Article

A Framework for Evaluating Stay Detection Approaches

Cornelia Schneider [1,*]**, Simon Gröchenig** [1] iD**, Verena Venek** [1]**, Michael Leitner** [2] **and Siegfried Reich** [1]

1 Salzburg Research Forschungsgesellschaft mbH, Mobile and Web-Based Information Systems, Competence Field e-Health, Jakob Haringer Straße 5/3, 5020 Salzburg, Austria; simon.groechenig@salzburgresearch.at (S.G.); verena.venek@salzburgresearch.at (V.V.); siegfried.reich@salzburgresearch.at (S.R.)

2 Department of Geography and Anthropology, Louisiana State University, E-104 Howe-Russell Complex, Baton Rouge, LA 70803, USA; mleitne@lsu.edu

* Correspondence: cornelia.schneider@salzburgresearch.at

Abstract: In recent years, sensors of mobile devices are increasingly used in the research field of Active and Assisted Living (AAL), in particular, for movement analysis. Questions, such as where users typically stay (and for how long), where they have been or where they will most likely be going to, are of utmost importance for implementing smart AAL services. Due to the plethora of application scenarios and varying requirements, the challenge is the identification of an appropriate stay detection approach. Thus, this paper presents a comprehensive framework covering the entire process from data acquisition, pre-processing, parameterization to evaluation so that it can be applied to evaluate various stay detection methods. Additionally, ground truth data as well as application field data are used within the framework. The framework has been validated with three different spatio-temporal clustering approaches (time-based/incremental clustering, extended density based clustering, and a mixed method approach). Using the framework with ground truth data and data from the AAL field, it can be concluded that the time-based/incremental clustering approach is most suitable for this type of AAL applications. Furthermore, using two different datasets has proven successful as it provides additional data for selecting the appropriate method. Finally, the way the framework is designed it might be applied to other domains such as transportation, mobility, or tourism by adapting the pre-selection criteria.

Keywords: significant locations; location retrieval; Active and Assisted Living; AAL; process model

1. Introduction

Due to the technological progress in recent years in sensor miniaturization, nowadays, high-quality sensors are integrated into almost every mobile device, including smartphones, smartwatches, or fitness trackers. Usually, these devices are equipped with at least one or more of the following sensors: accelerometer, gyroscope, magnetometer, and global navigation satellite system (GNSS) receiver. All of these sensors are used in different disciplines for mobile data acquisition. Thus, today, the focus is no longer on how to acquire mobile data with state of the art mobile devices, but rather, on analyzing mobile data. Despite constant progress in hardware development, performing time-aware analyses on powerful stationary computers and, particularly, on mobile devices, is still challenging for computationally intensive applications, such as movement analysis. In the context of mobile data analysis, the research field of movement pattern analysis is of particular interest to many disciplines, including transportation, mobility, tourism and health care [1–4].

In the research field of Active and Assisted Living (AAL), built-in sensors of mobile devices are used for movement analysis indoors and outdoors, for monitoring activities, and detecting emergency situations [5–7]. For example, outdoor movements of frail or cognitively impaired people can be monitored in order to assist them and/or their relatives in case they get lost or if they have fallen down. By using AAL services, these scenarios can be detected or even prevented. For the implementation of such services following questions need to be answered: (i) where does a user typically stay (and for how long); (ii) where has she/he been; or (iii) where will she/he most likely be going to. To answer these questions within this research, it was decided to rely on a person's historical movement data. With respect to computation and response time in case of an emergency, a model was created where stays and transitions of a person are extracted. Based on the information where a person typically stays or moves in-between, this research is able to propose either stays or transition corridors where a person may be at a certain point in time. In this context, the challenge was the identification of a suitable stay detection approach as such a standard selection process is currently missing. Usually, stay detection approaches are evaluated with ground truth data using different measures like recall, precision, and F-measure, as well as entropy and purity [8–13]. In some research additional measures like spatio-temporal accuracy, user's forgetfulness, or user assessment are applied [10,12,13].

A selection framework offering the possibility of using ground truth data as well as data collected in a particular application field by covering the whole process from data acquisition, pre-processing, parameterization to evaluation does not yet exist. Therefore, this paper proposes an evaluation framework for stay detection approaches taking into account the aforementioned aspects. As stochastic and grid-/cell-based approaches have already been investigated in previous work [13,14] spatio-temporal clustering approaches are now investigated here. The proposed framework is validated by comparing three different spatio-temporal clustering approaches using two different datasets, namely ground truth data as well as application data from a field trial conducted with older (partly cognitively impaired) people. To the best of our knowledge, such a comprehensive framework has not yet been published.

The remainder of this paper is structured as follows: in Section 2, related work is reviewed. Section 3 defines general parameters and describes the framework. In Section 4, the application of the framework is presented. The paper closes with a discussion (Section 5) and conclusions drawn from the results (Section 6).

2. Related Work

In the last few years, several stay detection approaches for different positioning systems such as GPS, WLAN, Bluetooth, RFID, or GSM have been developed. Some of these are tailored approaches for a dedicated positioning system and others are positioning system independent. The common aim of all approaches is the analysis and generation of location histories.

2.1. Stay Detection—An Approximation

In the scientific literature, different terms for stay detection are used. In Ashbrook and Starner [15] they refer to it as 'finding significant locations', whereas Hariharan and Toyama [16] or Ye et al. [17] speak about 'extracting/detecting stays/stay points'. Hightower et al. [9], Kang et al. [8], or Changqing, et al. [10] use 'place extraction, identification or discovery', whereas Liu and Seah [4] prefer the term 'location retrieval'. For the terms 'place', 'stay', and 'stay point' clear definitions exist in the literature. A place is a location where an object stays, whereas a 'stay' refers to the event where an object is located at a place for a certain amount of time [4,16]. A 'stay point' is the geographical representation of a 'stay' in form of a quadruple (latitude, longitude, t_{in}, t_{out}), where t_{in} and t_{out} denote the check-in and check-out time at the dedicated location [4]. The time period between t_{in} and t_{out} is defined as 'stay duration'. Thus, in the following, 'stay detection' indicates the approach for detecting stays and 'stay point' is used when the geographical representation of a 'stay' is meant.

2.2. Stay Detection Approaches

According to the literature, different approaches have been developed for the purpose of stay detection from movement data. While originally straightforward and less accurate approaches have been used (e.g., [18]), more advanced stay detection approaches have been implemented to obtain higher accuracy most recently. Approaches for stay detection can be divided into (i) clustering; (ii) grid-/cell-based; and (iii) stochastic methods, that deal with spatio-temporal information. Furthermore, these approaches can be classified with respect to their movement perspective as (i) Eulerian and (ii) Lagrangian approaches. The Eulerian approach concentrates on locations fixed in space and observes objects passing by, whereas the Lagrangian approach adopts the perspective of a moving object, when it moves along a trajectory [19,20]. For this reason, common spatio-temporal clustering and stochastic approaches can be classified as Lagrangian. In contrast, grid-/cell-based approaches can be categorized as Eulerian approaches, in most cases. For positioning systems providing a geographical point as positioning result, e.g., GPS or WLAN (if a position is determined by triangulation), Lagrangian approaches are more common. Systems where a specific area (polygon) or cell is the positioning result, e.g., GSM or Bluetooth beaconing, Eulerian approaches are preferred.

In the literature, clustering approaches for stay detection include (i) partition-based clustering such as k-means [10,11,15]; (ii) hierarchical agglomerative clustering [21]; (iii) time-based/incremental clustering [8,22,23]; (iv) density-based clustering [10,11]; and (v) mixed-method approaches [11,17]. Examples for grid-/cell-based approaches can be found in [13,24,25]. Examples for stochastic approaches are discussed in [14,26,27].

2.3. Data Acquisition

For data acquisition, mobile phones or customary GPS loggers have frequently been used. Across studies, sampling rates have varied between 1 and 60 s. It is notable that the selection of the sampling rate does not depend on the stay detection approach or the expected minimum duration of a stay to be detected, but mainly on the battery life of the collection device.

Many approaches have been evaluated with ground truth data (GTD) collected during corresponding studies, however, with different sample sizes. 'GTD' in this context means that study participants' tag additional information, such as coordinates of a visited place ('tagged place'), as well as the time of entering and leaving the place ('tagged stay'), which results in an annotated GTD set. In some studies data were collected for only some days or hours by one or two persons [8,23], while in other cases at least four persons collected data for several weeks or even months [9,10,12,15]. Some approaches were evaluated using existing, third-party datasets instead of self-generated data [11,17,27]. In contrast to GTD, these third-party or application field data (AFD) are collected without any additional information (uncontrolled recording) by ordinary people, for instance, during a field trial or ordinary system usage. 'AFD' can be differentiated between personalized and anonymous data. In case of movement data, anonymous means that information about the collector is anonymized and/or trajectories are modified so that home or working places cannot be reconstructed. In contrast, personalized data include some information about the collector, like home, work place, and/or even original trajectories.

An evaluation procedure combining GTD and AFD has not yet been applied. Concerning data annotation, there have been differences in GTD as well as in existing datasets. In some cases, people were asked to write diaries [9], while in other studies they were subsequently asked to assess the results and label their locations [12,15,17]. Both diaries and subsequent assessments were also used together by Changqing et al. [10].

2.4. Data Pre-Processing

When using positioning systems like GPS or WLAN, different measurement and interpolation errors influencing the location accuracy have to be considered [28,29]. In such cases, different standard

pre-processing methods like filtering and/or smoothing are used. With respect to stay detection evaluation procedures, a pre-processing step has only been mentioned in [13].

2.5. Parameterization

In the literature, it is often described which parameters are used for a stay detection approach but not why or how these parameters have been selected. In other words, the parameter selection process is rarely described in detail. While in some studies parameter testing was reported [8,10,15], parameter values have only been indicated in other studies without any explanation why specific values were chosen [12,17,23]. General parameters used by almost all approaches are 'stay duration' and 'cluster distance'. While stay duration is defined as the timespan between t_{in} and t_{out}, cluster distance is used to overcome the issue that for the same location different positions can be measured due to location inaccuracy caused by the positioning system (see Section 2.4). Thus, cluster distance defines the maximum distance between two points that belong to the same location [12]. The cluster distance depends on the specific approach.

2.6. Established Evaluation Procedures and Measures

Common measures used for evaluating spatial and/or temporal accuracy of stay detection approaches are recall, precision, and F-measure, as well as entropy and purity [8–11,13,30]. Only in some cases, additional measures are used. In the research by Changqing et al. [10] and Montoliu et al. [12] users had the possibility to assess 'detected stays' determined by the stay detection algorithm. In some cases, users realized that they forgot to indicate stays in the baseline. Montoliu et al. [12] called this users' forgetfulness and introduced a correction factor for precision and F-measure. Runtime or rather computation time has been addressed by Changqing et al. [10], Gröchenig and Schneider [13], Porras Bernárdez [30], Venek et al. [14] as an additional measure. Dedicated spatio-temporal measures have been introduced in Gröchenig and Schneider [13] and Venek et al. [14].

The existing research on stay detection indicates that many approaches are available but dedicated frameworks for evaluating an appropriate approach considering different aspects, including data acquisition, data pre-processing, parameterization, and evaluation measures, do not exist. Additionally, stay detection approaches are usually either tested with GTD or AFD. An evaluation procedure combining both datasets has not yet been considered. Therefore, a comprehensive framework for the evaluation of stay detection approaches is presented and validated next.

3. Definition of the Framework

To our best knowledge, there is no widely accepted assessment method so far for evaluating stay detection approaches. Thus, this paper presents an attempt to provide an iterative framework using ground truth data as well as data collected in a particular application field by covering the whole process from data acquisition, pre-processing, parameterization to evaluation. The stay detection evaluation framework includes five steps: (i) Pre-selection; (ii) acquisition; (iii) pre-processing; (iv) parameterization; and (v) evaluation (Figure 1). Whereby iterations between acquisition and preprocessing (e.g., poor data quality), pre-selection and parameterization (e.g., suitable values for parameters cannot be determined for an approach), and pre-selection and evaluation (e.g., none of the selected approaches has convinced) are possible. Moreover, two different datasets (GTD and AFD) are used within the evaluation process. GTD is required for parameterization and extensive evaluation of the approaches, whereby AFD is used to check their suitability for the intended field of application. In the following, each individual step of the proposed framework, as well as the usage of the two datasets are explained in detail.
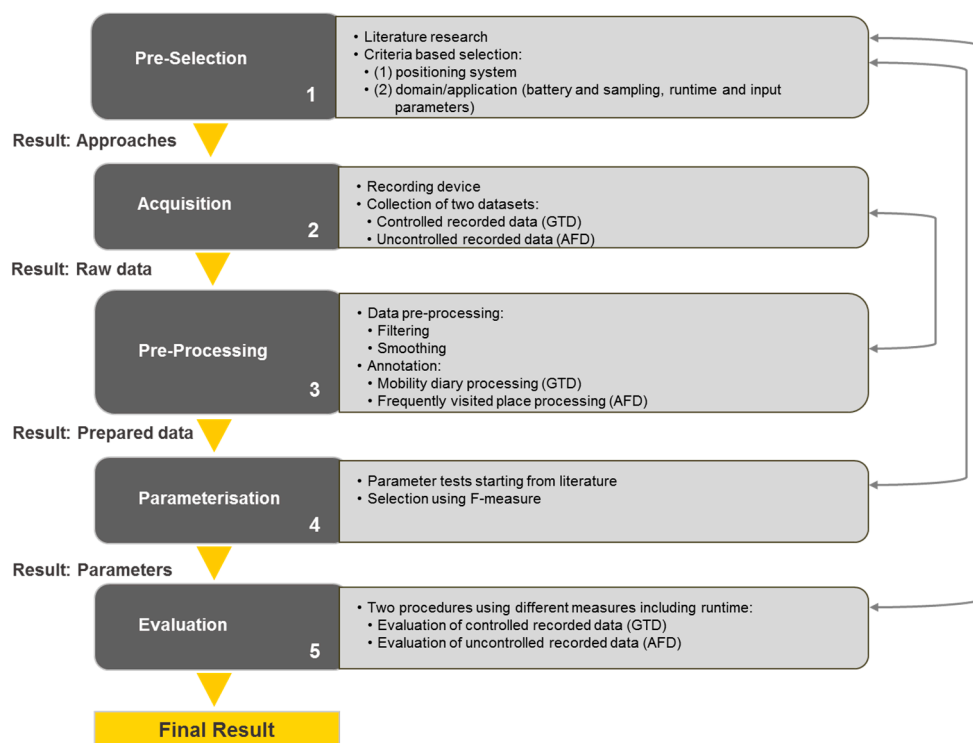
Figure 1. Stay detection evaluation framework.

3.1. Pre-Selection

Pre-selection includes a literature review for identifying relevant approaches using the following two main criteria: (1) Positioning system; and (2) domain/application. Whereby domain/application considers: (i) Battery and sampling; (ii) runtime; and (iii) input parameters of the approach. The selected positioning system (GPS, GSM, WLAN, etc.) should be known, since some approaches have been developed for and/or tested with a dedicated positioning system, like Laasonen et al. [24] and maybe needs to be adapted in order to be used with another positioning system. Additionally, the domain/application need to be considered. Some mobile services, for instance in the AAL domain, require long battery life between 12 and 18 h. Moreover, shape, size and weight of the device are essential [31]. Consequently, lower sampling rates are used and/or recording devices are automatically switched off, if a standstill is detected. In this case, the stay detection approach has to work properly, even with fewer available positions. For emergency applications, runtime also plays an important role. With respect to mobile AAL services, time-aware processing of the data may additionally be feasible. Moreover, it needs to be taken into account that for some stay detection approaches an a priori knowledge of specific input parameters is needed, such as the (expected) number of visited places (clusters), in case of the k-means method. Therefore, required input parameters should be comprehensible and, if required, independent of a priori knowledge.

3.2. Acquisition

For the proposed framework, data acquisition is an important step that involves deliberations about the recording device and the recording type. Regarding the recording device, it should at least enable adjustments of the setting dependent configurations, namely the standstill detection for switching the device off and setting the appropriate sampling rate. Furthermore, the device should support the sampling rate required by the intended application. For example, higher sampling rates are required when (additionally) analyzing different movement scenarios (e.g., normal or fast walking, slow or fast direction changes, short or long stops) compared to extracting stays

from trajectories (of pedestrians) [29]. The recording type provides information about the acquired data that can be categorized as either controlled or uncontrolled. 'Controlled recorded data', like GTD, consist of recorded positioning data and metadata, such as a diary of stays. In contrast, 'uncontrolled recorded data' like AFD include only recorded positioning data without any additional information. If uncontrolled recorded data are used, then they have to be personalized for the evaluation of stay detection approaches (e.g., frequently visited places of the collectors are disclosed and tagged). This additional information can then be used within the framework for evaluation purposes. For enabling parameterization, extensive evaluation, and the examination of the suitability for the intended field of application, the data acquisition procedure should collect both datasets, including ground truth and application field datasets.

3.2.1. Collection of Controlled Recorded GTD

The following procedure for collecting GTD is proposed, device selection, identification of trial participants (data collectors), and the definition of rules for data recording and documenting mobility. For collecting GTD a recording device, being able to record at least one day with a high sampling rate of less than 5 s [29], needs to be chosen. Additionally, four persons (see Section 2.3) with different mobility behavior (means of transportation, activity level with respect to their number of visited places) willing to track their daily movements and document them in a mobility diary for at least four weeks have to be identified. To ensure optimal recording, the following rules should be followed: (i) weekly almanac update; (ii) daily charging of the device; (iii) awaiting GPS signal after leaving a building; (iv) downloading data from recording device, if necessary; and (vi) setting a morning reminder to use the device. Moreover, each data collector has to document all visits to places lasting longer than 15 min, reporting the exact location, as well as the check-in and check-out time.

3.2.2. Collection of Uncontrolled Recorded AFD

Since for the collection of (application dependent) uncontrolled recorded data (AFD) none rules exist, the following tasks are suggested: (i) device selection; (ii) identification of trial participants (data collectors); and (iii) definition of setting reminders. An appropriate recording device for the intended application, being able to record at least one day with the stated sampling rate without user interaction (automatic background recording), has to be identified. Moreover, 15 to 30 trial participants within a dedicated field of application, willing to be tracked for several weeks and providing their permission by signing an informed consent form, have to be recruited. To ensure that data collectors do not forget to charge and carry the device, daily reminders need to be defined, e.g., on the collectors' smartphones.

3.3. Pre-Processing

Pre-processing is particularly important (1) for coping with measurement and interpolation errors caused by the positioning system (see Section 2.4); and (2) for data annotation, which can be classified into annotating controlled and uncontrolled recorded data.

3.3.1. GPS Data Pre-Processing

For GPS data pre-processing a two-step pre-processing approach following Gröchenig and Schneider [13] is proposed. In a first step, duplicated track points (both time stamp and coordinates) are removed. Additionally, track points before and after tunnels (or other tracking interruptions) and outliers are filtered out. In a subsequent step, data are smoothed using a triangular kernel function to determine weighting factors in order to produce more natural tracks and reduce spikes.

3.3.2. Data Annotation

When annotating controlled and uncontrolled recorded data the following two procedures are proposed: (1) mobility diary processing and (2) frequently visited place processing. The procedure for

frequently visited places only works, if at least one frequently visited place is known (e.g., address or coordinates).

Mobility diary processing of controlled recorded GTD: In a first step, the total amount of visited places has to be determined and each place has to be labeled with a unique ID for all trial participants. Thereafter, each participant needs to identify Open Street Map (OSM) elements (node, way, or relationship) of all visited places. To ensure that places are tagged identically, the coordinates are set at the centroid of each OSM element. Subsequently, for each stay lasting at least 15 min, a tagged place with check-in and check-out times is created.

Frequently visited places processing of uncontrolled recorded AFD: First, a frequently visited place (e.g., known home or working place) is identified using geocoding or reverse geocoding methods. Next, the OSM coordinates of this place are determined as described above for GTD. Thereafter, stays lasting for at least 15 min at the frequently visited place are identified and tagged.

3.4. Parameterization

Besides general parameters like minimum stay duration or cluster distance, individual parameters are essential for each approach to work properly. As already explained in Section 2.5, the parameter selection process is not explained in many cases. As a pre-defined parameterization process does not exist, the following process using GTD is proposed. After implementing the approach to be investigated, parameter selection starts with parameter values suggested in the literature. Lower and higher values for each parameter will be tested until results worsen (number of detected and overlapping stays compared to ground truth). Subsequently, the F-measure will be calculated for identified parameter sets, which will then be ranked according to their F-measure. Finally, parameters with the highest F-measure of all datasets will be selected.

3.5. Evaluation

The most commonly used evaluation measures are precision and recall (see Section 2.6). Additional measures are rarely used. Spatio-temporal measures have been introduced by Gröchenig and Schneider [13] and Venek et al. [14]. As time-aware computation plays an important role for many application fields, runtime should additionally be considered. Since two different datasets can be evaluated, a distinction needs to be made between the evaluation of controlled and uncontrolled recorded data. Thus, two different evaluation strategies were developed. In contrast to other work [10,12,15,17] and regarding our intended application, the proposed evaluation procedure does not consider subjective assessment of results by data collectors.

For both GTD and AFD datasets, a spatio-temporal evaluation approach that additionally considers runtime, is proposed. Due to known limitations of positioning systems (see Section 2.4), exact overlaps of tagged stay points and detected stay points will hardly occur. Therefore, for both datasets a distance and time threshold is introduced. Whereby the 'distance threshold' is the spatial and the 'time threshold' the temporal tolerance between a tagged and detected stay point. If the distance between a tagged and a detected place exceeds either the distance or the time threshold, the detected place is not linked to a tagged place. As proposed by Gröchenig and Schneider [13], the value of the distance threshold is set to 53 m (considering the GPS error and the average diameter of a tagged place) and the value of the time threshold to 15 min. Considering the fact that no interest exists in short stops, like stays at an intersection before crossing it and that medium stay times like shopping, are defined with 13 to 15 min [32], the minimum stay duration is also set to 15 min.
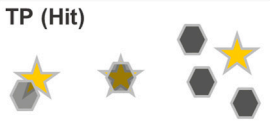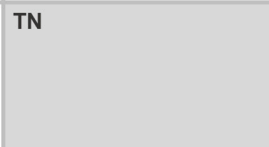
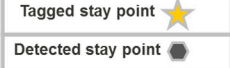#### 3.5.1. Evaluation of Controlled Recorded GTD

In a first step, five possible results of the application of stay detection approaches to GTD are defined. Results can be that (1) a stay point is detected, where there is none; (2) the tagged stay point is not detected; (3) the tagged stay point is detected and represented by more than one detected stay

point; (4) the detected stay point can be assigned to more than one tagged stay point; and (5) the tagged stay point is detected.

In the following, three different evaluation measures, namely the confusion matrix, spatio-temporal quality measures, and runtime, are defined and quantified to determine these five results.

Confusion matrix: Generally, a confusion matrix consists of four measures, namely true positive (TP), true negative (TN), false positive (FP), and false negative (FN) [33]. As a typical two-class problem is not investigated here, the confusion matrix is applied without a TN (see Figure 2). Regarding stay detection, a TN represents all locations where neither detected nor tagged stay points occur. The illustrated confusion matrix captures all identified results except a detected stay point with multiple tagged stay points since a TP depends on the number of correctly identified tagged stay points.

Figure 2. Confusion matrix without true negatives (TN).

Based on the confusion matrix, recall, precision, and F-measure [33] are determined with respect to detected and tagged stay points. To determine thresholds for these measures, three main issues need to be considered, including that (1) GPS datasets can be error-prone even after pre-processing; (2) if space and time are considered; there can be deviations in one or both dimensions; and (3) collectors can make mistakes during data annotation. Each of these issues bias the results. For instance, the approach might detect more than one stay, if GPS measurement errors, which occurred more frequently during a single tagged stay, remain even after pre-processing. Experiences have shown [13,14] that values for recall, precision, and F-measure > 0.65 seem appropriate to determine an appropriate stay detection approach. Thus, thresholds are set to 0.65.

Several measures for GTD, to evaluate spatial and temporal accuracy of stay detection approaches, have been developed [13,14]. However, the following two have proven to be useful: (1) Spatial accuracy; and (2) the degree of spatio-temporal correctness.

1. Spatial accuracy Q_{sa} measures the quality of the spatial detection accuracy with respect to tagged places. This measure depends on the degree of detection correctness q_t (see Formula (1)) for each tagged place, which is computed by using the mean Euclidean distance between each tagged and corresponding assigned detected place(s) $\bar{d}_t$ and the distance threshold denoted by r. If the mean distance of all corresponding detected places is smaller than half the distance threshold, the tagged place is considered as correctly detected. If the mean distance lies between half and the actual distance threshold, the detection correctness degree decreases starting from 0.5 by $1/r$. Otherwise and if the tagged place is not recognized by any detected places, the tagged place is not detected correctly. The sum of degrees of correctness of each tagged place divided by the total number of tagged places T

of the GTD gives the accuracy of the correctly assigned detected places (see Formula (2)). This means that detected places without associated tagged places are not considered. A value of 1.0 indicates that locations of all tagged and assigned detected places are identical. A value of 0.0 indicates that all detected places do not meet the distance threshold criteria [13,14]. Spatial accuracy of sufficiently performing stay detection approaches should be more than 0.6.

$$q_t := \begin{cases} 1 & \text{if } \overline{d}_t \leq \frac{r}{2} \\ 1 - \frac{\overline{d}_t}{r} & \text{if } \frac{r}{2} < \overline{d}_t \leq r \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$Q_{sa} := \frac{\sum_t^T q_t}{T} \tag{2}$$

2. The degree of spatio-temporal correctness Q_{stc} measures the amount of correctly detected stay points regarding time information of all D spatially assigned detected stay points. Therefore, the mean of the time difference from tagged and corresponding detected stay point $\Delta t_{t,d}$ is determined by using entry and exit times (see Formula (3)). D divides the number N_{corr} of detected stay points, which are recognized as correctly detected, when they are within the time threshold $\Delta t_{t,d} \leq 900$ s (see Formula (4)). This determines the proportion of temporally correct detected stay points. A value of 1.0 indicates that all detected stay points correspond to assigned tagged stay points, while a value of 0.0 means none of the detected stay points corresponds to time entries in the GTD [13]. Regarding adequate stay detection, the degree of spatio-temporal correctness should exceed a value of 0.65.

$$\Delta t_{t,d} = \frac{\left| (t_{entry,d} - t_{entry,\,t}) \right| + \left| (t_{exit,\,d} - t_{exit,\,t}) \right|}{2} \tag{3}$$

$$Q_{stc} := \frac{N_{corr}}{D} \tag{4}$$

The runtime is measured without pre-processing steps in order to only assess the duration of the stay detection. A fixed threshold for runtime cannot be specified, as it depends on the amount of data. For about 500,000 track points, a value of ≤ 5 s is suggested.

3.5.2. Evaluation of Controlled Recorded AFD

In contrast to controlled recorded data, visited places are only partially known and, thus, an extensive evaluation is not possible. Hence, at least a single frequently visited place like home or work place should be known, so that they can be assigned as tagged stay points. The resulting tagged stay points are then compared with detected stay points. The runtime for AFD is measured in the same way as for controlled recorded data.

For evaluating spatio-temporal accuracy of AFD two ratios considering distance and time are used, namely the spatial- and spatio-temporal ratio.

1. Spatial ratio R_s refers to the number of detected places N_{dp} that are within the distance threshold divided by the number of tagged places N_{tp} (see Formula (5)). If the ratio is higher than 1.0 more places are detected than tagged; a value lower than 1.0 indicates that fewer places are detected than tagged. The value of the spatial ratio should be between 0.75 and 1.25.

$$R_s = \frac{N_{dp}}{N_{tp}} \tag{5}$$

2. Spatio-temporal ratio R_{st} refers to the number of spatio-temporal overlapping detected stay points N_{dsp} meeting the distance and the time threshold criteria divided by tagged stay points N_{tsp} (see Formula (6)). A value of 1.0 indicates that for all tagged stay points overlapping detected stay

points were found. A value of 0.0 indicates that for no tagged stay point an overlapping detected stay point was found. For evaluation purposes, the spatio-temporal ratio should exceed 0.65.

$$R_{st} = \frac{N_{dsp}}{N_{tsp}} \qquad (6)$$

In summary, the proposed stay detection evaluation framework consists of five steps: (i) Pre-selection; (ii) acquisition; (iii) pre-processing; (iv) parameterization; and (v) evaluation. Besides general steps, which are the same for all approaches, the evaluation framework also allows adaptations for taking into account domain- or application-specific features (pre-selection), as well as requirements that depend on the selected method (parameterization) (see Table 1).

Table 1. General and adaptable parts of the framework.

Framework Step	General	Domain/Application Specific	Approach Dependent
Pre-selection		X	
Acquisition	X		
Pre-processing	X		
Parameterization			X
Evaluation	X		

4. Application and Validation of the Framework

In this section, the proposed framework is applied and validated with GTD and AFD collected during a field trial with older people testing a mobile AAL system. In addition, the aim is to assess spatio-temporal approaches in this project, including stochastic and grid-/cell-based approaches [13,14] that have already been investigated in previous works.

4.1. Pre-Selection

As suggested by the pre-selection procedure of the framework, a literature search was conducted and three spatio-temporal approaches were identified. It is important that for the research in this application (1) GPS data are used; (2) a battery life of $\geq$12 h is expected; (3) time-aware runtime is required; and (4) approaches where an a priori knowledge of specific input parameters (e.g., number of clusters) is not needed are preferred. Partition-based and hierarchical agglomerative approaches were not considered for validation, since for the first approach the number of clusters has to be known in advance and for the second approach the runtime is poor [34]. In the following, the three selected spatio-temporal approaches are briefly presented considering the positioning system, battery life, sampling rate, runtime, and input parameters.

4.1.1. Time-Based/Incremental Clustering (TIC)

Time-Based/Incremental Clustering (TIC) by Kang et al. [8] is a positioning system independent approach. The positioning system used by Kang et al. [8] in Place Lab [35]. This approach was tested in this research using WLAN and GSM infrastructure. Additionally, the authors deliberately developed TIC to run on resource-limited (mobile) devices as a background task. To save battery, TIC is able to deal with lower sampling rates. However, it has to be noted that stays lasting exactly the minimum stay duration may not be recognized with a sampling rate larger than 1 s. Since this issue is expected to occur rarely, also due to known limitations of positioning systems, it is neglected.

The aim of TIC is the automatic extraction of stays. For spatio-temporal stay detection, Kang et al. [8] made several considerations for the following parameters: (i) minimum duration of a stay at a certain location; (ii) minimum distance between stays; and (iii) a time parameter for detecting outliers and determining whether an object is moving or not. The approach is able to extract 'relevant' stays autonomously and ignores 'irrelevant' stays, such as stops at traffic lights from GPS

traces. TIC works incrementally. If a detected stay is within a defined distance to an existing detected stay, the two stays are merged.

The implementation used in this paper is similar to the one developed by Kang et al. [8] with one exception, namely the option of considering short 'relevant' stays of two or three minutes. This was not considered, since the proposed framework defines a fixed minimum stay duration to apply to all approaches (see Section 3.5).

4.1.2. (Extended) Density-Based Clustering (xDBSCAN)

Density-based clustering approaches such as DBSCAN [36] or OPTICS [37] do not consider spatial and temporal parameters simultaneously, except ST-DBSCAN [38]. However, the ST-DBSCAN requires the fulfilment of spatial and temporal clustering criteria at the same time. Otherwise, no clusters are created. This means that if a GPS receiver loses its signal within a building, ST-DBSCAN will never generate a cluster. This is because the two GPS points for entering and leaving the building are not sufficient to meet the minimum number of points used under normal conditions. In contrast, when using DBSCAN (that only considers the spatial criterion), a cluster can be generated, if the signal loss occurs at the same place for several times. Thus, a two-step positioning system independent approach using DBSCAN with its usual parameters (ε, *minPts*) and a time extraction (xDBSCAN) was implemented. For battery savings, xDBSCAN is able to deal with lower sampling rates whereby the minimum number of points *minPts* has to be redefined. A place is detected, if *minPts* are within the maximum distance ε in a DBSCAN cluster.

In a first step, DBSCAN creates spatial clusters and in a second step, the trajectories within these clusters are cut for generating stay points with entry and exit times. Furthermore, clusters that occur at path intersections are eliminated because of the minimum stay duration parameter.

In this paper the Apache DBSCAN implementation was used [39] for the first step. OPTICS would have been an alternative to DBSCAN. Since time-awareness and, thus, runtime are important in the proposed framework, it was decided to use DBSCAN. In contrast to OPTICS, there are possibilities to enhance DBSCAN in case runtime issues occur [40–42]. The second step of this approach was implemented by us. In this step, stay points are extracted from clusters.

4.1.3. Mixed-Method Approach (MIXED)

The MIXED approach is a two-step approach, tested with GPS data and introduced by Ye et al. [17] within the GeoLife project [43]. Regarding battery savings and sampling rate, the approach is similar to TIC. In contrast to Kang et al. [8], Ye et al. [17] did not merge clustered stays directly, but performed this in a second step using OPTICS. Since this approach is similar to TIC in many respects, a similar runtime is also expected. The main parameters used within this approach are the cluster distance and the minimum stay duration.

The first step of the implementation used in this paper is equal to the first step of Ye et al. (2009). For the second step, DBSCAN instead of OPTICS was used, due to the reasons explained above.

4.2. Acquisition

To test the entire framework, two datasets, including GTD and AFD, were collected.

4.2.1. Controlled Recorded GTD

It was decided to apply QStarz GPS loggers for recording good quality data (2 × GPS Travel Recorder BT-Q1000X (QStarz, Taipei, Taiwan) and 2 × GPS Data Recorder CR-Q1100P (QStarz, Taipei, Taiwan)). The same receivers are often used for mobility tracking [28,29,44]. Furthermore, these loggers are able to record single days at high sampling rates. A sampling rate of 3 s was used in this research. Due to the sufficient battery capacity of loggers, the standstill detection was not activated. Four trial participants gave their informed consent for tracking and documenting their mobility for more than five weeks. This is according to the rules stated for controlled data recording (see Table 2).

Table 2. Details of ground truth datasets for four trial participants.

Property	Person A	Person B	Person C	Person D
Days of recording	39	40	41	38
Number of track points	418,095	489,131	436,958	415,104
Main mode of transportation	walking, car	train, walking	cycling	car
Unique visited places	27	35	43	26
Recorded trips between places	131	109	157	104

4.2.2. Uncontrolled Recorded AFD

The data were collected within two six-week field trials testing a mobile AAL system for older (partly cognitively impaired) people. From these field trials, datasets of 31 informed participants could be used for the study. Since field trials aimed at hardware testing, as well, different devices (Laipac S911 Bracelet Locator HC (Laipac Technology Inc., Richmond Hill, Canada), Google Nexus 5 (Google Inc., Menlo Park, CA, USA), Samsung Xcover 1 and 2 (Samsung Electronics (UK) Ltd., Suwon, Korea)) with different sampling rates (5 to 15 s) were used. Devices recorded data automatically in the background and used standstill detection to save battery life. Thus, fewer track points were recorded compared to GTD. Despite daily reminders for wearing and charging, the planned 42 tracked test days per trial participant were not achieved. On average, only 18.5 days' worth of data were collected. The main reasons were that (i) participants turned on the device only when they went outside; (ii) they became ill; (iii) they felt uncomfortable while wearing the device; or (iv) they forgot to charge it. While eight trial participants did always stay near their home, the others left their home more or less regularly. In total, this resulted in 153 home stays lasting at least 15 min.

4.3. Pre-Processing

In this subsection, the framework addresses movement data pre-processing and processing of data annotation. Data pre-processing was applied to both GTD and AFD datasets in the same way. With respect to data annotation, the framework differentiates between controlled and uncontrolled recorded data (see Section 3.3.2). In case of controlled recorded GTD, the total amount of visited places was determined and each unique place was labelled with an ID, for each mobility diary. For each place, data collectors were asked to identify corresponding OSM elements in order to retrieve locations of tagged places. Regarding the uncontrolled recorded AFD, the home address of each trial participant was known and set as a frequently visited place. Stays at home that took at least 15 min were identified and set as tagged stay points.

4.4. Parameterization

Parameterization is essential for each approach to work properly. The individual parameter selection process for all approaches was discussed in detail in the parameterization procedure of the proposed framework. General parameters, such as minimum stay duration, are the same for all approaches. Different parameter values, starting with values mentioned in the literature were tested, applied to the GTD, and selected with respect to the obtained F-measures.

4.4.1. TIC

Parameters for this approach are *cluster distance, merging distance, minimum stay duration*, and *moving time* for detecting outliers and determining whether an object is moving or not. For all parameters, except *minimum stay duration*, which is already set to 15 min, suitable values needed to be determined. Kang et al. [8] proposed *cluster distance* values between 30 and 50 m and Hightower et al. [9] of 300 m. Thus, *cluster distance* values between 25 and 350 m, in increments of five meters, were tested. Kang et al. [8] specified the *merging distance* with one-third of the *cluster distance*. First tests in

our research indicated that one-third of the *cluster distance* may not be sufficient for GPS data, thus, the *merging distance* was extended to two-thirds of the *cluster distance* and additionally tested.

Since no details were given in the literature for the *moving time* parameter, values between 10 and 120 s in increments of ten were tested. The *moving time* parameter was determined when testing *cluster distances* between 30 and 50 m and *merging distances* between 7 and 33 m. When comparing the resulting F-measures, values between 80 and 100 s scored highest, hence, we selected 90 s for this parameter.

The final *cluster distance* values were applied to the parameterization procedure (see Section 3.4). For the GTD, a *cluster distance* between 90 and 140 m and a *merging distance* of two-thirds the *cluster distance* achieved the best F-measures. The final combination selected was a *cluster distance* value of 125 m and a corresponding *merging distance* of 83 m.

4.4.2. xDBSCAN

In the literature, parameters ε (*cluster distance*) and *minPts* of DBSCAN are often selected by applying OPTICS to the data. The example of an OPTICS diagram applied to GTD is shown in Figure 3. Low frequencies next to each other represent GPS points that are spatially located near each other and indicate clusters. Figure 3 includes one large cluster on the left-hand side, with several smaller clusters to be recognized on the right-hand side. While the height of frequencies can be used to determine a proper value for ε, *minPts* can be derived from the number of frequencies within an identified cluster. The first significant peak (encircled in Figure 3) corresponds to a low GPS point density, identifying points outside a cluster. With a height of 0.45 (WGS84 degrees $\times$ 10,000), the ε corresponds to approximately 4–5 m. For ε, values between 0.01 and 80 m were tested. From the OPTICS diagram, the cluster with the lowest number of points can be used to determine a reference value for *minPts*, which is 50. Thus, *minPts* values between 3 and 800 were evaluated. For all combinations of the two parameters (ε and *minPts*), F-measures were compared and acceptable results were achieved for ε between 1 and 40 m, and for *minPts* between 100 and 400. Finally, 1 m for ε and 200 for minPts were used.
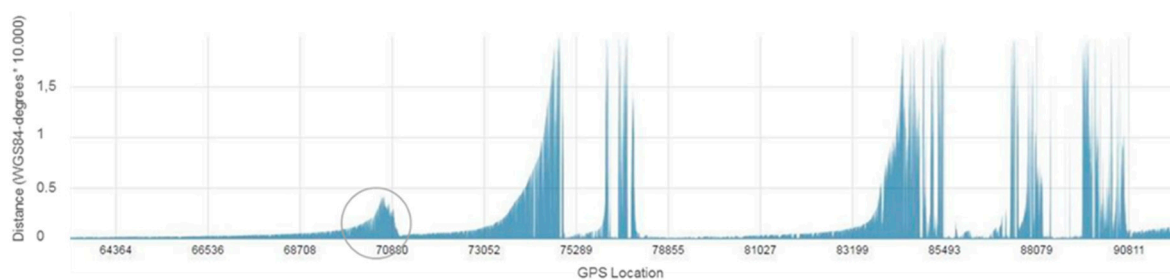

Figure 3. OPTICS diagram with the first significant peak encircled.

4.4.3. MIXED

The first step of the MIXED approach relies on two main parameters, namely *cluster distance* and *minimum stay duration*. Ye et al. [17] set the *cluster distance* to 200 m. According to the literature and our experiences with TIC, distance values between 25 and 350 m in increments of five meters were tested and F-measures calculated. Based on results, it was determined that a *cluster distance* value between 175 and 250 m would be appropriate. The finally selected *cluster distance* measured 200 m, which is also in accordance with values stated in the literature.

Since DBSCAN is used for the merging step of the MIXED approach, parameters for ε and *minPts* needed to be determined, as well. It was decided to set ε to the value of the distance threshold, namely 53 m. To mind a places visited only once, the parameter *minPts* was set to 1.

4.5. Evaluation

As described by evaluation procedures for controlled and uncontrolled recorded datasets, the corresponding values were determined. Runtime for evaluation purposes was measured on a Lenovo T450s with 2.6 GHz, 12 GB RAM and SSD storage.

Table 3 compares measurement values for assessing controlled recorded GTD. Results show that most evaluation measures of both TIC and MIXED exceed defined thresholds. Runtimes of both approaches are almost equal and there are some variations in the other measures over the four trial participants. The worst results are achieved by using xDBSCAN, where most values are below the defined thresholds.

Table 3. Ground truth data (GTD) evaluation results for Time-Based/Incremental Clustering (TIC), xDBSCAN, and MIXED (bold values indicate that thresholds are met).

Approach	Person A (39 Days, 418,095 Points)	Person B (40 Days, 489,131 Points)	Person C (41 Days, 436,958 Points)	Person D (38 Days, 415,104 Points)
TIC (125 m, 83 m)				
Recall (0.65)	**0.71**	**0.65**	0.45	**0.73**
Precision (0.65)	**0.70**	0.49	0.38	0.52
F-measure (0.65)	**0.71**	0.56	0.41	0.61
Q_{sa} (0.6)	**0.65**	**0.61**	0.56	**0.60**
Q_{stc} (0.65)	**0.79**	**0.72**	0.48	**0.78**
Runtime (5 (s))	**0.28**	**0.23**	**0.20**	**0.18**
xDBSCAN ($\varepsilon = 1$, $minPts = 200$)				
Recall (0.65)	0.37	0.29	0.43	0.59
Precision (0.65)	**0.76**	0.57	**0.71**	**0.69**
F-measure (0.65)	0.50	0.39	0.54	0.64
Q_{sa} (0.6)	0.46	0.26	0.24	0.31
Q_{stc} (0.65)	0.60	0.44	0.62	**0.71**
Runtime (5 (s))	45,105.38	65,665.74	49,194.14	58,378.40
MIXED (200 m)				
Recall (0.65)	**0.70**	**0.65**	**0.72**	**0.74**
Precision (0.65)	**0.70**	0.50	0.62	0.54
F-measure (0.65)	**0.70**	0.57	**0.67**	0.62
Q_{sa} (0.6)	0.56	0.60	0.56	0.54
Q_{stc} (0.65)	**0.77**	**0.69**	**0.80**	**0.83**
Runtime (5 (s))	**0.28**	**0.24**	**0.24**	**0.21**

AFD evaluation results (Table 4) are very similar to GTD results (Table 3) with respect to xDBSCAN, particularly considering runtime when processing more than 10,000 points. Evaluation results of TIC and MIXED show a higher variability for AFD than for GTD, with TIC producing better results compared to MIXED. In general, results for the entire evaluation procedure of the framework indicate that TIC meets threshold criteria and runtime performance more often with both datasets (GTD and AFD) than MIXED.

Table 4. Application field data (AFD) evaluation results for TIC, xDBSCAN, and MIXED (bold values indicate that thresholds are met).

Record	TIC (125 m, 83 m)			xDBSCAN ($\varepsilon = 10$, $minPts = 400$)			MIXED (200 m)		
	R_s (0.75 − 1.25)	R_{st} (0.65)	*Runtime* (5 (s))	R_s (0.75 − 1.25)	R_{st} (0.65)	*Runtime* (5 (s))	R_s (0.75 − 1.25)	R_{st} (0.65)	*Runtime* (5 (s))
Participant 1 (22 days, 3743 points)	0.00	0.00	**0.00**	0.50	0.00	3.65	0.00	0.00	**0.00**
Participant 2 (5 days, 10,731 points)	**1.00**	**1.00**	**0.01**	**1.00**	0.33	27.16	**1.00**	**1.00**	**0.01**
Participant 3 (17 days, 25,837 points)	**0.91**	0.64	**0.02**	0.73	0.37	157.46	**1.00**	**0.82**	**0.02**

Table 4. *Cont.*

Record	TIC (125 m, 83 m)			xDBSCAN ($\varepsilon = 10$, *minPts* = 400)			MIXED (200 m)		
	R_s (0.75 − 1.25)	R_{st} (0.65)	*Runtime* (5 (s))	R_s (0.75 − 1.25)	R_{st} (0.65)	*Runtime* (5 (s))	R_s (0.75 − 1.25)	R_{st} (0.65)	*Runtime* (5 (s))
Participant 4 (10 days, 8225)	**1.00**	**1.00**	**0.00**	0.00	0.00	17.33	1.50	0.50	**0.01**
Participant 5 (34 days, 25,490 points)	**1.06**	**0.88**	**0.02**	**0.94**	0.41	87.69	1.47	0.47	**0.02**
Participant 6 (14 days, 10,013 points)	**1.11**	**0.67**	**0.01**	0.67	0.11	19.54	**1.00**	0.44	**0.00**
Participant 7 (17 days, 9426 points)	**1.00**	**1.00**	**0.00**	**1.25**	**0.75**	22.66	**1.25**	**0.75**	**0.01**
Participant 8 (36 days, 25,549)	**1.00**	**1.00**	**0.01**	**1.00**	0.00	158.65	2.00	0.00	**0.01**
Participant 9 (39 days, 42,083 points)	**1.00**	**1.00**	**0.02**	**1.00**	0.60	397.10	**1.00**	**1.00**	**0.02**
Participant 10 (7 days, 3518 points)	**1.00**	**1.00**	**0.00**	0.00	0.00	3.19	0.50	0.50	**0.00**
Participant 11 (11 days, 13,274 points)	**1.00**	**1.00**	**0.01**	**0.75**	0.50	30.74	**1.00**	0.50	**0.01**
Participant 12 (10 days, 1801 points)	**1.00**	**1.00**	**0.00**	0.00	0.00	0.80	1.60	0.60	**0.00**
Participant 13 (18 days, 43,402 points)	**0.75**	0.63	**0.02**	**0.75**	0.13	464.51	**0.75**	0.25	**0.02**
Participant 14 (7 days, 6076 points)	**1.00**	**1.00**	**0.00**	0.00	0.00	9.10	**1.00**	**1.00**	**0.00**
Participant 15 (18 days, 55,591 points)	**0.80**	0.40	**0.02**	0.40	0.00	793.90	0.20	0.00	**0.03**
Participant 16 (35 days, 77,304 points)	**1.00**	**1.00**	**0.03**	**1.00**	**1.00**	1510.00	**1.00**	**1.00**	**0.04**
Participant 17 (21 days, 1529 points)	0.43	0.14	**0.00**	0.00	0.00	0.55	0.14	0.00	**0.00**
Participant 18 (5 days, 14,959 points)	**1.00**	**1.00**	**0.01**	0.00	0.00	56.43	**1.00**	**1.00**	**0.01**
Participant 19 (39 days, 98,866 points)	**1.00**	**1.00**	**0.04**	**1.00**	**1.00**	2425.15	**1.00**	**0.80**	**0.05**
Participant 20 (28 days, 16,454 points)	**1.00**	**1.00**	**0.01**	**1.00**	**1.00**	63.81	**1.00**	**1.00**	**0.01**
Participant 21 (3 days, 850 points)	0.57	0.29	**0.00**	0.00	0.00	0.18	0.43	0.14	**0.00**
Participant 22 (7 days, 1064 points)	0.00	0.00	**0.00**	0.00	0.00	0.28	0.67	**0.67**	**0.00**
Participant 23 (10 days, 18,916 points)	**1.00**	**1.00**	**0.01**	1.29	0.57	90.02	**1.00**	**1.00**	**0.01**
Participant 24 (3 days, 126 points)	**1.00**	**1.00**	**0.00**	0.00	0.00	0.01	2.00	0.00	**0.00**
Participant 25 (40 days, 52,314)	**1.00**	**1.00**	**0.02**	**1.00**	**1.00**	642.63	**1.00**	**1.00**	**0.03**
Participant 26 (25 days, 140,612 points)	**0.82**	0.64	**0.06**	0.55	0.27	4428.31	**0.91**	0.55	**0.07**
Participant 27 (35 days, 11,577 points)	**1.00**	**1.00**	**0.01**	**1.00**	0.00	32.44	1.00	1.00	**0.01**
Participant 28 (11 days, 10,588 points)	**1.00**	**1.00**	**0.00**	0.00	0.00	28.23	1.00	1.00	**0.00**
Participant 29 (13 days, 3536 points)	**1.00**	**1.00**	**0.00**	0.00	0.00	2.86	**1.17**	0.50	**0.00**
Participant 30 (28 days, 33,379 points)	**1.00**	**0.67**	**0.01**	0.67	0.50	256.46	0.17	0.00	**0.02**
Participant 31 (8 days, 157 points)	0.00	0.00	**0.00**	0.00	0.00	0.01	0.00	0.00	**0.00**

5. Discussion

Other than the development, application, and validation of the proposed framework for stay detection approaches, the performance of three different spatio-temporal clustering approaches for mobile AAL systems for older people was assessed. Based on the evaluation results of the proposed framework (see Tables 3 and 4), the xDBSCAN is clearly not appropriate for the intended application, mainly due to the poor runtime. Additionally, no preference between TIC and MIXED can be made solely based on GTD, since neither of the two approaches show a convincing performance among all proposed measures. For this reason, the comprehensive evaluation of the AFD included in the framework is justified.

Compared to Hightower et al. [9], the TIC performed better than the other two approaches considering the entire framework. However, the TIC had significantly more problems with the data of person C than MIXED. This can be explained by the fact that C lives next to a mountain so that C is often localized more than the cluster distance of 125 m away from the tagged home place. This is not a problem for MIXED due to a cluster distance of 200 m. Considering both datasets, the cluster distance of 125 m for TIC works fine. This justifies the use of two different datasets. However, an advantage of the MIXED approach is the lower number of parameters. Compared to the grid-/cell-based cookie-cutter approach [13], the TIC approach is nearly similar regarding quality measures but superior regarding runtime. Thus, the TIC approach will be selected for future field operations.

Regarding the validation of the proposed framework, all procedures were successfully applied to each of the spatio-temporal clustering approaches for determining the most appropriate approach. In contrast to other work [10,12,15,17], a subjective result assessment by data collectors in the proposed framework was not considered. However, if an assessment of the controlled recorded data by collectors is possible, it can be easily integrated into the evaluation process by defining a corresponding evaluation procedure.

In contrast to Montoliu et al. [12], users' forgetfulness during the tagging process was also not considered. Nevertheless, the combination of GPS measurement errors and the consideration of both space and time, address this issue when setting evaluation thresholds for evaluation measures.

6. Conclusions

In this paper, a comprehensive framework for evaluating stay detection approaches is presented. In comparison to previous work [10,12], this framework covers the entire process from pre-selection, data acquisition, pre-processing, parameterization, to evaluation. Furthermore, the framework offers the possibility to use both ground truth as well as application field data for the intended field of application. In addition to known evaluation measures, spatio-temporal quality measures are applied (Q_{sa} and Q_{tc} for GTD and R_s and R_{st} for AFD). Since each step of the framework is discussed in detail in this paper, it should be possible to apply this framework to other stay detection approaches and datasets, no matter whether they are positioning system independent or not. Even framework extensions are possible, for example, the integration of data assessment through data collectors. Moreover, the use of two different datasets has proven to be justified and effective. The selection of a suitable approach would have been challenging using GTD only. With the inclusion of both datasets (GTD and AFD), the TIC was identified as the most suitable approach for our intended field of application and will be used in our future projects. This should make it possible to propose either stays or transition corridors where a person may be at a certain daytime to rescuers in case of an emergency (e.g., a person is lost without a GPS device).

Author Contributions: Cornelia Schneider, Simon Gröchenig, and Verena Venek designed the framework. The framework was reviewed by Michael Leitner and Siegfried Reich. Cornelia Schneider contributed to data collection. Data annotation was done by Cornelia Schneider, Simon Gröchenig, and Verena Venek. Simon Gröchenig and Cornelia Schneider implemented the stay detection approaches. Simon Gröchenig, Verena Venek, and Cornelia Schneider implemented the evaluation measures. Data analysis and interpretation was done by Cornelia Schneider, Verena Venek, Simon Gröchenig, Michael Leitner, and Siegfried Reich.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Castelli, G.; Mamei, M.; Rosi, A. The Whereabouts Diary. In *Location- and Context-Awareness SE—11*; Lecture Notes in Computer Science; Hightower, J., Schiele, B., Strang, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4718, pp. 175–192. ISBN 978-3-540-75159-5.
2. Vazquez-Prokopec, G.M.; Stoddard, S.T.; Paz-Soldan, V.; Morrison, A.C.; Elder, J.P.; Kochel, T.J.; Scott, T.W.; Kitron, U. Usefulness of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus. *Int. J. Health Geogr.* **2009**, *8*, 68. [CrossRef] [PubMed]
3. Brunauer, R.; Rehrl, K. Deriving driver-centric travel information by mining delay patterns from single GPS trajectories. In Proceedings of the 7th ACM SIGSPATIAL International Workshop on Computational Transportation Science—IWCTS '14, Dallas/Fort Worth, TX, USA, 4 November 2014; ACM Press: New York, NY, USA, 2014; pp. 25–30.
4. Liu, Y.; Seah, H.S. Points of interest recommendation from GPS trajectories. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 953–979. [CrossRef]
5. Shoval, N.; Wahl, H.-W.; Auslander, G.; Isaacson, M.; Oswald, F.; Edry, T.; Landau, R.; Heinik, J. Use of the global positioning system to measure the out-of-home mobility of older adults with differing cognitive functioning. *Ageing Soc.* **2011**, *31*, 849–869. [CrossRef]
6. Ni, Q.; Hernando, A.B.G.; de la Cruz, I.P. The Elderly's Independent Living in Smart Homes: A Characterization of Activities and Sensing Infrastructure Survey to Facilitate Services Development. *Sensors* **2015**, *15*, 11312–11362. [CrossRef] [PubMed]
7. Schneider, C.; Reich, S.; Feichtenschlager, M.; Willner, V.; Henneberger, S. Selbstbestimmtes Leben trotz Demenz. *HMD Prax. Wirtschaftsinf.* **2015**, *52*, 572–584. [CrossRef]
8. Kang, J.H.; Welbourne, W.; Stewart, B.; Borriello, G. Extracting places from traces of locations. In Proceedings of the 2nd ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots—WMASH '04, Philadelphia, PA, USA, 1 October 2004; ACM Press: New York, NY, USA, 2004; Volume 9, p. 110.
9. Hightower, J.; Consolvo, S.; LaMarca, A.; Smith, I.; Hughes, J. Learning and Recognizing the Places We Go. In *UbiComp 2005: Ubiquitous Computing SE—10*; Lecture Notes in Computer Science; Beigl, M., Intille, S., Rekimoto, J., Tokuda, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3660, pp. 159–176. ISBN 978-3-540-28760-5.
10. Changqing, Z.; Frankowski, D.; Ludford, P.; Shekhar, S.; Terveen, L. Discovering personally meaningful places : An interactive clustering approach. *ACM Trans. Inf. Syst.* **2007**, *25*, 12. [CrossRef]
11. Cao, X.; Cong, G.; Jensen, C.S. Mining significant semantic locations from GPS data. *Proc. VLDB Endow.* **2010**, *3*, 1009–1020. [CrossRef]
12. Montoliu, R.; Blom, J.; Gatica-Perez, D. Discovering places of interest in everyday life from smartphone data. *Multimed. Tools Appl.* **2013**, *62*, 179–207. [CrossRef]
13. Gröchenig, S.; Schneider, C. A Cookie-Cutter Approach for Determining Places and Stays from Movement Data. *GI_Forum* **2016**, *1*, 53–64. [CrossRef]
14. Venek, V.; Brunauer, R.; Schneider, C. Evaluating the Brownian Bridge Movement Model to Determine Regularities of People's Movements. *GI_Forum* **2016**, *2*, 20–35. [CrossRef]
15. Ashbrook, D.; Starner, T. Using GPS to learn significant locations and predict movement across multiple users. *Pers. Ubiquitous Comput.* **2003**, *7*, 275–286. [CrossRef]

16. Hariharan, R.; Toyama, K. Project Lachesis: Parsing and Modeling Location Histories. In *Geographic Information Science SE—8*; Lecture Notes in Computer Science; Egenhofer, M., Freksa, C., Miller, H., Eds.; Springer: Berlin/Heidelberg, Germany, 2004; Volume 3234, pp. 106–124. ISBN 978-3-540-23558-3.

17. Ye, Y.; Zheng, Y.; Chen, Y.; Feng, J.; Xie, X. Mining Individual Life Pattern Based on Location History. In Proceedings of the 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, Taipei, Taiwan, 18–20 May 2009; pp. 1–10.

18. Marmasse, N.; Schmandt, C. Location-Aware Information Delivery withComMotion. In *Handheld and Ubiquitous Computing SE—12*; Lecture Notes in Computer Science; Thomas, P., Gellersen, H.-W., Eds.; Springer: Berlin/Heidelberg, Germany, 2000; Volume 1927, pp. 157–171. ISBN 978-3-540-41093-5.

19. Smouse, P.E.; Focardi, S.; Moorcroft, P.R.; Kie, J.G.; Forester, J.D.; Morales, J.M. Stochastic modelling of animal movement. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2010**, *365*, 2201–2211. [CrossRef] [PubMed]

20. Merki, M.; Laube, P. Detecting reaction movement patterns in trajectory data. In Proceedings of the AGILE '2012 International Conference on Geographic Information Science, Avignon, France, 24–27 April 2012; pp. 24–27.

21. Krumm, J.; Rouhana, D. Placer: Semantic Place Labels from Diary Data. In Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing—UbiComp '13, Zurich, Switzerland, 8–12 September 2013; ACM Press: New York, NY, USA, 2013; p. 163.

22. Schmid, F.; Richter, K. Extracting Places from Location Data Streams. In Proceedings of the International Workshop on Ubiquitous Geographical Information Services, Münster, Germany, 20–23 September 2006; Available online: http://cindy.informatik.uni-bremen.de/cosy/staff/richter/pubs/schmidrichter_UbiGIS.pdf (accessed on 11 September 2017).

23. Hu, D.H.; Wang, C. GPS-based Location Extraction and Presence Management for Mobile Instant Messenger. *Embed. Ubiquitous Comput.* **2007**, 309–320. [CrossRef]

24. Laasonen, K.; Raento, M.; Toivonen, H. Adaptive On-Device Location Recognition. In *Pervasive Computing SE—21*; Lecture Notes in Computer Science; Ferscha, A., Mattern, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2004; Volume 3001, pp. 287–304. ISBN 978-3-540-21835-7.

25. Agamennoni, G.; Nieto, J.; Nebot, E. Mining GPS data for extracting significant places. In Proceedings of the 2009 IEEE International Conference on Robotics and Automation (ICRA '09), Kobe, Japan, 12–17 May 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 855–862.

26. Kim, M.; Kotz, D.; Kim, S. Extracting a Mobility Model from Real User Traces. In Proceedings of the 25th IEEE International Conference on Computer Communications, Barcelona, Spain, 23–26 April 2006.

27. Scellato, S.; Musolesi, M.; Mascolo, C.; Latora, V.; Campbell, A.T. NextPlace: A spatio-temporal prediction framework for pervasive systems. In Proceedings of the 9th International Conference on Pervasive Computing Pervasive, San Francisco, CA, USA, 12–15 June 2011; pp. 152–169. [CrossRef]

28. Ranacher, P.; Brunauer, R.; Trutschnig, W.; van der Spek, S.; Reich, S. Why GPS makes distances bigger than they are. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 316–333. [CrossRef] [PubMed]

29. Schneider, C.; Zutz, S.; Rehrl, K.; Brunauer, R.; Gröchenig, S. Evaluating GPS sampling rates for pedestrian assistant systems. *J. Local. Based Serv.* **2016**, *9725*, 1–28. [CrossRef]

30. Porras Bernárdez, F.D. Extraction of User's Stays and Transitions from GPS Logs: A Comparison of Three Spatio-Temporal Clustering Approaches. Master's Thesis, Technische Universität Wien, Wien, Austria, 2016.

31. Schneider, C.; Henneberger, S. Electronic Spatial Assistance for People with Dementia: Choosing the Right Device. *Technologies* **2014**, *2*, 96–114. [CrossRef]

32. Schneider, W.; Hennig, A. *Lexikon Kennzahlen für Marketing und Vertrieb*; Springer: Berlin/Heidelberg, Germany, 2008; ISBN 978-3-540-79861-3.

33. Leroy, G. *Designing User Studies in Informatics*; Health Informatics; Springer: London, UK, 2011; ISBN 978-0-85729-621-4.

34. Andreopoulos, B. Clustering Categorical Data. In *Data Clustering*; Chapman & Hall/CRC Data Mining and Knowledge Discovery Series; Aggarwal, C.C., Reddy, C.K., Eds.; Chapman and Hall/CRC: Palm Beach County, FL, USA, 2013; p. 652. ISBN 978-1-4665-5821-2.

35. LaMarca, A.; Chawathe, Y.; Consolvo, S.; Hightower, J.; Smith, I.; Scott, J.; Sohn, T.; Howard, J.; Hughes, J.; Potter, F.; et al. Place Lab: Device Positioning Using Radio Beacons in the Wild. In *Pervasive Computing SE—8*; Lecture Notes in Computer Science; Gellersen, H.-W., Want, R., Schmidt, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3468, pp. 116–133, ISBN 978-3-540-26008-0.

36. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, OR, USA, 2–4 August 1996; Simoudis, E., Han, J., Fayyad, U.M., Eds.; AAAI Press: Palo Alto, CA, USA, 1996; pp. 226–231.

37. Ankerst, M.; Breunig, M.M.; Kriegel, H.-P.; Sander, J. OPTICS: Ordering points to identify the clustering structure. In Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data—SIGMOD '99, Philadelphia, PA, USA, 31 May–3 June 1999; ACM Press: New York, NY, USA, 1999; pp. 49–60.

38. Birant, D.; Kut, A. ST-DBSCAN: An algorithm for clustering spatial–temporal data. *Data Knowl. Eng.* **2007**, *60*, 208–221. [CrossRef]

39. Apache Apache Commons Math 3.6.1 API. Available online: http://commons.apache.org/proper/commons-math/javadocs/api-3.6.1/index.html (accessed on 25 July 2017).

40. Zhou, A.; Zhou, S.; Cao, J.; Fan, Y.; Hu, Y. Approaches for scaling DBSCAN algorithm to large spatial databases. *J. Comput. Sci. Technol.* **2000**, *15*, 509–526. [CrossRef]

41. Borah, B.; Bhattacharyya, D.K. An improved sampling-based DBSCAN for large spatial databases. Proceedings of International Conference on Intelligent Sensing and Information Processing, Chennai, India, 4–7 January 2004; pp. 92–96. [CrossRef]

42. Schubert, E.; Sander, J.; Ester, M.; Kriegel, H.P.; Xu, X. DBSCAN Revisited, Revisited. *ACM Trans. Database Syst.* **2017**, *42*, 1–21. [CrossRef]

43. Microsoft GeoLife: Building Social Networks Using Human Location History. Available online: https://www.microsoft.com/en-us/research/project/geolife-building-social-networks-using-human-location-history/ (accessed on 25 July 2017).

44. Zheng, Y.; Chen, Y.; Li, Q.; Xie, X.; Ma, W.-Y. Understanding transportation modes based on GPS data for web applications. *ACM Trans. Web* **2010**, *4*, 1–36. [CrossRef]