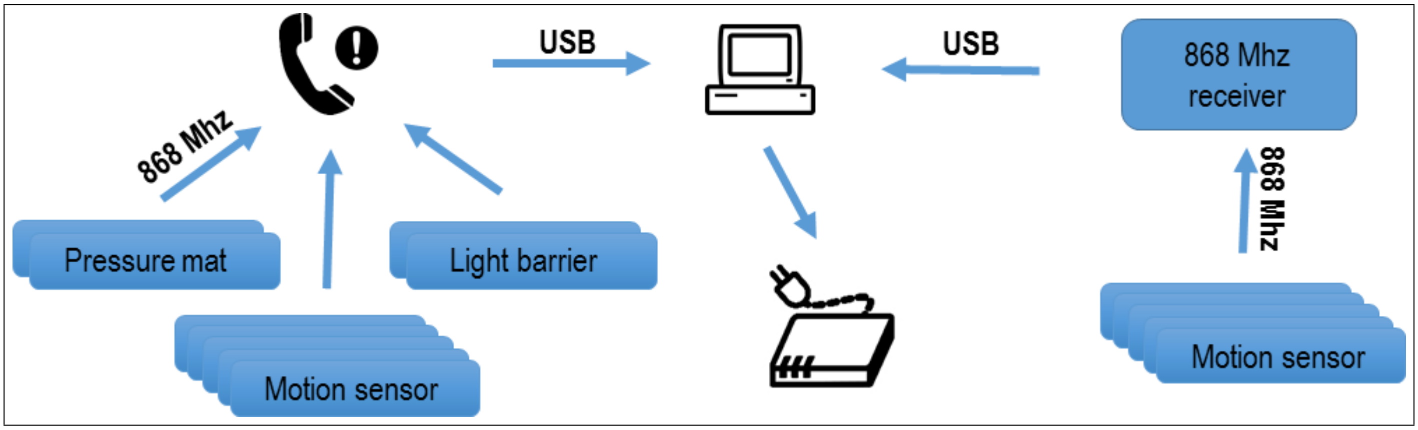

Figure 4. Hardware setup for the field trial including emergency phone, router, radio receiver and sensors.

Apart from the sensor graph, which had to be implemented after installation, the software setup was equal across all households. The number of sensors varied based on the size of the apartment.

V. Results

A. Home Lab

Tracking accuracy using all sensors is 90.3%. This is the percentage of all sensor events across all time frames that are correctly associated to any of the targets. Correct association is determined by the labels in the dataset, which have been added manually after reviewing video recordings. The accuracy of individual time frames ranges from 62.1% to 99.5%, with a median of 93.1%. After reviewing the individual time frames, it can be seen that the variance in accuracy between time frames can be largely attributed to the varying complexity of the recorded activities. In some time frames, the residents spend little time in the same room, but move independently across the lab, thus making it easy for the tracking algorithm to separate the data. In other time frames, the residents spend much time close to another.

The median rate of false associations (events that are falsely associated to another target) per time frame is 5.88%, making false associations the largest source of error. It is clear why: While the other errors (wrongly discarding the data as noise (0.44%) and failing to associate the data to any target (0.59%)) commonly affect singular sensor events, false associations usually cause a large number of consecutive events to be associated falsely.

1) Subset Data: Figure 6 shows tracking accuracy across sensor subsets. The sets vary by the number of included sensors (x-axis) and their selection criterion. As can be seen, tracking accuracy generally decreases with reduced sensor count. This is to be expected as the resolution of the tracking space decreases and situations with overlapping motion increases. Down to 30 sensors, tracking accuracy decreases only slightly for all but one sensor set. For the set of sensors with much overall activity, accuracy even increases slightly. The graph also shows that performance variation increases with a decreasing number of sensors. While tracking accuracy varies between 85.8% and 90.6% with 40 sensors, with 20 sensors accuracy ranges between 59.5% and 88.2%.

Datum	Uhrzeit		Bettruhe	Abwesenheit	Besuch	Personenzahl
	von	bis				
28.08.15	14:20	23:30		X		1
29.08.15	1[illegible]:00	22:00		X		2
30.08.15	06:20	15:45		X		2

Figure 5. Example excerpt of an activity log showing date, time (from, to), bed rest, absence, visit and number of people.

2) Clustered Data: Tracking accuracy on the clustered data set is 77.2%. This result is considerably lower than results with ten individual sensors. However, using the selection criteria explained above, more than half of the 10-sensor sets either removed whole rooms from the data, or all data from one of the residents. As a result, the evaluation using ten individual sensors is considered invalid. Considering that each of the rooms the clusters represent could be monitored using a single well-placed sensor, the tracking accuracy is surprisingly high relative to the number of sensors used.

In terms of estimation of people present, the algorithm was correct 84.5% of the time. This means that, for 84.5% of sensor events and tracking updates, the number of tracks spawned by the algorithm matches the actual number of people present. As previous studies have shown, this performance is at least as good as with the full, unclustered set of sensors [42].

B. Field Trial

Across all households, the number of people present was correctly estimated 84.0% of the time. This means that, for 84.0% of sensor events and tracking updates, the number of tracks spawned by the algorithm matches the number of people present according to the activity log. This includes data from three households and all motion sensors. This matches the results of the clustered home lab data. The average time for the algorithm to recover (to restore the correct number of tracks) after a tracking error amounts to 9.7 sensor events, or 101 seconds.

Upon further inspection, it can be seen that the largest share of incorrect estimations lies in the night time. It becomes obvious that, when an individual is inactive, their track profile will slowly fade during updates caused by activity of other

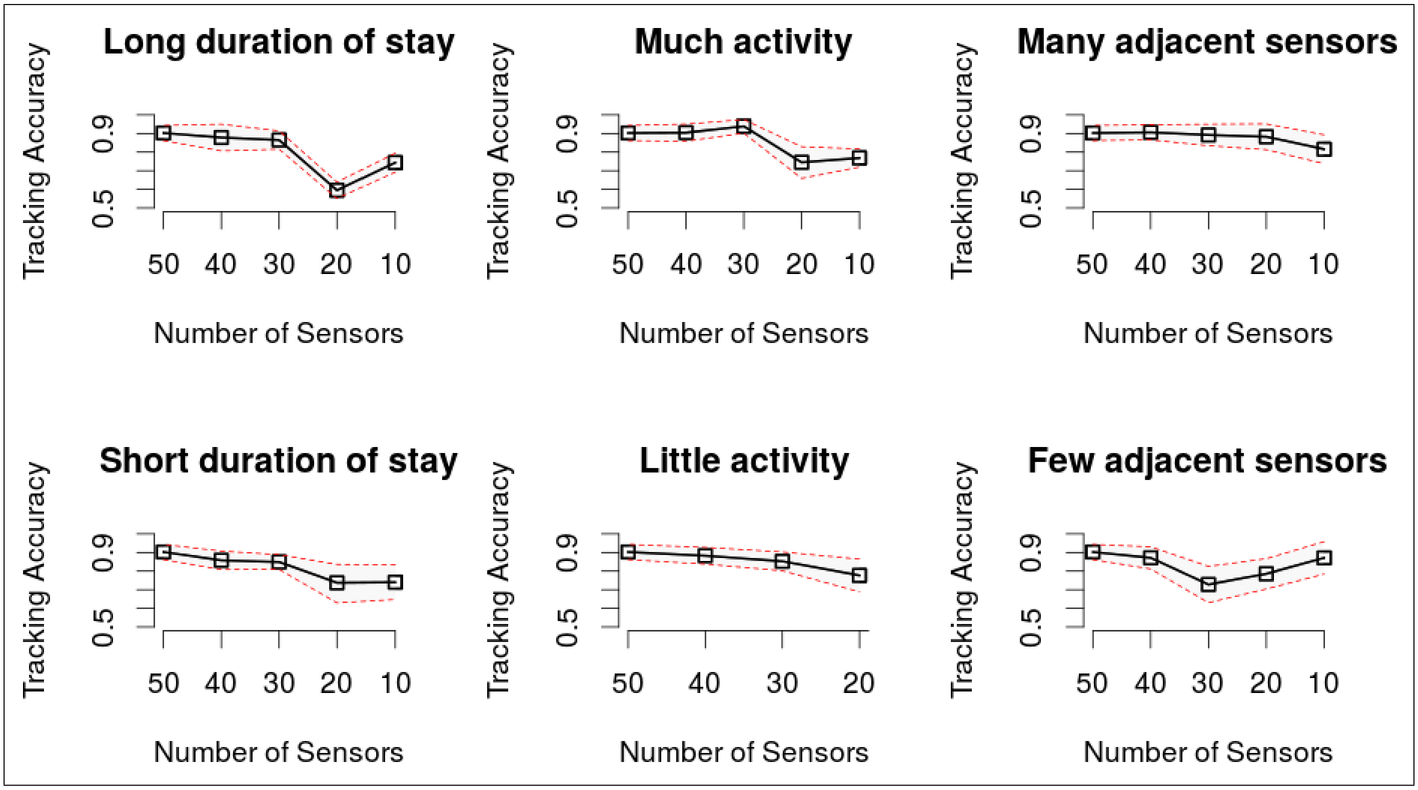

Figure 6. Tracking performance across sensor groups.

individuals. This happens, for example, when one of the residents is restless at night, or when two individuals are monitored by one and the same sensor.

Furthermore, the estimation error in terms of their impact on home security and ambulant care scenarios is considered: The algorithm could be used to activate or deactivate an alarm system when the house is empty, or to place an automated emergency call when a sole individual shows unusual behavior. Thanks to the light barriers and pressure mats, it is known when an individual left the apartment. Thus, the timestamps of the outer pressure mat with an appropriate change in the number of tracks in the output of the algorithm are compared. On average, it took 40 seconds between a person leaving the apartment and the appropriate fading of their track path. Since the number of updates depends on the amount of activity, this amount of time varied up to 174 seconds.

VI. Conclusion

Technical systems to measure mobility parameters and to implement safety concepts in the home are one possibility to help caretakers manage the future challenges imposed by current and future changes in demography and care. In the upcoming decades, there will be continuously more people requiring care and fewer young people to carry out this care.

The article at hand describes an algorithm to enable previously developed concepts of care support and health monitoring in the home for households of more than one person. This is achieved through a multi-target tracking algorithm in a space monitored by binary sensors. The system utilizes ambient binary sensor technologies, which is considered less obtrusive than video cameras and microphones, and more suitable for older people. It enables the separation of sensor data generated by multiple persons in smart home environments without the need for identifying sensors. The approach makes it possible to install services of home security and ambulant care support in multi-person households, which previously have only been possible to use in single-person households.

The algorithm makes use of a graph consisting of sensors as nodes and their spatial relations as vertices. Compared to other related works, the algorithm works particularly well in low-resolution settings (i.e., with few binary sensors). It was shown that tracking accuracy can be improved by placing sensors based on activity characteristics. For example, sensors with many neighboring sensors provide a consistently higher accuracy than those with few, and sensors in places where the duration of stay is long on average prove to be less beneficial than those where duration of stay is short.

It was shown that the decrease of tracking accuracy resulting from smaller sets of sensors (i.e., decreased target space resolution) can be largely absorbed by selective placement of sensors. It was shown that tracking two targets in a network of 20 sensors or more can be achieved for over 90% of the time.

It must be noted that differences in tracking performance may not only be due to advantageous sensor placement, but also due to favorable data: While tracking in space with many adjacent sensors works well, it neglects in part space where tracking might be particularly difficult but useful, such as in narrow hallways. The share of total events covered by the different subsets of sensors range from 11 to 98%.

The algorithm was also successfully used in a field trial and it was shown that it is possible to correctly estimate the number of people present over long periods of time. The algorithm was also found to be useful for applications of home security and technical care support, as it swiftly determined when a person has left or entered the monitored area. A conclusion of whether the error rates are sufficiently small for use in practical applications cannot be given, because this depends largely on the application itself. For example, in order to collect mobility data of an individual during times they are alone at home, only a small percentage of the overall time is required to collect this data. However, an accuracy of more than 90% may still not be sufficient for the data to be used in medical decision making. It was shown that the implementation of the algorithm works reliably with different sensors and sensing technologies to separate sensor data of two residents in a shared apartment. The number of people present was correctly calculated 84.0% of the time. This is less than expected, but can largely be explained by difficulties differentiating between absence and inactivity, such as when people are sleeping.

The results of the field trial show that the tracking algorithm is sensitive to differences in the amount of activity between individuals. This is due to the fact that the number of updates is based on the number of sensor events received, and every track that is not assigned any event during an update fades, i.e. , a smoothing function to this track's probability distribution is applied. When the peak probability of a track falls below a certain threshold, it is removed from the hypothesis space. This is done to enable fading of tracks when individuals leave the house without the need to rely on special sensors monitoring the entrance and exit doors. However, this is cause for tracking errors when one individual is inactive, such as during night. Neither the installation of additional sensors nor the omission of the smoothing function seem viable options at this point, and the problem remains unsolved.

VII. Future Work

The home lab experiment gives insight into the importance of sensor placement for multi-target tracking using binary sensors. The next step must be to find the ideal sensor setup for the data used in this evaluation, which may be a mixture of the sensor subsets and criteria examined here, and may contain criteria which were not yet considered.

Furthermore, the algorithm's performance with more than two targets must be evaluated. The tracking of two to three individuals is sufficient for most applications this approach is targeted towards. However, the approach is not theoretically limited in its applicability and studies with three or more individuals may provide better insight into the limitations of the algorithm.

It is further planned to include identifying information in the algorithm so as to not only associate the data to tracks, but to identify the target. This could be easily achieved using the labelled data from the home lab (c.f. Nillius et al. [43]), but ultimately it may be possible to assign tracks to individuals using unsupervised clustering techniques, using the meta information of sensor placement and expected number of people present. This way, the sensor graph could be replaced by individual motion models, further improving tracking accuracy.

Acknowledgment

The authors would like to thank the team of the Center for Advanced Studies in Adaptive Systems (CASAS) at the Washington State University for making their data publicly available. The authors would further like to thank the *Cicely* consortium for making the field trial possible.

This work was funded by the German Ministry for Education and Research (BMBF) within the research project *Cicely* (grant 16SV5896).

References

[1] S. M. Müller and A. Hein, "Multi-Target Data Association in Binary Sensor Networks," in AMBIENT 2015, The Fifth International Conference on Ambient Computing, Applications, Services and Technologies, July 2015, pp. 79–84, [Last access: 2016-06-15]. [Online]. Available: http://www.thinkmind.org/index.php?view=article&articleid=ambient_2015_3_50_40044

[2] United Nations, "World mortality report," Tech. Rep., 2013. [Online]. Available: http://www.un.org/en/development/desa/population/publications/pdf/mortality/WMR2013/World_Mortality_2013_Report.pdf

[3] J. Vaupel and H. Lundstrom, "The future of mortality at older ages in developed countries." 1994.

[4] G. E. Stockmayer, "The demographic foundations of change in us households in the twentieth century," Ph.D. dissertation, University of California, Berkeley, 2004.

[5] M. Eisenmenger, O. Pötzsch and B. Sommer, "Germany's population by 2050, Results of the 11th coordinated population projection," Federal Statistical Office, Tech. Rep., 2006.

[6] S. Ruggles, M. Sobek, T. Alexander, C. A. Fitch, R. Goeken, P. K. Hall, M. King, and C. Ronnander, "Integrated public use microdata series: Version 3.0 [machine-readable database]." 2004.

[7] H. Rothgang, R. Müller, and R. Unger, Themenreport "Pflege 2030": was ist zu erwarten - was ist zu tun? Bertelsmann Stiftung, 2012.

[8] W. H. Organization et al., "The european health report 2009: Health and health systems," 2009.

[9] A. H. Association et al., The AHA annual survey database. American Hospital Association, 2005.

[10] T. Frenken, B. Vester, M. Brell, and A. Hein, "atug: Fully-automated timed up and go assessment using ambient sensor technologies," in Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on. IEEE, 2011, pp. 55–62.

[11] A. Fleury, M. Vacher, and N. Noury, "Svm-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results," Information Technology in Biomedicine, IEEE Transactions on, vol. 14, no. 2, 2010, pp. 274–283.

[12] A. Sixsmith, N. Johnson, and R. Whatmore, "Pyroelectric ir sensor arrays for fall detection in the older population," in Journal de Physique IV (Proceedings), vol. 128. EDP sciences, 2005, pp. 153–160.

[13] S. Katz, A. B. Ford, R. W. Moskowitz, B. A. Jackson, and M. W. Jaffe, "Studies of illness in the aged: the index of adl: a standardized measure of biological and psychosocial function," Jama, vol. 185, no. 12, 1963, pp. 914–919.

[14] A. Mihailidis and G. R. Fernie, "Context-aware assistive devices for older adults with dementia," Gerontechnology, vol. 2, no. 2, 2002, pp. 173–188.

[15] E.-E. Steen, T. Frenken, M. Frenken, and A. Hein, "Functional assessment in elderlies homes: Early results from a field trial," in Ambient Assisted Living. Springer, 2014, pp. 3–17.

[16] N. C. Krishnan and D. J. Cook, "Activity recognition on streaming sensor data," Pervasive and mobile computing, vol. 10, 2014, pp. 138–154.

[17] D. Wilson and C. Atkeson, "Simultaneous tracking and activity recognition (star) using many anonymous, binary sensors," in Pervasive computing. Springer, 2005, pp. 62–79.

[18] M. Addlesee, R. Curwen, S. Hodges, J. Newman, P. Steggles, A. Ward, and A. Hopper, "Implementing a sentient computing system," Computer, vol. 34, no. 8, 2001, pp. 50–56.

[19] T. Kanade, R. Collins, A. Lipton, P. Burt, and L. Wixson, "Advances in cooperative multi-sensor video surveillance," in Proceedings of DARPA Image Understanding Workshop, vol. 1, no. 3-24. Citeseer, 1998, p. 2.

[20] F. Krüger, M. Kasparick, T. Mundt, and T. Kirste, "Where are my colleagues and why? tracking multiple persons in indoor environments," in Intelligent Environments, International Conference on. IEEE, 2014, pp. 190–197.

[21] S. Oh and S. Sastry, "Tracking on a graph," in Proceedings of the 4th international symposium on Information processing in sensor networks. IEEE Press, 2005, p. 26.

[22] D. Marinakis, G. Dudek, and D. Fleet, "Learning sensor network topology through monte carlo expectation maximization," in Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on. IEEE, 2005, pp. 4581–4587.

[23] M. Ogawa, R. Suzuki, S. Otake, T. Izutsu, T. Iwaya, and T. Togawa, "Long-term remote behavioral monitoring of the elderly using sensors installed in domestic houses," in Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, 2002. Proceedings of the Second Joint, vol. 3. IEEE, 2002, pp. 1853–1854.

[24] T. Barger, D. Brown, and M. Alwan, "Health-status monitoring through analysis of behavioral patterns," Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, vol. 35, no. 1, 2005, pp. 22–27.

[25] B. Logan, J. Healey, M. Philipose, E. Tapia, and S. Intille, "A long-term evaluation of sensing modalities for activity recognition," in UbiComp 2007: Ubiquitous Computing. Springer, 2007, pp. 483–500.

[26] E.-E. Steen, T. Frenken, M. Eichelberg, M. Frenken, and A. Hein, "Modeling individual healthy behavior using home automation sensor data: Results from a field trial," Journal of Ambient Intelligence and Smart Environments, vol. 5, no. 5, 2013, pp. 503–523.

[27] D. Podsiadlo and S. Richardson, "The timed" up & go": a test of basic functional mobility for frail elderly persons." Journal of the American geriatrics Society, vol. 39, no. 2, 1991, pp. 142–148.

[28] J. H. Shin, B. Lee, and K. S. Park, "Detection of abnormal living patterns for elderly living alone using support vector data description," Information Technology in Biomedicine, IEEE Transactions on, vol. 15, no. 3, 2011, pp. 438–448.

[29] M. Skubic, R. D. Guevara, and M. Rantz, "Testing classifiers for embedded health assessment," in Impact Analysis of Solutions for Chronic Disease Prevention and Management. Springer, 2012, pp. 198–205.

[30] M. Floeck and L. Litz, "Activity-and inactivity-based approaches to analyze an assisted living environment," in Emerging Security Information, Systems and Technologies, 2008. SECURWARE'08. Second International Conference on. IEEE, 2008, pp. 311–316.

[31] A. Crandall and D. Cook, Human Behavior Recognition Technologies, H. W. Guesgen and S. Marsland, Eds. IGI Global, Mar. 2013.

[32] T. Frenken, E.-E. Steen, M. Brell, W. Nebel, and A. Hein, "Motion pattern generation and recognition for mobility assessments in domestic environments." in Proceedings of the 1st Int. Living Usability Lab Workshop on AAL Latest Solutions, 2011, pp. 3–12.

[33] S. Müller, A. Helmer, E.-E. Steen, M. Frenken, and A. Hein, "Automated clustering of home sensor networks to functional regions for the deduction of presence information for medical applications," Wohnen–Pflege–Teilhabe–Besser leben durch Technik , 2014.

[34] S. Blackman, "Multiple hypothesis tracking for multiple target tracking," Aerospace and Electronic Systems Magazine, IEEE, vol. 19, no. 1, 2004, pp. 5–18.

[35] S. Müller and A. Hein, "Multi-target tracking in home sensor networks," in Proc. of the Workshop on AI Problems and Approaches for Intelligent Environments. Prague, Czech Republic: CEUR-WS.org, Aug. 2014, to appear.

[36] D. Reid, "An algorithm for tracking multiple targets," Automatic Control, IEEE Transactions on, vol. 24, no. 6, 1979, pp. 843–854.

[37] A. Crandall and D. Cook, "Tracking systems for multiple smart home residents," Human Behavior Recognition Technologies, 2010.

[38] D. Cook and M. Schmitter-Edgecombe, "Assessing the quality of activities in a smart environment," Methods of information in medicine, vol. 48, no. 5, 2009, p. 480.

[39] A. Pauls, S. Simon, M. Schwarz-Eywill, V. Gerdes, and C. Scheve, "Intelligentes monitoring von subjektiven und objektiven parametern in der spezialisierten ambulanten palliativversorgung (sapv) – das bmbf-projekt cicely," Zeitschrift für Palliativmedizin, vol. 15, no. 03, 2014, p. PC207.

[40] ——, "Kurze informationswege in der spezialisierten ambulanten palliativversorgung (sapv) – das bmbf-projekt cicely," Zeitschrift für Palliativmedizin, vol. 15, no. 03, 2014, p. V60.

[41] V. Gerdes, S. Simon, A. Pauls, H. Rolfes, and C. Scheve, "Unterstützung der selbstbestimmung am lebensende: Empowerment für patienten in der spezialisierten ambulanten palliativversorgung (sapv) – das bmbf-projekt cicely," Zeitschrift für Palliativmedizin, vol. 15, no. 03, 2014, p. V135.

[42] S. M. Müller, E.-E. Steen, and A. Hein, "Inferring multi-person presence in home sensor networks," in Ambient Assisted Living. Springer, 2016, pp. 47–56.

[43] P. Nillius, J. Sullivan, and S. Carlsson, "Multi-target tracking-linking identities using bayesian network inference," in Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, vol. 2. IEEE, 2006, pp. 2187–2194.