Detecting Anomaly and Its Sources in Activities of Daily Living

Salisu Wada Yahaya¹ · Ahmad Lotfi¹ · Mufti Mahmud¹

Received: 14 September 2020 / Accepted: 2 December 2020 / Published online: 4 January 2021
© The Author(s) 2020

Abstract

To support the independent living and improve the quality of life for the increasing ageing population, system for monitoring their daily routine and detecting anomalies in the routine is required. Existing anomaly detection systems are unable to identify the sources of the abnormalities, thereby hindering the development of adaptive monitoring systems with reduced false prediction rate. In this paper, an approach for identifying the sources of abnormalities in human activities of daily living is proposed. Anomalies are detected by modelling the existing activity data representing the usual behavioural routine of an individual to serve as a baseline model. Subsequent activities deviating from the baseline are then classified as outliers or anomalies. An ensemble of one-class support vector machine, isolation forest, robust covariance estimator and local outlier factor is utilised for the anomaly detection achieving an accuracy of 98%. The proposed approach for identifying anomaly sources is based on the concept of similarity measure using distance functions. Two methods for measuring the pairwise distance of the features of the activity data termed as one vs one similarity measure and one vs all similarity measure are proposed. Experimental evaluation of the proposed approach on activities of daily living datasets has shown the credibility of the proposed approach for utilisation in an in-home monitoring system.

Keywords Anomaly detection · Activities of daily living · Similarity measure · Local outlier factor · Isolation forest · One-class support vector machine

Introduction

In the context of ambient assisted living (AAL), to promote the independent living of older adults, it is essential to monitor their activities of daily living (ADL) such as mobility, eating, maintaining personal hygiene and continence. Detecting abnormality in these activities is of paramount importance, especially for older adults since it allows family and carers (formal or informal) to be aware of the need for intervention. Abnormality in individual activities can be termed as any significant change or deviation from the usual routine of that individual [7]. These changes can be an early indication of health decline, which may be detrimental to wellbeing. For example, studies have shown that early indicators of mild cognitive impairment (MCI) such as Dementia are identifiable through changes in behavioural

routine such having a frequently interrupted sleep, having high activity level during the night and less activity during the day, forgetfulness and confusion in carrying out daily routines etc. [7].

Considering the rise in the population of older adults, increase in the cost of care and the desire of the older adults to stay in their homes for as long as possible (rather than being looked after in care homes), the most feasible approach to improve their quality of life and promote independent living is through an in-house monitoring system capable of detecting abnormalities in their daily activities [11, 30, 33, 40].

Due to the variability of activity from one individual to another, there exists no standard dataset representing normal or abnormal behaviours. This makes abnormality detection a very difficult task. A data-driven approach seems promising as it involves modelling each individual's routine independently based on the individual's activity data. Researchers have taken the approach of modelling the usual behavioural routine of individuals to serve as a baseline model. Subsequent activities are compared to baseline and significant deviations are classified as *outliers* or *anomalies* [6, 27, 50].

✉ Salisu Wada Yahaya
salisu.yahaya2015@my.ntu.ac.uk

¹ Department of Computer Science, School of Science and Technology, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, UK

This approach is also referred to as “Novelty Detection” or “One-Class Classification” [38].

The limitation of the existing anomaly detection models is their inability to identify the anomaly sources. The term “Source(s) of Anomaly” is used to refer to the feature of the data entry responsible for the classification of that entry as an outlier. For example, for activity-based anomaly, detection system for sleeping routine, different features of the sleeping activity such as the time, duration, day or week, interruptions, activity transitions, proceeding or succeeding activities can be the source of the abnormality. As an illustration, Fig. 1 shows a sample plot representing a sleeping activity pattern with two features (i.e. start time and duration) selected from several features for easy visualisation. The clusters “X1”, “X2” and “X3” represent the normal activity routines while A, B, and C indicate abnormalities. Existing computational models can identify the anomalies but not the sources of the anomalies. The sources can be observed from Fig. 1 as follows: (1) the anomaly source for “A” is the duration feature indicating insufficient sleep or oversleeping, (2) the anomaly source for “B” is the start time feature indicating going to bed early or late while (3) the anomaly source for “C” is both the start time and duration features.

The main contribution of this paper is the proposal of a similarity measure approach for identifying the sources of anomalies in human activities. The proposed exploratory approach is based on a pairwise distance measure of the features extracted from the activity data. Two approaches for measuring the pairwise distance are presented and termed as one vs one similarity measure (OOSM) and one vs all similarity measure (OASM). Statistical measures are then applied to estimate the threshold of the extracted features with features exceeding the threshold predicted as the

Fig. 1 A sample plot representing sleeping activity with three outliers A, B and C

anomaly sources. An ensemble of novelty detection models consisting of one-class support vector machine (OS-SVM), isolation forest (iForest), local outlier factor (LOF) and robust covariance estimator (RCE) is applied for the detection of outliers in the data. The proposed similarity measure approach is then applied to the predicted outliers to identify their sources.

The rest of this paper is organised as follows: “[Related work](#)” provides a review of related studies, while “[Anomaly detection in human activities](#)” gives an overview of the anomaly detection models. In “[Proposed similarity measure approach to identify anomaly source](#)”, the similarity measure approach is presented. “[Experimentation and discussion](#)” contains experimental results. A conclusion of the work is then provided in “[Conclusion](#)”.

Related Work

Research in the field of ambient intelligence (AmI) has been receiving tremendous attention, primarily due to the need to promote the independent living of the increasing ageing population [1]. The advent of miniaturised sensing devices allows for an easy collection of data for human activity recognition (HAR) and behaviour modelling. The work by Thapliyal et al. [45] took a close look at smart home technologies and their applications for improved wellbeing of the older adults. Activity recognition is primarily focused on interpreting the collected sensor data into a human-readable form [23]. Behaviour modelling deals with the utilisation of the activity data to build a profile of the monitored individual. Different methodologies have been applied for monitoring and recognition of human activities using ambient sensors, wearable sensors or both [3, 9, 19, 31, 39, 41].

In ADL context, approaches for detecting abnormalities in human activities are broadly classified into three categories, namely: (1) score-based approach, (2) classification approach and (3) outlier detection approach.

The score-based approach involves clinical assessment of the older adults health by a clinical expert. Certain aspects of their daily routines (such as mobility, cognitive health, etc.) are assessed over some time (e.g. every 6 months) and a score is assigned for each of the categories during the assessment. A computation model (usually a regression model) is then trained to map the data collected from ambient sensors in the home environment during the period of the assessment to the assigned score. The aim is to have the trained model predict future health score from the future ADL data. The clinical assessments are carried out using a questionnaire or by mere observation of an expert. While this approach seems promising, it is subjective and highly dependent on the period in which the assessment is performed. For instance, a wrong score may be assigned to an individual

which may not reflect the actual health status of the subject, especially if the evaluation is performed when the subject is not in a good state of mind.

Based on the above concept, the author in [16] proposed clinical assessment using activity behaviour (CAAB). The proposed approach is applied to ADL dataset collected for the duration of over 2 years from 18 different smart homes. To evaluate the approach, statistical correlation is established between the score assigned by the expert and the prediction of the CAAB regression model. Similar work is carried out by [5] using the same dataset. The only difference is that this clinical assessment encompasses more areas than CAAB, such as money and self-management, home daily living, travel and event memory, and social skills. This large set of activities is referred to as incremental activities of daily living-compensation (IADL-C). In addition to assigning a health score, a classifier is trained to predict changes in recurrent health scores. In the extended version of the paper, the author attempts to reduce the class imbalance problem by oversampling the minority class [2]. The results obtained show a correlation between the expert's assigned scores and the model's prediction. The classifier performs poorly in predicting changes in recurrent assessment.

In the classification approach, ADL anomaly detection is treated as a binary classification problem. Binary classification task requires training data for the two classes to be available. Due to the absence of training data representing normal and abnormal activities, the approach taken is to use the collected data as the training samples for the normal class, while synthetic data are generated for the abnormal class. The synthetic data are generated to reflect common anomalies relating to MCI. For instance, since frequently interrupted sleep and confusion in the performance of daily activities are attributed to early symptoms of Dementia, the data is generated to mimic these symptoms. The authors in [7] used this approach. They artificially insert anomalous events into the collected data such as random events in between sleeping activity to reflect a disturbed sleep and altering the order of activity sequence to simulate confusion in the carrying out the activity, etc. A classifier is trained on the data to efficiently detect anomalies. The major drawback of this approach is that for the model to detect any anomaly, a synthetic data simulating that anomaly must be generated. Generate every variation of an anomaly is nearly impossible due to the variability of anomalies across individuals.

A combination of convolutional neural network (CNN) and long-short term memory (LSTM) is used to detect synthetic anomalies in [7]. The ADL data are converted into a 2D image map and fed into a CNN to learn data encoding while LSTM learns the activity sequence. In [22], an approach for early detection of anomalous behaviour in smart homes based on causal association rule is proposed. The identified causes are then utilised with Markov Logic

network for identifying the risk of anomaly occurrence at a later time, and for recommending suitable actions to avoid future occurrences. In [51], the authors perform a comparative analysis of LSTM, CNN, CNN-LSTM and Autoencoder for accurate detection of ADL anomalies. The validation is performed on public datasets while employing an oversampling technique to alleviate the non-uniformity in the data distribution.

The outlier-based approach differs from the classification approach as only one set of training data is required. The collected ADL data are used as the training sample representing normal behaviour for building the baseline model. Subsequent activity data are compared to the baseline, and significant variations are classified as outliers. This concept is adopted due to its less prone to errors and versatility compared to the classification approach since there is no need to generate every variation of an anomaly. Similar to the score-based and classification-based approaches, the collected data must reflect the usual behavioural routine of the individual. Data that is contaminated with outliers could lead to poor performance by the model.

The authors in [25] utilised ADL data collected with ambient sensors to detect abnormal activity instances using different entropy measures. Activities with entropy value exceeding a certain range are predicted as outliers. Abnormality detection algorithms based on hidden state conditional random field (HCRF) are proposed in [47]. Data collected using ambient sensors are utilised for the validation of the proposed approach, and a comparison with SVM is carried out. Similarly, the authors in [4] proposed a system for promoting the safety of autistic children indoor. This system is based on a 3D-CNN and LSTM applied to a video stream to predict activities with physical irregularities. This system applies to older adults monitoring. A Kinect depth sensor is used in [37] for identifying deviating activities that could constitute an abnormality in a smart home environment. A data-driven technique is used to define fuzzy sets over attributes of the occupant's behaviour and a fuzzy inference engine with a membership function is used to identify an abnormal pattern.

In [27, 49], OC-SVM is used to detect ADL anomalies. An approach for creating an ensemble of outlier detection models is proposed and applied for the detection of anomalies in ADL dataset in [50]. Density-based spatial clustering of applications with noise (DBSCAN) is used for the detection of ADL anomalies in [18] by clustering the duration, time and number of triggered sensor events of the activities. Activities with irregular durations, time, and the number of triggered events are classified as anomalous. DBSCAN is also used in [24] by classifying activities that fall outside the cluster as abnormal. Researchers have also utilised temporal models for identification of ADL anomalies for activities that are temporal in nature.

An approach is proposed for detecting temporal relation between activities and detecting abnormality in them [28]. Modelling the sequential pattern of activities can allow for the detection of anomalies since anomalies are novel patterns. In [6], recurrent neural network (RNN) is utilised while Hidden Markov Model (HMM) is used in combination with a Fuzzy rule-based system (FRBS) for the detection of anomalies in human activities in [20]. Anomalies relating to room occupancy is detected using Self Organising Map (SOM) in [35], while echo state network (ESN) is used in [30].

The existing system for anomaly detection in ADL often operate in an offline manner. To address this limitation, Meng et al. [32] proposed online daily habit modelling and anomaly detection (ODHMAD) model capable of performing real-time activity recognition and anomaly detection from ambient sensors data. Computational models for anomaly detection have been applied in a non-ADL context as well. In [44], anomalies in human activity is detected from a video stream using a CNN with an adaptive compression technique. Similarly, CNN-based approach for anomaly detection in images is proposed in [21].

Outlier detection models are based on different methodologies. Some of the models are based on a distance measure that involves estimating the nearest neighbours of the data points. The data points with close neighbours are classified as normal and vice versa. One of the drawbacks of this approach is that it is computationally expensive in high dimensional space [38]. In [15], this distance approach is used for the detection of a disease outbreak. In density-based approaches, data points in the region of low density are classified as outliers and normal otherwise [38]. The approach assumes that the data are of a certain distribution, which makes it practically impossible for real-life application since the data distribution is not known a priori. To mitigate this, non-parametric approaches are proposed in which the distributions are estimated from the training sample [38].

To enhance the performance of the outlier detection models, ensemble models are used since each individual model is good at certain characteristic feature. Dib et al. [17] applied an ensemble of machine learning models to monitor structural health by detecting damages using guided waves generated by the building sensors. An ensemble of novelty detection models based on a consensus vote is proposed in [50] and applied for the detection of ADL anomalies, while in [36], an ensemble approach based on similarity measures is proposed.

From the reviewed literature, it can be seen that anomaly detection is relevant to different application areas. A summary of the reviewed works relating to ADL anomaly detection is presented in Table 1.

Anomaly Detection in Human Activities

The proposed workflow for identifying the sources of abnormalities in human activities presented in 5 steps is shown in Fig. 2. To detect the anomaly sources, the underlining ADL data must contain outliers identifiable using computational models. Detecting abnormalities is a complex task due to the variability of activities from one individual to another, thereby leading to the absence of standardised datasets representing normal and abnormal activities. Additionally, the existence of anomalies in different activities makes the anomaly detection task rather arduous. To address this, a data-driven activity-based approach is adopted whereby the computational models for anomaly detection are trained on data representing the activities of interest.

The initial set of data collected from the home environment representing the usual behavioural routine of the monitored individual is used to train the novelty detection models to serve as a baseline. The baseline model is then utilised on subsequent activity data to detect deviating routines that constitute an abnormality. In this context, the result of the model for each activity is considered as either *normal* or *anomaly*. The observations predicted as anomalies are then explored further to identify the actual source of the abnormality as depicted in Fig. 2. In the remaining part of this section, six computational models for behaviour modelling and abnormality detection are explained. The exploratory similarity measure approach for identifying the anomaly sources (stage 5 of the workflow) is presented in “[Proposed similarity measure approach to identify anomaly source](#)”.

One-Class Support Vector Machine (OC-SVM)

The OC-SVM is an unsupervised novelty detection algorithm proposed by Scholkopf et al. [43]. To identify outliers, assuming the data have an underlying probability distribution P , the problem can be framed into a minimisation of a quadratic function.

Let $A = \{a_1, \dots, a_n\}$ be an n set of d -dimensional data ($A \in \mathbb{R}^d$), and $\Phi : \mathbb{R}^d \rightarrow F$ be a non-linear mapping from a data space $\mathbb{R}^d$ to a feature space F . The support vector separating the data is computed by solving the quadratic problem:

$$\min_{w \in F, \xi \in \mathbb{R}^d, \rho \in \mathbb{R}} \frac{1}{2} \|w\|^2 + \frac{1}{vn} \sum_{i=1}^n \xi_i - \rho, \quad (1)$$

subject to

$$(w \cdot \Phi(a_i)) \geq \rho - \xi_i, \quad \xi_i \geq 0, \quad i = 1, \dots, n,$$

where n is the training sample size, $v \in (0, 1)$ is a trade-off parameter for the expected fraction of outliers, ξ is a slack

Table 1 Studies on ADL anomaly detection

References	Data type	Modelling approach	Evaluation
[28]	Inferred ADL data	Temporal relation of activities	Percentage of correct prediction
[16]	Inferred ADL data	CAAB and regression model	Correlation and RMSE
[22]	Inferred ADL data	Causal association rule and Markov logic network	Accuracy, precision and recall
[49]	Inferred ADL data	OC-SVM	Accuracy
[2, 5]	Inferred ADL data	IADL-C, Regression and classification model	RMSE, correlation, F-Score, accuracy, sensitivity and ROC-AUC
[7]	Raw binary data	CNN + LSTM	Sensitivity, specificity
[30]	Raw binary and inferred ADL data	ESN	RMSE
[24]	Inferred ADL data	DBSCAN	Precision and recall
[27]	Inferred ADL data	OC-SVM	Error rate, F-measure and accuracy
[6]	Raw binary data	Vanilla RNN, GRU, LSTM	Accuracy
[18]	Inferred ADL data	DBSCAN	EER
[20]	Inferred ADL data	HMM + FRBS	Accuracy, precision, recall, F-measure, ROC-AUC
[50]	Inferred ADL and synthetic data	Ensemble (OC-SVM, LOF, iForest, RCE)	Accuracy
[34, 35]	Raw binary data	SOM	Accuracy
[25]	Raw binary data	Entropy measures	Accuracy, precision and recall
[47]	Inferred ADL data	HCRF & SVM	Accuracy
[4]	ADL data (video stream)	3D-CNN + LSTM	Accuracy, precision, recall, F-measure and specificity
[37]	ADL data (depth)	Fuzzy inference system	Percentage error rate
[32]	Inferred ADL data	(ODHMAD)	Precision, false alarm and false prediction rate
[51]	Inferred ADL data	LSTM, CNN, CNN-LSTM and autoencoder	Accuracy

RMSE root mean square error, *MAE* mean absolute error, *EER* event error rate, *RCE* robust covariance estimation

Fig. 2 A workflow of the anomaly detection and source identification process

variable, w and ρ are the hyper-plane parameters in the feature space F .

To solve the Eq. 1, the decision function $f(a)$ which determines if a_i is an outlier is obtained as:

$$f(a) = \sum_{i=1}^n \alpha_i k(a_i, a) - \rho, \quad (2)$$

where α_i is a Lagrange multiplier for the vectors a_i and $k(a_i, a) = \Phi$ is the kernel function for the non-linear mapping.

$$k(a_i, a) = \Phi(a_i) \cdot \Phi(a). \quad (3)$$

The radial basis function (RBF) kernel with spread parameter γ is used:

$$K(a_i, a) = \exp(-\gamma \|a_i - a\|^2) \quad (4)$$

$$f(a) = \begin{cases} +1 & a \text{ is NOT an outlier} \\ -1 & a \text{ is an outlier} \end{cases} \quad (5)$$

The prediction of the model is determined from Eqs. 3 and 5. A more details description of OC-SVM can be found in [43].

Isolation Forest

The iForest is based on the premise that outliers are few and different, making them susceptible to isolation [29]. This makes the detection of anomalies based on the concept of isolation rather than using distance or density measures.

Let $A = \{a_1, \dots, a_n\}$ be an n set of d -dimensional data ($A \in \mathbb{R}^d$). An isolation tree (iTree) is built from subsample instances ψ of $A' \subset A$ by recursively dividing A' with a randomly selected attribute q and a split value p , until either: (1) the node has only one instance or (2) all data at the node have the same values. The iTree is a binary tree with each node having exactly zero or two child nodes. Assuming all instances are distinct, each instance is isolated to an external node when an iTree is fully grown, the number of external nodes is ψ , the number of internal nodes is $(\psi - 1)$ and the total number of nodes of an iTree is $(2\psi - 1)$. To detect outliers, a ranking of the data that reflects the degree of anomaly is performed through the sorting of the data points according to their path lengths or anomaly scores. The anomalies are points that are ranked at the top of the list. The different terminologies for the iForest are defined as follows:

- **Isolation Tree (definition)** Let T be a node of an iTree. T is either an external node with no child, or an internal node with one test and exactly two child nodes (T_l, T_r). A test at node T consists of an attribute q and a split value p such that the test $q < p$ determines whether the data point belongs to either T_l or T_r .
- **Path length (definition)** The path length $h(a)$ of a point a is a measure of the number of edges a traverses an iTree from the root node to an external node such that a short path length indicates a high degree of susceptibility to isolation and a long path length indicating low susceptibility.
- **Anomaly score (definition)** The anomaly score $s(a, \psi)$ for a data point a of data set with n instances is given as:

$$s(a, \psi) = 2^{-\frac{E(h(a))}{c(\psi)}}. \quad (6)$$

$E(h(a))$ is the average path length ($h(a)$), $c(\psi)$ is the average path length of unsuccessful binary search, and ψ is the subsample size from $A = \{a_i\}_{i=1, \dots, n}$.

$$c(\psi) = 2H(\psi - 1) - 2(\psi - 1)/\psi, \quad (7)$$

where $H(i)$ is a harmonic number estimated as $\ln(i) + 0.5772156649$ (Euler's constant is 0.5772156649). The interpretation of the outcome of the anomaly score is summarised below with a more detailed discussion in [29].

- The instances are anomalies if the value of s is close to 1 ($E(h(a)) \rightarrow 0, s \rightarrow 1$).

- The instances are not anomalies if the value of s is much smaller than 0.5 ($E(h(a)) \rightarrow (\psi - 1), s \rightarrow 0$) and
- The entire instances have no distinct anomalies if the values of s is approximately 0.5 ($E(h(a)) \rightarrow c(\psi), s \rightarrow 0.5$).

Local Outlier Factor

The local outlier factor (LOF) is an unsupervised approach for outlier detection that is based on density estimation of a data point relative to its nearest neighbours. Since the density around an outlier differs significantly from the density around its neighbours, data points with relatively lower density to its neighbours estimated in a form of a score are considered as outliers [8].

Let $A = \{a_1, \dots, a_n\}$ be an n -sample dataset. Let $d(a, b)$ denote the distance between two objects a and b . Assuming $a = a_i$ is an observation in A , the k -distance of a denoted by $D_k(a)$ is the distance between a and its k -nearest neighbours while the k -distance neighbourhood of a denoted by $N_k(a)$ contains every element whose distance from a is not greater than k -distance $D_k(a)$ of a given as:

$$N_k(a) = \{x \in A | d(a, x) \leq D_k(a)\}, \quad (8)$$

x is an entry in A , which is a neighbour of a

The k -reachability distance $RD_k(a)$ of the object a w.r.t an its nearest neighbours a' is given as:

$$RD_k(a, a') = \max\{D_k(a), d(a, a')\}. \quad (9)$$

The local reachability density $LRD_k(a)$ of a is given as:

$$LRD_k(a) = \frac{|N_k(a)|}{\sum_{a' \in N_k(a)} RD_k(a, a')}. \quad (10)$$

The average of the ratio of the local reachability density of a and its k -nearest neighbours is presented as:

$$LOF_k(a) = \frac{\sum_{a' \in N_k(a)} \frac{LRD_k(a')}{LRD_k(a)}}{|N_k(a)|}. \quad (11)$$

The lower the local reachability density $LRD_k(a)$ of a the higher the local reachability density $LRD_k(a')$ of its nearest neighbours a' , the higher the LOF $LOF_k(a)$, making the a an anomaly. A more detailed discussion of LOF can be found in [8].

Covariance Matrix Estimation

In this approach, outliers are detected using the maximum likelihood estimators (MLE) for the mean and the covariance matrix of the normal (non-outlying) data. Observations with

abnormally large distances (Mahalanobis distances) are predicted as outliers [42].

Let $A = \{a_1, \dots, a_n\}$ be an n -sample dataset of d -dimension and each observation $a_i = \{a_{i1}, \dots, a_{id}\}$. Assuming the data follows a multivariate normal distribution, The mean ($\hat{\mu}$) and covariance matrix ($\hat{\Sigma}$) estimated using MLE are given as:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n a_i \quad (12)$$

$$\hat{\Sigma} = \sum_{i=1}^n (a_i - \hat{\mu})(a_i - \hat{\mu})^T / (n - 1) \quad (13)$$

The Mahalanobis distance of the observations are computed as:

$$MD(a, \hat{\mu}, \hat{\Sigma}) = \sqrt{(a - \hat{\mu})^T \hat{\Sigma}^{-1} (a - \hat{\mu})} \quad (14)$$

The estimations above assume that the observations are not contaminated as few outliers could lead to poor performance. Alternative robust approaches termed as Robust covariance estimation (RCE) for estimating the location (mean) and scatter (covariance) such as using minimum covariance determinant (MCD) are proposed [42].

Ensemble of Detectors Based on Similarity Measure

Ensemble approaches allow for the aggregation of individual models for better performance. Ensemble models for outlier detectors based on similarity measures are proposed in [36]. The authors proposed two approaches namely, ensemble of detectors with correlated votes (EDCV) and ensemble of detectors with variability votes (EDVV). The proposed approaches merely estimate the appropriate weight for each model in the ensemble using the individual model's results for a given dataset. The weight estimation in EDCV is based on correlation coefficient while the weight estimation in EDVV is based on a measure of mean absolute deviation (MAD) of the individual model output. Using the estimated weights, a score for the data entries are calculated to determine if the data points are outliers.

Let $A = \{a_1, \dots, a_n\}$ be an n set of d -dimensional data ($A \in \mathbb{R}^d$). Let $G = \{g_1, \dots, g_m\}$ be a set of m outlier detection models. a_i is an observation in A and g_i is a model in G . The weight $W(g_i)$ for g_i is estimated using Equation 15 and Equation 16 for the EDCV and EDVV respectively.

$$W(g_i) = \frac{(\sum_{i=1}^m C_i) - 1}{m - 1}, \text{ for EDCV,} \quad (15)$$

C is a matrix of the correlation coefficients of A for the corresponding model g_i and m is the number of models in the ensemble.

$$W(g_i) = \frac{(\sum_{i=1}^m \mathbb{D}_i)}{m - 1}, \text{ for EDVV,} \quad (16)$$

$\mathbb{D}$ is a matrix of the MAD of A for the corresponding model g_i and m is the number of models in the ensemble

The final score for a data entry a_i is calculated as:

$$S(a_i) = \frac{(\sum_{j=1}^m F(i, j) * V(i, j) * W(j))}{m}, \quad (17)$$

$S(a_i)$ is the final outlier score, m the number of models, $F(i, j)$ is the outlier score, $V(i, j)$ is the label (output), and $W(j)$ is the weight of the j th model respectively.

Consensus Novelty Detection Ensemble

As part of our earlier contribution for robust detection of ADL anomalies, an ensemble approach is proposed in [50] termed consensus novelty detection ensemble (CNDE). The approach is based on consensus votes obtained for an activity and the computation of a normality score that determines the data label (either outliers or inliers).

Let $A = \{a_1, \dots, a_n\}$ be an n set of d -dimensional data ($A \in \mathbb{R}^d$). Let $G = \{g_1, \dots, g_m\}$ be a set of m outlier detection models. The Normality Score $\mathcal{N}(a)$ which determine if a is an outlier is calculated as:

$$\mathcal{N}(a) = \text{Avg}(\chi(a), \xi(a)) \quad (18)$$

where

$$\xi(a) = \frac{1}{m} \sum V(a) \quad (19)$$

and

$$\chi(a) = \frac{1}{m} \sum_{i=1}^m \varphi_i(a) * \omega_i \quad (20)$$

$\chi(a)$ is the Combined Internal Consensus Score (CICS) and $\xi(a)$ is the Combined External Consensus Score (CECS) for the data point a . m is the ensemble model size, $V(a)$ is the aggregated votes of a by the models. $\varphi_i(a)$ represents the Internal Consensus Score (ICS) while ω^i is the weight of the i th model respectively, calculated as:

$$\varphi(a) = \frac{1}{k} \sum_{i=1}^k u \quad (21)$$

$$\omega = \tilde{\omega} - \frac{e}{n} \tilde{\omega}, \quad (22)$$

where $\varphi(a)$ is the ICS, k is the number of folds the training data is split into to train k sub-models, u is the vote from the sub-models. ω denotes the final weight (after penalisation), $\tilde{\omega}$ is the initial weight (before penalisation usually initialised

to 1), e is the false prediction rate for each model in the ensemble respectively and n is the training sample size.

A threshold value ρ is introduced for the Normality Score $\mathcal{N}(a)$ estimated from the distribution of $\mathcal{N}(a)$. The function $f(a)$ determines the model outcome as:

$$f(a) = \begin{cases} \mathcal{N}(a) < \rho & x \text{ is an outlier} \\ \text{otherwise} & x \text{ is NOT an outlier} \end{cases} \quad (23)$$

The Normality Score $\mathcal{N}(a)$ value ranges from 0 to 1 ($0 \leq \mathcal{N}(a) \leq 1$) with lower score value ($\mathcal{N}(a) \rightarrow 0$) signifying outliers. An elaborated discussion of this ensemble approach can be found in [50].

The presented anomaly detection models are applied to ADL datasets for the detection of outliers. Additionally, an ensemble of models based on a majority vote approach is considered and the obtained results are presented in Sect. 5.

Proposed Similarity Measure Approach to Identify Anomaly Source

Computational models identify outliers by measuring the similarity between data points such that entries with a significant variance in the measured similarities are predicted as anomalies. The process of identifying the sources of anomaly involves (a) estimating the similarity matrix of the training data, (b) estimating the threshold of the similarity matrix, (c) computing the similarity matrix of the anomalous observation and d) calculating the similarity score of the anomalous observation to identify the source as shown in the schematic diagram in Fig. 3.

Different similarity measures exist in the literature for both numerical and categorical data with distance functions such as Euclidean, Minkowski, Manhattan distances commonly used as a measure of similarity for numerical data [48]. In general, machine learning models for both supervised and unsupervised learning utilises distance measures. For example, K-nearest neighbour (KNN) uses distance functions such as Euclidean and Manhattan to measure the distances between a data point and its neighbours, with instances having relatively close neighbours

Fig. 3 A schematic diagram of the proposed similarity measure approach

classified as observations of the same class [26, 48]. Clustering algorithms such as K-means uses distance measure to estimate the proximity of the data entries to their assigned cluster centroids [48].

Two approaches for the distance measurement are proposed and termed as; one vs one similarity measure (OOSM) and one vs all similarity measure (OASM). The approaches involve the pairwise distance measurement of the corresponding features of the dataset. We utilised three distance functions, namely Euclidean, Chebyshev and Canberra distance satisfying the fundamental distance measure requirements as follows [48]:

- Non-negativity rule: $D(x, y) \geq 0$.
- Symmetry rule: $D(x, y) = D(y, x)$.
- Identity of indiscernible rule: $D(x, y) = 0$ iff $x = y$.

In the OOSM, the distance is measured between corresponding features of the data entries one at a time. For example, the distance of the i th feature of the first entry is a_1 is measured against the i th feature of the second entry a_2 , then with that of the third entry a_3 and vice-versa across all the features as a given in Eq. 27. In the OASM, the distance of the corresponding features is measured all together in one single expression. For example, the distance of the i th feature of the first entry a_1 is measured against the i th feature of all the remaining entries as presented in Eq. 31.

One vs One Similarity Measure (OOSM)

Let $A = \{a_1, a_2, \dots, a_n\}$ be a d -dimensional set of n normal activity data used for training the anomaly detection model. The similarity matrix $\mathbb{S}(A)$ is calculated as the distance of the pairwise features of the observations. Expressing the data A in a matrix form:

$$A = \begin{bmatrix} a_1^1 & - & - & a_1^d \\ - & - & - & - \\ - & - & - & - \\ a_n^1 & - & - & a_n^d \end{bmatrix} \quad (24)$$

Given the similarity matrix $\mathbb{S}(A)$ for A :

$$\mathbb{S}(A, A) = \mathbb{S}(A) = \begin{bmatrix} \mathbb{S}(a_1, A) \\ - \\ - \\ \mathbb{S}(a_n, A) \end{bmatrix} \quad (25)$$

The similarity $\mathbb{S}(a_i, A)$ for the i th observation of A is can be presented as:

$$\mathbb{S}(a_i, A) = [a_i^1 \quad - \quad - \quad a_i^d] \begin{bmatrix} a_1^1 & - & - & a_1^d \\ - & - & - & - \\ - & - & - & - \\ a_n^1 & - & - & a_n^d \end{bmatrix} \quad (26)$$

using OOSM approach, $\mathbb{S}(a_i, A)$ is calculated as:

$$\mathbb{S}(a_i, A) = \begin{bmatrix} \partial(a_i^1, a_1^1) & - & - & \partial(a_i^d, a_1^d) \\ - & - & - & - \\ - & - & - & - \\ \partial(a_i^1, a_n^1) & - & - & \partial(a_i^d, a_n^d) \end{bmatrix}, \mathbb{S}(a_i, A) \in \mathbb{R}^{n \times d} \quad (27)$$

where $\partial(a_i^k, a_j^k)$ is a 1-dimensional pairwise distance calculated using the three distance functions as:

$$\text{Euclidean: } \partial(a_i^k, a_j^k) = \sqrt{(a_i^k - a_j^k)^2} = |a_i^k - a_j^k| \quad (28)$$

$$\text{Chebyshev: } \partial(a_i^k, a_j^k) = \max |a_i^k - a_j^k| = |a_i^k - a_j^k| \quad (29)$$

$$\text{Canberra: } \partial(a_i^k, a_j^k) = \frac{|a_i^k - a_j^k|}{|a_i^k| + |a_j^k|} \quad (30)$$

One vs All Similarity Measure (OASM)

The similarity matrix $\mathbb{S}(a_i, A)$ in Eq. 26 is calculated as:

$$\mathbb{S}(a_i, A) = \left[\partial(a_i^1, a_{j=1, \dots, n}^1) \quad - \quad - \quad \partial(a_i^d, a_{j=1, \dots, n}^d) \right], \mathbb{S}(a_i, A) \in \mathbb{R}^{1 \times d} \quad (31)$$

The one-dimensional pairwise distance $\partial(a_i^k, a_{j=1, \dots, n}^k)$ is a calculated as:

$$\text{Euclidean: } \partial(a_i^k, a_{j=1, \dots, n}^k) = \sqrt{\sum_{j=1}^n (a_i^k - a_j^k)^2} \quad (32)$$

$$\text{Chebyshev: } \partial(a_i^k, a_{j=1, \dots, n}^k) = \max_{j=1}^n |a_i^k - a_j^k| \quad (33)$$

$$\text{Canberra: } \partial(a_i^k, a_{j=1, \dots, n}^k) = \sum_{j=1}^n \frac{|a_i^k - a_j^k|}{|a_i^k| + |a_j^k|} \quad (34)$$

The threshold $\vec{\mathbb{S}}(A)$ of the features (columns) of $\mathbb{S}(A, A)$ in Eq. 25 is a row vector estimated from the positively (right) skewed distribution of the entries using median and inter-quartile range rule for outlier detection [46].

$$\vec{\mathbb{S}}(A) = [\delta^1 \quad - \quad - \quad \delta^d], \vec{\mathbb{S}}(A) \in \mathbb{R}^{1 \times d} \quad (35)$$

$$\delta^k = \alpha^k + e\beta^k, \quad (36)$$

where α, β are the median and interquartile range of the k th feature while e is a constant (usually set to 1.5).

Given a set of d -dimensional outliers $B = \{b_1, b_2, \dots, b_m\}$, the anomaly source for each observation b_i is estimated by estimating its similarity with the non-outlying data given as:

$$\mathbb{S}(b_i, A) = [b_i^1 \quad \dots \quad b_i^d] \begin{bmatrix} a_1^1 & \dots & a_1^d \\ \vdots & \ddots & \vdots \\ a_n^1 & \dots & a_n^d \end{bmatrix} \quad (37)$$

$\mathbb{S}(b_i, A)$ is calculated using either OOSM in Eq. 27 or using OASM in Eq. 31. The similarity score $\tilde{\mathbb{S}}(b_i, A)$ for b_i is a row vector obtained by calculating the feature mean of $\mathbb{S}(b_i, A)$ from Eq. 37.

$$\tilde{\mathbb{S}}(b_i, A) = [\rho^1 \quad \dots \quad \rho^d], \tilde{\mathbb{S}}(b_i, A) \in \mathbb{R}^{1 \times d} \quad (38)$$

$$\rho^k = \begin{cases} \text{mean}(\mathbb{S}(b_i^k, A)) & \text{for the OOSM} \\ \mathbb{S}(b_i^k, A) & \text{for the OASM} \end{cases} \quad (39)$$

where ρ^k is the score for the k th feature of the observation b_i .

The function $f(\rho^k)$ that determines if the k feature is the anomaly source is expressed as:

$$f(\rho^k) = \begin{cases} 0 \leq \rho^k \leq \delta^k & \text{The } k\text{th feature is NOT the anomaly source} \\ \text{otherwise} & \text{The } k\text{th feature is the anomaly source} \end{cases} \quad (40)$$

The anomaly source is identified in Eq. 40. If the similarity score of the feature ρ^k has an abnormally large variance to the feature threshold δ^k , the k th feature of the observation is the anomaly source. Figure 3 shows a schematic diagram of the proposed similarity measure while its procedures are outlined in Algorithm 1.

Algorithm 1 Similarity Measure

Input: Training set $A = \{a_i\}, i = 1, \dots, n$
 Outlier observation b_j
Output: Features list (sources) $f = \{\}$

```

1: procedure ANOMALY SOURCE
2:    $\mathbb{S}(A) \leftarrow$  Compute similarity of  $A$ 
3:    $\tilde{\mathbb{S}}(A) = \{\delta^k\}, k = 1, \dots, d \leftarrow$  Estimate threshold of  $\mathbb{S}(A)$ 
4:    $\mathbb{S}(b_j, A) \leftarrow$  Compute similarity of outlier  $b_j$ 
5:    $\tilde{\mathbb{S}}(b_j, A) = \{\rho^k\}, k = 1, \dots, d \leftarrow$  Compute score of  $\mathbb{S}(b_j, A)$ 
6:   for  $k = 1 : d$  do
7:     if  $\rho^k \geq 0$  and  $\rho^k \leq \delta^k$  then
8:        $f \leftarrow \text{index}(k)$   $\triangleright k^{\text{th}}$  feature is a source
9:     else
10:       $f \leftarrow f$   $\triangleright k^{\text{th}}$  feature is NOT a source
11:    end if
12:  end for
13:  return  $f$ 

```

Experimentation and Discussion

This section contains the description of the datasets employed for the validation of the proposed approach. The preprocessing technique, experimental scenarios and the obtained results are presented in this section.

Datasets and Preprocessing

Two real datasets utilised for the evaluation consists of data collected as part of this research work from Nottingham Trent University (NTU) smart home facility termed as “SmartNTU” and a publicly available dataset from Center for Advanced Studies in Adaptive System (CASAS) repository termed as “CASAS HH111”. The datasets are collected using low-cost, non-invasive ambient sensors such as passive infrared sensors (PIR) installed in various location of the smart home to identify the presence of individuals,

Fig. 4 A sample plot of raw binary data from ambient sensors

Table 2 A sample of interpreted ADL data

Start time	End time	Activity	Location
21:19:31	21:24:11	Toileting	Toilet
21:25:01	22:14:17	Watching TV	Living room
22:14:52	07:13:07	Sleeping	Bedroom
07:15:47	07:24:20	Preparing meal	Kitchen
–	–	–	–

pressure sensors on sofas and beds to identify sleeping or sitting activity, door entry sensors to detect door opening and closing events, etc. The use of these sensors for data collection is due to their non-intrusive nature, thereby posing fewer privacy concerns, and are found to be generally more acceptable compared to visual sensors (e.g. cameras) [13].

The CASAS HH111 dataset consists of the daily activity data for a volunteer older adult living alone for a duration of 50 days [12]. During the data collection period, the recorded activities include but not limited to sleeping, toileting, eating and kitchen-related activities, etc. Similarly, the SmartNTU data consist of the recorded activity data for an individual living alone for the duration of 72 days.

A sample plot of the binary data generated by the ambient sensors (e.g. PIR sensor) is shown in Fig. 4. The attributes of the sensor data include a timestamp used to represent the start and end time of an activity, a unique device identifier for identifying the sensor location, event value, etc. Different computational models are proposed in the literature for data interpretation [14]. The interpreted activity each consist of several attributes such as start time, end time, activity type and the location of the performed activity. Table 2 shows a sample of the interpreted activities from the collected sensor data.

To efficiently evaluate the performance of the computational models and reduce bias due to class imbalance in the datasets, the datasets are oversampled using an approach proposed in [10] known as synthetic minority over-sampling technique (SMOTE).

Since the anomaly detection approach is activity dependent (i.e. anomalies are detected in activities of interest), the datasets are filtered, and sleeping activities are selected for modelling and detection of abnormalities. Additionally, artificial anomalies in sleeping activity simulated based on the data distribution are induced. The following features are extracted from the sleeping activity:

- **Activity start time** This represents the hours and minutes in which the activity begins. The hour is converted from $[0, \dots, 23]$ range to a lower scale of $[-11, \dots, 11]$.
- **Activity duration** This is a representation of the activity duration (in min) obtained by subtracting the start time of the activity from its end time.

- **Interruption(s)** This is used to represent the transitions from the sleeping activity to other activities. If the duration of the transition is less than 1 h, the activity is considered interrupted, otherwise, the end of the activity.
- **Interruption length** This represents the duration of all the recorded interruptions (in min) in a single observation. It is obtained by summing up length (duration) of the individual interruptions.
- **Day of activity** This represents the day the activity is performed, ranging from 0 to 6 for the 7 days of the week. This is relevant since individuals might have different activity routine for the different days of a week.
- **Weekend or weekday** This indicates if the activity is performed on a weekday or a weekend with 0 representing weekdays and 1 representing weekends.

Fig. 5 Result of the anomaly detection models for SmartNTU dataset **Fig. 6** Result of the anomaly detection models for CASAS HH111 dataset

For the computational models to achieve optimal results, the extracted features are normalised due to the variability in their respective scales.

Experimental Scenario and Result

The proposed approach for identifying the sources of anomaly applies to data already classified as outliers by the anomaly detection models described in Sect. 3. The four models presented namely, OC-SVM, iForest, LOF and RCE as well the ensemble approaches such as majority votes (MV), EDCV, EDVV and CNDE are evaluated on both the SmartNTU and CASAS HH111 datasets. The model contamination rate parameter (i.e. the rate of outliers in the training data) is set to a minimal value of 0.1 across the models.

The model training is performed using the activity data for the first 31 days, and the data for the remaining days are used for the validation. Data observation predicted as outliers are confirmed from the ground truth of the simulated and oversampled data as well as the outliers manually annotated.

The result of the abnormality detection models for the SmartNTU dataset is shown in Fig. 5, while that of the CASAS HH111 dataset is shown in Fig. 6. It can be seen that in both cases, the ensemble models outperformed the individual models as expected. For the separate individual models, OC-SVM performs better than iForest, LOF and RCE across the two datasets. Overall, the CNDE ensemble approach we proposed in [50] achieved the overall best performance with the accuracy of 98% and 96% for the SmartNTU and CASAS HH111 datasets, respectively.

These results indicate that the models are able to identify abnormalities in the sleeping activity, which may be due to the variability in the extracted features. To identify the features that are the likely source of the identified anomalies, the anomalous observations are explored using the proposed similarity measure.

Utilising the anomaly detection results presented in Figs. 5 and 6, the observations from both datasets that are predicted as outliers are further explored using the proposed similarity measure to identify the anomaly sources. After the oversampling of the datasets with SMOTE prior to the model training, the respective datasets contains approximately 100 anomalous observations with a relatively equal number of sources in 4 of the most discriminating features that are identified using principal component analysis (PCA). The four features are the start time, duration, interruption, and interruption length. The outlying observations are explored and the features of the data that are the likely anomaly sources are identified. Since multiple features can be identified as the source, identification of at least one anomaly source per activity is sufficient. In a scenario where multiples anomaly sources (features) are identified, a true

positive prediction is assigned if one of the predicted sources matches the ground truth.

To evaluate the performance of the similarity measure approaches, different metrics are measured in comparison to the ground truth. The ground truth of the anomaly sources is identified through the manual observation and annotation of the outliers on a feature by feature basis, and from the annotated data used for anomaly detection. For the artificially induced anomalies, this is easily achieved since the altered features of the data entries are known. The features that are altered to simulate an abnormally large variance for a given data entry are labelled as the sources of the anomaly for that entry. The four possible prediction outcomes for any given entry relative to its features are given as follows:

- **True positive (TP)** When a feature of the data entry labelled as the anomaly source is correctly predicted as the source by the model.
- **False positive (FP)** When a feature of the data entry that is not labelled as the anomaly source is wrongly predicted as the source by the model.
- **True negative (TN)** When a feature of the data entry labelled as not the anomaly source is correctly predicted as not the source by the model.
- **False negative (FN)** When a feature of the data entry labelled as the anomaly source is wrongly predicted as not the source by the model.

The precision, recall and accuracy of the similarity measure approaches are computed using the expressions in Eqs. 41–43. The precision and recall are computed on a feature by feature basis, while the accuracy is computed across all the features for the given dataset.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (41)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (42)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (43)$$

To further evaluate the proposed approaches, synthetic data are generated with outlying entries for the duration of 100 days. The features of the generated data are based on the distribution of the midpoints for the first and second quartile of the SmartNTU dataset. The simulated outliers sources are set as the third quartile of the data distribution of the features. These data are then utilised for the validation of the proposed approaches.

The result obtained for the similarity measure approaches based on Euclidean distance is presented in Fig. 7 showing

Fig. 7 Result for the similarity measures based on Euclidean distance on; **a** SmartNTU dataset, **b** CASAS HH111 dataset, **c** Synthetic dataset

the precision and recall for the different features. Similarly, Figs. 8 and 9 shows the results for the Chebyshev and Canberra distances. It can be seen that similarity measure approaches are able to identify the anomaly sources across the validation datasets.

The summary of the performance metrics are tabulated in Tables 3 and 4. It can be seen from Table 3 that the obtained result is slightly better for the SmartNTU dataset

compared to that of the CASAS HH111 dataset. However, the result for the Synthetic data surpasses that of the real datasets. This is expected since the synthetic data is carefully crafted to fit the model's requirement and therefore, should not be used as a measure of accuracy since the models can be biased towards the data. The average accuracy in Table 4 is grouped based on the distance function and based on the similarity measure approach across

Fig. 8 Result for the similarity measures based on Chebyshev distance on; **a** SmartNTU dataset, **b** CASAS HH111 dataset, **c** synthetic dataset

the two real datasets (SmartNTU and CASAS HH111) while excluding the result of the synthetic data in the computation due to the biased nature of the result. From the tabulated results, Euclidean distance function has the best performance while Canberra distance has the least. Additionally, the OASM approach outperforms the OOSM approach across the datasets. The OASM approach has an additional advantage of having a linear run-time complexity $O(n)$ over the OOSM approach that has a quadratic run-time complexity $O(n^2)$.

The identified anomaly sources (e.g. start time, duration etc.) may not have a significant meaning in the context of human activity recognition and behaviour modelling. To better understand the anomalies, a mapping is created from the extracted feature space to the real-life sources. This means that the extracted features are translated and attributed to real-life causes of the anomalies. For example, if the anomaly source is the start time, then the real-life source could be an indication of going to bed early or later than usual. Similarly, if the anomaly source is the duration, then this can be

Fig. 9 Result for the similarity measures based on Canberra distance on; **a** SmartNTU dataset, **b** CASAS HH111 dataset, **c** synthetic dataset

attributed to having insufficient sleep or oversleeping. The mapping may vary depending on the extracted features and their ground truth representation. A sample of the translated anomaly sources are summarised in Table 5 indicating the identified source and its real-life mapping observed from the data. To the best of our knowledge, no other exploratory approach exist for identifying anomaly sources in this context thereby making a comparative analysis of the proposed similarity measure nearly impossible.

Additionally, the datasets used for this validation consist of small data samples and fewer features. In a scenario where the dataset is significantly large or containing a large feature set, the proposed approach for estimating the similarity matrices may be time-consuming. This is because the similarity is measured for each outlying observation against the entire training samples across all the features. A heuristic approach for selecting fewer samples

Table 3 Overall accuracy of the similarity measure approaches

Dataset	Distance function	OOSM	OASM
SmartNTU	Euclidean	0.85	0.86
	Chebyshev	0.85	0.86
	Canberra	0.75	0.86
CASAS HH111	Euclidean	0.70	0.85
	Chebyshev	0.70	0.83
	Canberra	0.25	0.58
Synthetic	Euclidean	0.95	0.98
	Chebyshev	0.95	0.98
	Canberra	0.87	0.98

Table 4 Average accuracy of the similarity measures grouped by distance function and by similarity measure (SmartNTU and CASAS HH111 datasets)

Description	Measure	Average acc.
Based on distance function	Euclidean	0.82
	Chebyshev	0.81
	Canberra	0.61
Based on similarity measure	OOSM	0.68
	OASM	0.81

of the training data to be used for the similarity estimation can improve the efficiency of the methodology.

Conclusion

This paper proposes an exploratory approach for identifying sources of abnormalities in human activities based on a similarity measure approach. The proposed approach involves the pairwise distance measurement of the extracted features from human activity dataset in the form of a similarity matrix. A statistical approach is then applied to estimate the optimal threshold of the features and for the identification of the anomaly source with an achieved average accuracy of 0.82. As part of the future work plan, the proposed methodology will undergo extensive evaluation for the identification of multiple anomaly sources. Furthermore, approaches for optimising the data selection process for the similarity matrix computation will be explored. In addition to the identification of anomaly sources, the possibility of utilising the proposed approach for anomaly detection will be explored.

Acknowledgements This research project is supported by Petroleum Technology Development Fund and Nottingham Trent University.

Compliance with Ethical Standards

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Table 5 Identified anomaly sources for SmartNTU and CASAS HH111 dataset

Dataset	Day	Data feature	Interpreted source	Detailed description
SmartNTU	34, 46 and 55	Duration	Less sleeping	The individual has insufficient sleep compared to his/her usual sleep duration
	51	Duration	Over sleeping	The individual sleeps for a longer duration than usual
	39 and 69	Interruption	Interrupted sleep	The individual experienced several unusual interruptions while sleeping
	59	Start time	Going to bed late	The individual goes to bed late compared to his/her usual bedtime
CASAS HH111	50	Duration	Less sleeping	The individual has insufficient sleep compared to his/her usual sleep duration
	39 and 49	Interruption	Interrupted sleep	The individual experienced several unusual interruption
	37, 44 and 46	Interruption length	Interrupted sleep	The individual experienced a long transition during sleeping time. This may be an indication of performing an activity in the middle of the night
	32, 33 and 47	Start time	Going to bed early	The individual goes to bed earlier than usual

References

1. Acampora G, Cook DJ, Rashidi P, Vasilakos AV. A survey on ambient intelligence in healthcare. *Proc IEEE*. 2013;101(12):2470–94.
2. Alberdi A, Weakley A, Schmitter-Edgecombe M, Cook DJ, Aztiria A, Basarab A, Barrenechea M. Smart home-based prediction of multidomain symptoms related to Alzheimer's disease. *IEEE J Biomed Health Inf*. 2018;22(6):1720–31.
3. Anderez DO, Lotfi A, Pourabdollah A. Eating and drinking gesture spotting and recognition using a novel adaptive segmentation technique and a gesture discrepancy measure. *Expert Syst Appl*. 2020;140:112888.
4. Manocha A, Singh R. An intelligent monitoring system for indoor safety of individuals suffering from Autism Spectrum Disorder (ASD). *J Ambient Intell Human Comput*. 2019. <https://doi.org/10.1007/s12652-019-01277-3>.
5. Aranendi AA, Weakley A, Goenaga AA, Schmitter-Edgecombe M, Cook DJ. Automatic assessment of functional health decline in older adults based on smart home data. *J Biomed Inform*. 2018;81:119–30.
6. Arifoglu D, Bouchachia A. Activity recognition and abnormal behaviour detection with recurrent neural networks. *Procedia Comput Sci*. 2017;110:86–93.
7. Arifoglu D, Bouchachia A. Detection of abnormal behaviour for dementia sufferers using Convolutional Neural Networks. *Artif Intell Med*. 2019;94:88–95.
8. Breunig MM, Kriegel H-P, Ng RT, Sander J. LOF: identifying density-based local outliers. *SIGMOD Rec*. 29, 2 (June 2000), pp 93–104. <https://doi.org/10.1145/335191.335388>
9. Capela NA, Lemaire ED, Baddour N, et al. Evaluation of a smart-phone human activity recognition application with able-bodied and stroke participants. *J NeuroEngineering Rehabil*. 2016;13:5. <https://doi.org/10.1186/s12984-016-0114-0>.
10. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. *J Artif Intell Res*. 2002;16:321–57.
11. Chernbumroong S, Cang S, Atkins A, Yu H. Elderly activities recognition and classification for applications in assisted living. *Expert Syst Appl*. 2013;40(5):1662–74.
12. Cook DJ, Crandall AS, Thomas BL, Krishnan NC. Casas: A smart home in a box. *Computer*. 2013;46(7):62–9.
13. Cook DJ, Krishnan N. Mining the home environment. *J Intell Inf Syst*. 2014;43(3):503–19.
14. Cook DJ, Schmitter-Edgecombe M, Dawadi P. Analyzing activity behavior and movement in a naturalistic environment using smart home techniques. *IEEE J Biomed Health Inf*. 2015;19(6):1882–92.
15. Dai X, Bikdash M. Distance-based outliers method for detecting disease outbreaks using social media. *SoutheastCon*. 2016;2016:1–8.
16. Dawadi PN, Cook DJ, Schmitter-Edgecombe M. Automated cognitive health assessment from smart home-based behavior data. *IEEE J Biomed Health Inf*. 2016;20(4):1188–94.
17. Dib G, Karpenko O, Koricho E, Khomenko A, Haq M, Udpa L. Ensembles of novelty detection classifiers for structural health monitoring using guided waves. *Smart Mater Struct*. 2018;27(1):15003.
18. Fahad LG, Rajarajan M. Anomalies detection in smart-home activities. In: *Proceedings of the 2015 IEEE 14th international conference on machine learning and applications (ICMLA)*, Miami, FL, 2015, pp 419–422. <https://doi.org/10.1109/ICMLA.2015.195>.
19. Fan X, Xie Q, Li X, Huang H, Wang J, Chen S, Xie C, Chen J. Activity recognition as a service for smart home: ambient assisted living application via sensing home. In: *Proceedings of the 2017 IEEE International Conference on AI & Mobile Services (AIMS)*, Honolulu, HI, 2017, pp. 54–61. <https://doi.org/10.1109/AIMS.2017.29>.
20. Forkan ARM, Khalil I, Tari Z, Fofou S, Bouras A. A context-aware approach for long-term behavioural change detection and abnormality prediction in ambient assisted living. *Pattern Recogn*. 2015;48(3):628–41.
21. Gong M, Zeng H, Xie Y, Li H, Tang Z. Local distinguishability aggrandizing network for human anomaly detection. *Neural Netw*. 2020;122:364–73.
22. Hela S, Amel B, Badran R. Early anomaly detection in smart home: a causal association rule-based approach. *Artif Intell Med*. 2018;91:57–71.
23. Hernandez N, Lundström J, Favela J, McChesney I, Arnrich B. Literature review on transfer learning for human activity recognition using mobile and wearable devices with environmental technology. *SN Comput sci*. 2020;1:66. <https://doi.org/10.1007/s42979-020-0070-4>.
24. Hoque E, Dickerson RF, Preum SM, Hanson M, Barth A, Stankovic JA. Holmes: a comprehensive anomaly detection system for daily in-home activities. In: *Proceedings of the 2015 international conference on distributed computing in sensor systems*, Fortaleza, 2015, pp. 40–51. <https://doi.org/10.1109/DCOSS.2015.20>.
25. Howedi A, Lotfi A, Pourabdollah A. An entropy-based approach for anomaly detection in activities of daily living in the presence of a visitor. *Entropy*. 2020;22:845.
26. Hu LY, Huang MW, Ke SW, et al. The distance function effect on k-nearest neighbor classification for medical datasets. *Springer-Plus*. 2016;5:1304. <https://doi.org/10.1186/s40064-016-2941-7>.
27. Jakkula V, Cook DJ. Detecting anomalous sensor events in smart home data for enhancing the living experience. In: *Proceedings of the 7th AAAI Conference on artificial intelligence and smarter living: the conquest of complexity*. AAAI Press; 2011. pp. 33–37.
28. Jakkula V, Cook DJ, Crandall AS. Temporal pattern discovery for anomaly detection in a smart home. In: *Proceedings of the 2007 3rd IET international conference on intelligent environments*, Ulm, 2007, pp 339–345. <https://doi.org/10.1049/cp:20070390>
29. Liu FT, Ting KM, Zhou Z-H. Isolation-based anomaly detection. *ACM Trans Knowl Discov Data*. 2012;6(1):3. <https://doi.org/10.1145/2133360.2133363>.
30. Lotfi A, Langensiepen C, Mahmoud SM, Akhlaghinia MJ. Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour. *J Ambient Intell Humaniz Comput*. 2012;3(3):205–18.
31. Damaševičius R, Vasiljevas M, Šalkevičius J, Woźniak M. Human activity recognition in AAL environments using random projections. *Comput Math Methods Med*. 2016;4073584. <https://doi.org/10.1155/2016/4073584>.
32. Meng L, Miao C, Leung C. Towards online and personalized daily activity recognition, habit modeling, and anomaly detection for the solitary elderly through unobtrusive sensing. *Multimedia Tools Appl*. 2017;76(8).
33. Ni Q, GarcíHernando AB, De la Cruz IP. The elderly's independent living in smart homes: A characterization of activities and sensing infrastructure survey to facilitate services development. *Sensors*. 2015;15(5):11312–62.
34. Novák M, Biñas M, Jakab F. Unobtrusive anomaly detection in presence of elderly in a smart-home environment. In: *2012 ELEKTRO, Rajeck Teplice*, 2012, pp. 341–344. <https://doi.org/10.1109/ELEKTRO.2012.6225617>
35. Novák M, Jakab F, Lain L. Anomaly detection in user daily patterns in smart-home environment. *J Sel Areas Health Inf*. 2013;3(6):1–11.
36. Pasillas-Díaz JR, Ratté S. An unsupervised approach for combining scores of outlier detection techniques, based on similarity measures. *Electron Notes Theor Comput Sci*. 2016;329:61–77.

37. Pazhoumand-Dar H, Armstrong LJ, Tripathy AK. Detecting deviations from activities of daily living routines using kinect depth maps and power consumption data. *J Ambient Intell Human Comput.* 2020;11:1727–47. <https://doi.org/10.1007/s12652-019-01447-3>.
38. Pimentel MA, Clifton DA, Clifton L, Tarassenko L. A review of novelty detection. *Sig Process.* 2014;99:215–49.
39. Rahman MM, Islam MM, Ahmmed S, Khan SA. Obstacle and fall detection to guide the visually impaired people with real time monitoring. *SN Comput Sci.* 2020;1(4):219.
40. Rashidi P, Mihailidis A. A survey on ambient-assisted living tools for older adults. *IEEE J Biomed Health Inf.* 2013;17(3):579–90.
41. Ronao CA, Cho SB. Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Syst Appl.* 2016;59:235–44.
42. Rousseeuw PJ, Hubert M. Anomaly detection by robust statistics. *WIREs Data Min Knowl Discov.* 2018;8(2):e1236.
43. Schölkopf B, Platt JC, Shawe-Taylor JC, Smola AJ, Williamson RC. Estimating the Support of a High-Dimensional Distribution. *Neural Comput.* 2001;13(7):1443–71. <https://doi.org/10.1162/089976601750264965>.
44. Shreyas DG, Raksha S, Prasad BG. Implementation of an anomalous human activity recognition system. *SN Comput. Sci.* 2020;1:168. <https://doi.org/10.1007/s42979-020-00169-0>.
45. Thapliyal H, Kumar Nath R, Mohanty SP. Smart home environment for mild cognitive impairment population: solutions to improve care and quality of life. *IEEE Consum Electron Mag.* 2018;7(1):68–76.
46. Thomopoulos NT. *Statistical distributions.* New York: Springer International Publishing; 2017.
47. Tong Y, Chen R, Gao J. Hidden state conditional random field for abnormal activity recognition in smart homes. *Entropy.* 2015;17(3):1358–78.
48. Weller-Fahy DJ, Borghetti BJ, Sodemann AA. A survey of distance and similarity measures used within network intrusion anomaly detection. *IEEE Commun Surv Tutor.* 2015;17(1):70–91.
49. Yahaya SW, Langensiepen C, Lotfi A. Anomaly detection in activities of daily living using one-class support vector machine. In: Lotfi A, Bouchachia H, Gegov A, Langensiepen C, McGinnity M (eds) *Advances in computational intelligence systems.* UKCI 2018. *Advances in intelligent systems and computing*, vol. 840. Cham: Springer; 2019. https://doi.org/10.1007/978-3-319-97982-3_30.
50. Yahaya SW, Lotfi A, Mahmud M. A consensus novelty detection ensemble approach for anomaly detection in activities of daily living. *Appl Soft Comput.* 2019;83:105613.
51. Zerkouk M, Chikhaoui B. Spatio-temporal abnormal behavior prediction in elderly persons using deep learning models. *Sensors.* 2020;20:2359.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.