Chapter 7

Temporal Clustering for Behavior Variation and Anomaly Detection from Data Acquired Through IoT in Smart Cities

Vladimir Urosevic, Ana Kovacevic,
Firas Kaddachi and Milan Vukicevic

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.75203

Abstract

In this chapter, we propose a methodology for behavior variation and anomaly detection from acquired sensory data, based on temporal clustering models. Data are collected from five prominent European smart cities, and Singapore, that aim to become fully "elderly-friendly," with the development and deployment of ubiquitous systems for assessment and prediction of early risks of elderly Mild Cognitive Impairments (MCI) and frailty, and for supporting generation and delivery of optimal personalized preventive interventions that mitigate those risks, utilizing smart city datasets and IoT infrastructure. Low level data collected from IoT devices are preprocessed as sequences of activities, with temporal and causal variations in sequences classified as normal or anomalous behavior. The goals of proposed methodology are to (1) recognize significant behavioral variation patterns and (2) support early identification of pattern changes. Temporal clustering models are applied in detection and prediction of the following variation types: intra-activity (single activity, single citizen) and inter-activity (multiple-activities, single citizen). Identified behavioral variations and anomalies are further mapped to MCI/frailty onset behavior and risk factors, following the developed geriatric expert model.

Keywords: temporal clustering, IoT, smart cities, behavior recognition, anomaly detection

1. Introduction

Frailty and Mild Cognitive Impairment are common and inevitable conditions in the elderly citizen population defined as premature or accelerated physical and mental declines. These conditions are often an early indicator of more severe states, such as Alzheimer's disease. Control (delaying or decelerating) of the onset and progression of MCI/frailty is becoming one of the major tasks of global efforts in maintaining the functional independence and quality of life of the globally growing elderly population. In 2016, for the first time in history, estimated majority of the world population can expect to live into their sixties and beyond. Beside the global initiatives, strategies and action plans on healthy ageing conducted by the relevant key organizations such as World Health Organization (WHO), or United Nations (UN), the growing market trend for the so-called "silver economy" sector is booming. The increase of population aged 65 and over is projected to reach 28.1% of the whole population in the EU by 2050, they have a spending power higher than the generation segment aged 18 to 39, and they account for approx. 60% of total expenditures in the US and 50% in the UK, generating demand for new services and products, ranging from personalized care to age-friendly technologies and other solutions that enable the maintenance and prolongation of healthy, independent lives. Technologies and systems supporting innovative ways of influencing people's behavior and lifestyles at all ages also present a significant economic and business opportunity.

Geriatric practice has in this aim utilized different standardized instruments based on traditional data collection methods (administration of questionnaires, meter-based measurement or direct observation in controlled conditions) which are in most cases intrusive and demand citizens' presence in geriatric centers and a lot of time for data collection. More importantly, these methods do not enable real time monitoring of behavioral changes (e.g., data from questionnaires are collected on semi-annual or annual intervals) and thus prevent predictive and preventive interventions. Finally, data collected from such method is often subjective or incomplete.

With the goal to overcome the stated drawbacks**, many technological instruments and methods emerged**, aiming to automate as much as possible the detection and mitigation of behavior deteriorations and anomalies. Particularly, the recent development and expansion of wearable technologies/devices and Internet of Things (IoT) has enabled the build-up of infrastructures of smart devices that collect vast and heterogeneous volumes of various sensory data in smart cities. These social and technological infrastructural advances potentiate the public health and prevention aspects of smart cities, transforming the urban public health from a reactive to a predictive system. In the specific area of support for (active and healthy) ageing, the transformation and progress direction of particular interest is the expansion of concept of **ambient-assisted environments**, from currently predominant implementations in residential and social indoor spaces (homes, elderly care/community centers) to outdoor and public environments. However, there is still a large gap between potential and actual IoT data exploitation because of many challenges that have to be overcome before putting data driven predictive and preventive models in geriatric or healthcare practice [1, 2].

Research presented in this paper is mainly the part of **City4Age project** (www.city4ageproject.eu) that develops age-friendly **Cities and Environments in** deployments in six different prominent pilot smart cities in EU and beyond—Athens, Birmingham, Lecce, Madrid,

http://dx.doi.org/10.5772/intechopen.75203

Montpellier and Singapore. **The main goal** of this research and City4Age project is to develop a framework for predictive and preventive risk control of Frailty and MCI, as one of the core system infrastructure assets of age-friendly cities. **Specific goals** are development of methods based on smart cities IoT data for early risk detection and enrichment of traditional geriatric instruments. In order to achieve these goals, we are faced with several **challenges**: (1) identification and characterization of temporal behavioral patterns from sensor data, (2) Identification of behavior changes (transitions) and (3) Anomalous behavior and anomalous data detection.

Given that IoT data are collected from smart devices in form of unlabeled data streams, for initial behavior variation detection models, unsupervised machine learning techniques have to be employed. From data analytics point of view, behavior can be defined as alternating pattern of sequences of activities. Based on this definition, clustering is identified as natural technique for behavioral pattern recognition, change and anomaly detection. Clustering techniques that allow grouping objects into homogenous groups where objects in the same group are similar (intra-cluster distance is low) and objects between groups are dissimilar (inter-cluster distance is high). For building cluster models, we employed Hidden Markov models (HMMs), since they allow direct modeling of time series, provide framework for anomaly detection and have high degree of interpretability. Interpretability is very important property for incorporation of data driven models in healthcare practice and integration with domain knowledge of geriatricians.

Contributions of this chapter are twofold: (1) we propose a framework for behavior characterization, change and anomaly detection in IoT data in smart cities environment and (2) we provide first experimental evidence of usefulness of data driven modeling of behavioral data on collected City4Age IoT data.

2. State-of-the-art

World Health Organization (WHO) had recognized the importance as well as human and economic impact of age-friendly environments and launched the age-friendly Cities and Communities Programme that introduced the terms in 2006/2007, as the foundation initiative aimed for local and metropolitan governance and development levels. The European Commission (EC) supports the pursues of goals and objectives of age-friendly environments and sustainable development by numerous different instruments, primarily through R&D funding programs such as Horizon 2020 or specialized Active and Assistive Living (AAL) Programme. Important higher-level EC initiatives that foster innovation and concentrate stakeholder efforts are the recently established:

- European Innovation Partnership on Smart Cities and Communities (EIP on SCC), involving almost 400 committed cities and other partners, with a marketplace of specialized initiatives, solutions and tools.
- European Innovation Partnership on active and healthy ageing (EIP on AHA), first established EIP, in 2011, with specialized dedicated groups A3 for Functional decline & frailty, and D4 for age-friendly environments, among others.

These efforts increased research efforts in the area of smart city IoT data analytics through different projects.

Geriatric practice has in this aim utilized different standardized instruments based on traditional data collection methods (administration of questionnaires, meter-based measurement or direct observation in controlled conditions) and quantification and categorization of functional domains of daily life behavior and known frailty/MCI risk factors, such as Lawton IADL scale, Mini-Mental State Examination (MMSE), Fried Frailty Index, Nottingham Extended Activities of Daily Living, and numerous others. A comprehensive summary of such traditional generally **psychogeriatric instruments** and methods is provided in [3]. These instruments have evident major drawbacks, of late detection and problem identification (analysis and interpretation of questionnaires or conducting of exams can span intervals of months), and being generally ineffective, possibly subjective to a high degree, and costly for deployment.

The **City4Age project** (www.city4ageproject.eu), funded through the mentioned EC Horizon 2020 programme, is one of the pioneering efforts acting as a bridge between the mentioned two European Innovation Partnerships, EIP on SCC and EIP on AHA, contributing to specific and shared objectives and involving the committed participants from both Partnerships. The primary aim of the project is to enable fully Ambient Assisted age-friendly cities, through development and deployment of a range of ICT tools and services that will improve the unobtrusive early detection of MCI/frailty risks from heterogeneous IoT and smart city data sources at homes or on the move within the city, comprising the research and development work performed and results presented in this chapter as part of the work on the Data Analytics Platform. Coupled with the appropriate interventions—the developed tools will mitigate the detected risks as secondary aim. The developed system and components are being validated through in-situ deployments in six pilot smart cities.

Besides the City4Age project, there are numerous other related efforts in development of IoT driven systems for maintaining the functional independence and quality of life of the globally growing elderly population, or in development of data-driven health-related behavior recognition systems or platforms. Some of the recent relevant ones are the following:

The ActivAGE project (www.activageproject.eu), started in January 2017 and likewise funded through the Horizon 2020 programme, is a European multi-centric large-scale pilot on Smart Living Environments. The main objective is to build the first European IoT ecosystem across nine Deployment Sites (DS) in seven European countries, reusing and scaling up underlying open and proprietary IoT platforms, technologies and standards, and integrating new interfaces needed to provide interoperability across these heterogeneous platforms, that will enable the deployment and operation at large scale of active & healthy ageing IoT-based solutions and services.

Participatory Urban Living for Sustainable Environments (PULSE) project (www.pulseproject.eu), started in January 2016 and likewise funded through the Horizon 2020 programme, harvests open city data, and data from health systems, urban and remote sensors, and personal devices, to enable evidence-driven and timely management of public health events and processes, leveraging diverse data sources and big data analytics to transform urban public health from a reactive to a predictive system, and from a system focused on surveillance to an

inclusive and collaborative system supporting health equity. The clinical focus of the project is on chronic respiratory (asthma) and metabolic diseases (type 2 Diabetes), developing risk stratification models based on risk factors in each urban location (pilot deployment in five global cities—Barcelona, Birmingham, Paris, New York and Singapore), taking account of biological, behavioral, social and environmental risk factors, community resilience and well-being in cities.

IGERT project, ended in 2016, funded by the US National Science Foundation grant, comprises a multi-disciplinary doctoral training programme focused on designing and studying health-assistive smart environments, with particular emphasis on automatic monitoring and analysis of human health and behavior, unsupervised data-driven detection of activity/behavior and lifestyle changes, potential simulation/prediction of human behavior and activities, and enhancement of human physical and cognitive abilities [4].

Recently exhaustive and comprehensive reviews about temporal clustering algorithms and applications are published [5–7] and thus we will focus here only on the concepts that are closest to this research. As said behavior recognition, change and anomaly detection can be modeled naturally with clustering algorithms. Clustering techniques that allow grouping objects into homogenous groups where objects in the same group are similar (intra-cluster distance is low) and objects between groups are dissimilar (inter-cluster distance is high). Since the definition of clustering is based on the notion of similarity it is utterly important to define the notion of similarity and types of similarity measures. Unlike stationary data, time series have several aspects of similarity [7]:

Similarity in time: this is the simplest form of similarity with the assumption that instances that are close in time have similar values. This is a naïve assumption and is used for benchmark purposes in most of the cases.

Similarity in shape: these similarities disregard the time of occurrence of patterns. Using this definition, clusters of time-series with similar patterns of change are constructed regardless of time points—for example, extracting groups of elderly citizens who have a common pattern in their visits to pharmacy, regardless when these pharmacy visits occur in time-series. Dynamic time warping (DTW) is one of the mostly used dissimilarity measures of this type.

Structural similarity: the types of metrics used to find similarities in changes of time series. This is done by building models like AutoRegressive Moving Average (ARMA) or Hidden Markov models (HMMs), and the similarity is measured between parameters of models. In case of intra-activity behavior variations, structural similarity could recognize patterns such as: after three weeks of decrease of outdoor_time, citizen has registered increased outdoor_time, and this behavior is repeated every month.

3. City4Age analytic framework

Main challenges for the Data Science and Analytics in the related research in City4Age coincide with main generic challenges for the potential of collected IoT data from smart cities for health (and other personal) monitoring—volume and diversity of collected data is huge and

promising, but the development of formalized and applicable knowledge and learning models is lagging with adequate potential for interpretation, classification and exploitation. At the same time, the unobtrusively acquired dataset for each specific person over time is often sparse, incomplete and erroneous, and with high degree of variation caused by temporary sensor imprecisions or influence of external factors beyond the sensing or modeling scope. High-level City4Age Analytics and detection process flow is depicted on **Figure 1**. City4Age has from the beginning adopted the combined hybrid knowledge- and data-driven approach, with the initial contribution of the knowledge-based approach turning out somewhat overestimated in the meantime due to the issues mentioned above and consequent non-deterministic and volatile semantic integrity of "known" or presumed universal geriatric causalities and concepts. The main focus is thus currently on the **data driven** behavior change variation recognition and characterization, analysis of relative changes in time series data for each specific person since the start of the pilot monitoring and determining baseline referent points, values and features (individual geriatric care analytics), and subsequently on discovering correlations and underlying features and interdependencies in the complete studied and monitored populations and clusters and groups within it (group exploratory analytics), starting from minimal initial domain model knowledge.

Majority of the functional domains and parameters of daily life behavior and geriatric risk are nevertheless known and established, and formalized in the City4Age hierarchic computational model of geriatric behavior and risk [8].

The main model constructs and variables are based on the notions of Geriatric factors (GEF), representing monthly behavior characterizations from all various functional behavioral domain variables and known MCI/frailty risk indicators, on unified Likert scale, with 1 denoting the least favorable and five the most favorable behavior with respect to MCI and Frailty risk, a common and standard adopted representation in geriatric practice and many of the used traditional instruments and questionnaires. GEF are further structured on several hierarchic levels of decomposition (GES—geriatric sub-factors, GFG—geriatric factor groups), and can be synthesized or derived from "Measures," native numeric values generated by the

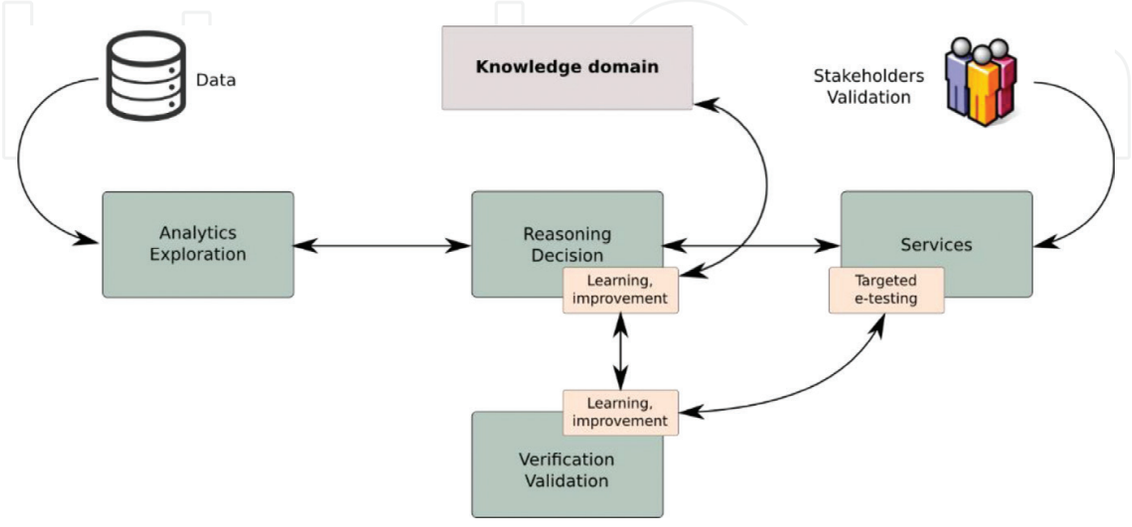

Figure 1. High-level City4Age analytics and detection process flow.

http://dx.doi.org/10.5772/intechopen.75203

various sensing technologies and methods for collecting data (e.g., daily number of walked steps, weekly number of visits to relatives, daily time in seconds spent in public transport, etc.), as exemplified on the diagram above. Measures are analyzed and processed using various algorithmic techniques and/or methods, some of which are the clustering and grading algorithms described in this chapter, but others have also been tried and tested, and therein is the flexibility and scalability of the model and the Analytics Framework, supporting the registration of various algorithmic methods through metadata and deploying them on the "detection" variables (GFG, GEF, GES, Measures) on various model hierarchy and derivation levels. Example of network representation of GES, GEF and Measures is presented on the diagram on **Figure 2** below. Relations between nodes shown on the diagram are not fixed/persistent, can variate according to different model configurations in different cities, or adaptively according to the results of the data-driven detection.

The results of the data-driven detection are in turn used for expanding and building-up the domain knowledge base. The structure (ontologies and semantics) and mechanisms for this are established [9] are in parallel ongoing development, and will be still more intensely in the future work, in the scope and frame of City4Age contributions and breakthroughs in establishment of data-driven geriatrics. The unobtrusively acquired temporal dataset on individual level, currently being acquired for each single elderly person, is highly likely to expand with the increase and improvement of deployment scope and reliability of data acquisition and detection infrastructure and technologies.

Figure 2. Example network representation of the City4Age geriatric model main constructs: Measures (purple), GES (green), and GEF (red).

4. Hidden Markov models for behavioral modeling of smart cities IoT data

As discussed, the main tasks of City4Age analytic framework are recognition of behavioral patterns, behavior changes (transitions) in time and anomaly detection. Additionally, models derived from data should be interpretable in order to integrate data driven insights with domain knowledge expertise. Hidden Markov models (HMMs) provide a framework for all main tasks and thus we employed these models for behavior variation analyses. Additionally, HMMs allow prediction of identified behavioral patterns in future and this adds predictive and preventive component in analytic framework. Here we will consider first order HMMs where each temporal state depends only on one previous state. This is strong assumption, but allows development of scalable models and real-time inference. **Figure 3** describes first order Markov chain where each state x depends on previous state (x-1) and observed data (y).

Hidden Markov models can be explained as total probability of X and Y by following formulae:

$$p(X,Y) = p(x_1)\prod_{T-1}^{t=1} p(x_{t+1} \mid x_t)\prod_{T}^{t'=1} p(y_{t'} \mid x_{t'}) \tag{1}$$

where $p(y_{t'} \mid x_t)$ represents observation probability, while $p(x_{t+1} \mid x_t)$ prepresents transition probability.

In our case observations are series of IoT sensory data while hidden states represent categorized, homogenous series parts (that wil.l be characterized as behavioral patterns or behaviors). This is why we use Gaussian HMMs that characterize states with Gaussian distributions. This is depicted on **Figure 4**.

Each HMM model is thus constituted from three elements:

1. Prior probability distribution of hidden states (vector π) that describes how frequently each state occurs in general.
2. Transition matrix ($A_{i,j}$) that describe the transition probabilities from one state to another.
3. Probability distribution functions (one for each state) with corresponding parameters. In our case Gaussian distributions are modeled and thus means and standard deviations are used for definition of hidden state (behavior) probability distribution. HMMs allow modeling of discrete data too, but in that case probability distributions are represented by conditional distributions.

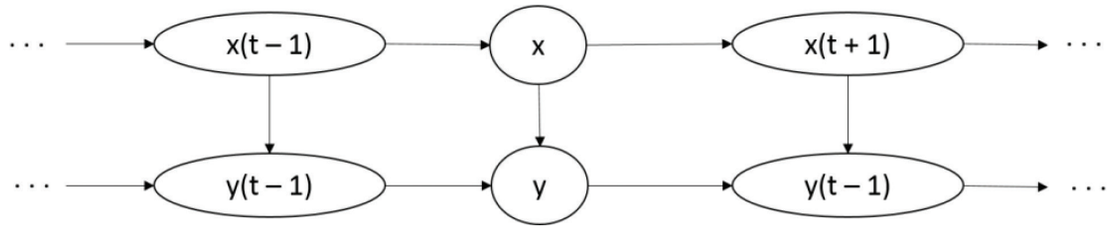

Figure 3. First-order Markov chain.

http://dx.doi.org/10.5772/intechopen.75203

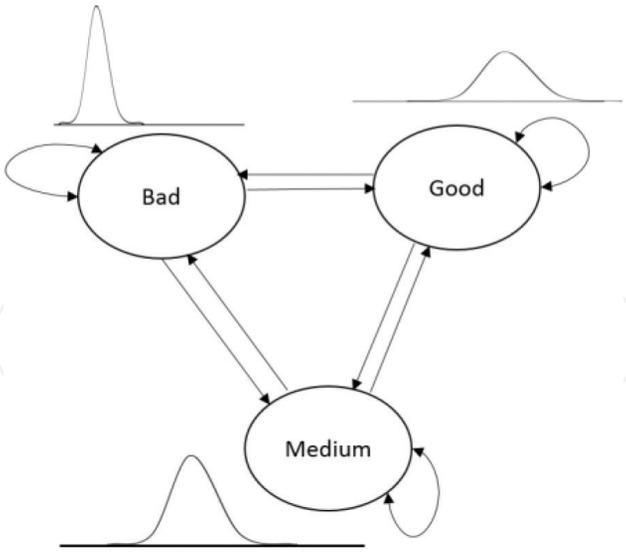

Figure 4. Behavior modeling with Gaussian HMMs.

Based on HMM definition, we can work on following tasks [10]:

Training—Learning parameters of HMM (A, B, and the prior distribution π), given a training sequence of observations $y_1, y_2, \ldots, y_T$. By solving this task, we will be able to characterize behavioral patterns (distributions). This task is solved by forward-backward algorithm.

Decoding—given an observation sequence and an HMM, determine the most probable hidden state (behavior) sequence. We used this task for state prediction and model evaluation. This task is solved by Viterbi (backward algorithm).

Likelihood—Calculation of probability that given sequence originates from given HMM model. In this research, we did not work on this task, since we built personalized behavioral models, but it will be used in later stages of the project when we will model behavior of groups of care recipients.

5. Framework for behavioral pattern recognition and change detection

Based on definition of behavior as pattern of sequences of activities and corresponding measure values, clustering algorithms emerge as natural algorithmic approach for behavioral pattern recognition and change detection. In City4Age setting, inputs for clustering algorithms are time series. These time series can be represented by values of activity measures, GES, GEF or Geriatric Score of care recipients. Based on time series, temporal clustering algorithms can identify patterns (similar time series values in consecutive time-steps) that are repeated over time. We characterize these patterns as behaviors and transition between patterns, behavior changes. Very important component in derivation of GES and GEF from activity measures are numerical indicators (NUIs). NUIs represent aggregations (e.g., mean, std., trend, etc.) of activity measure values on monthly level. This granularity level is convenient since it allows direct

conversion of NUIs to GES and GEF that are interpretable to geriatricians. However, monthly statistics in some cases do not capture important within month variations in time series.

This is why, in contrast to NUIs, clusters are not restricted to monthly level. Depending on input data, clusters can be identified on daily or monthly level. For example, if number_of_steps activity measure is clustered over time, model can identify similar groups of daily values: days with high values (i.e. average of 3000 steps with standard deviation of 200 steps) and days with low values (i.e. average of 600 steps with standard deviation of 100 steps). Similarly, cluster models can identify patterns of series of GES or GEF. For example: care recipient have periods of time where motility have average motility value of 4.1 with standard deviation of 0.2. So, behavioral patterns encapsulated in cluster models provide characterization of behavior on finer grade than monthly level. Additionally, the level of granularities does not have to be defined in advance (e.g., weeks).

For example, in first 22 days of January, care recipient had high values and high variability of number_of_steps, but in next 12 days he or she had low values with low variability. Even though, behavioral patterns described by clusters are not necessarily aligned with monthly representations of NUIs, GES and GEF, they can be exploited for definition of new NUIs that will capture within month behavior variations (e.g., care recipient showed improved behavior in last eight days of a month). NUIs based can be further graded as described in previous sub-section. Based on previous examples it is intuitively clear that clusters (behavior patterns) encapsulate smaller variations of time series and allows data driven discretization and characterization of discrete categories. This discretization allows easier inspection of behavioral changes (than observing unclustered series with many variations) and thus, results of clustering (cluster labels) can be directly represented on City4Age interactive dashboards or used as NUIs for derivations of GES and GEF (**Figure 5**). This way Geriatricians can access visualizations of natural groupings among series data, and label behaviors (e.g., normal, bad or good or on the Likert scale) or revise data driven grading of clusters.

It is important to emphasize that grouping of activities and/or citizens will play an important role in extending of City4Age interactive dashboards, with mentioned easier identification

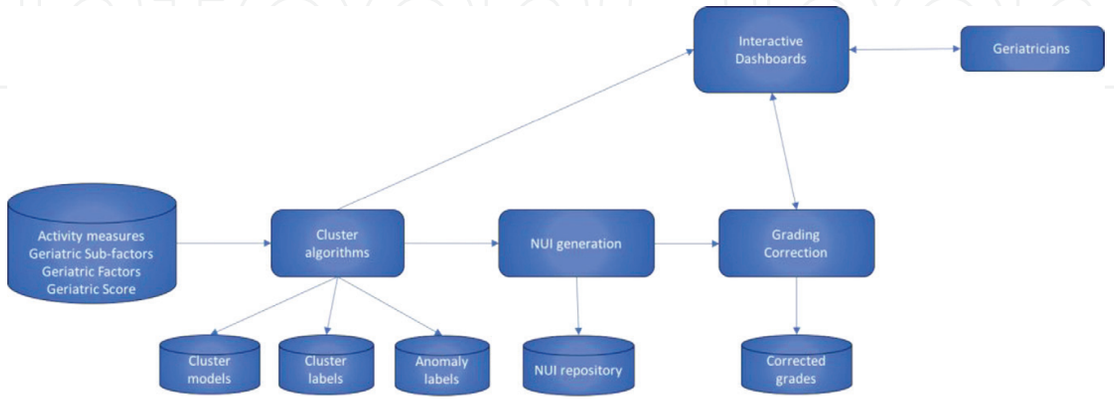

Figure 5. Cluster model flow.

http://dx.doi.org/10.5772/intechopen.75203

and labeling of patterns in streaming data. The labels will allow development of supervised models and further automation of City4Age analytics processes and improvement of alarming systems.

6. Experiments

In order to evaluate proposed framework we conducted experiments collected from City4Age Pilot sites. The main goal of the experimental evaluation was to evaluate usefulness of HMM models for behavior recognition, behavior change and anomaly detection in context of City4Age IoT data. We will describe process of data collection, modeling and evaluation in following text.

6.1. Data collection and preparation

Data used in experiments originate from the Birmingham Pilot. Data have been acquired by monitoring 3 Care recipients during a 6-month period (from January to July 2017, and ongoing). Sensory data are collected using Nokia Steel (e.g., Withings Activité) smartwatches. Nokia Steel tracks the following activities: sleep cycle, movements tracking, walked steps and distance, burned calories, elevation, heart rate, and optionally SP02 (peripheral capillary oxygen saturation, an estimate of the amount of oxygen in the blood, taken with additional pulse oximeter). Integration of sensor data with City4Age analytics platform is described in the following text. The proximity positioning data are gathered through smartphone BLE transceiver and relayed through the smartphone 4G connection to the City4Age Platform. Nokia/Withings API is used for initial pre-processing step on the sleep, activity, and other data obtained from the smartwatches, before sending to the City4Age Platform. So, input for building clustering algorithms in this research was sets of activity measures for each citizen. Summary of observed activity measures is presented in **Table 1**.

6.2. Experimental setup

The main goal of our experiments was to show that HMM models can be efficiently used for behavioral pattern recognition, behavior change detection and anomaly detection. In order to achieve this goal we faced several challenges: identification of adequate model evaluation (selection) measure, identify optimal number of behavioral states for each care recipient and each activity and finally to characterize identified behaviors (clusters or behavioral patterns). Since HMM models cannot implicitly learn optimal number of hidden states, we built HMM models with varying number of clusters (in the range 2–10) for each care recipient and each activity. Additionally, since there is no consensus for evaluation of cluster models in unsupervised setting, each model was evaluated with log likelihood, BIC and AIC evaluation measures. So setting we conducted 810 experiments in total (3 care recipients × 10 activities × 9 variations of state numbers × 3 evaluation measures). Each experiment lasted for 15–24 s (including learning and evaluation). Since HMM is one of the most scalable algorithm from

Geriatric sub-factor	Activity	Measure unit
Walking	WALK_STEPS	# of steps
	WALK_DISTANCE	meters
Quality of sleep	SLEEP_LIGHT_TIME	seconds
	SLEEP_DEEP_TIME	seconds
	SLEEP_AWAKE_TIME	seconds
	SLEEP_WAKEUP_NUM	seconds
	SLEEP_TOSLEEP_TIME	seconds
Physical activity	PHYSICALACTIVITY_SOFT_TIME	seconds
	PHYSICALACTIVITY_MODERATE_TIME	seconds
	PHYSICALACTIVITY_INTENSE_TIME	seconds
	PHYSICALACTIVITY_CALORIES	# of calories

Table 1. Observed activity measurements.

Probabilistic Graphical models family (it is frequently used for signal processing and speech recognition) it allows adaption for much larger series as City4Age streaming data arrives. After building models, they are applied to activity measure time series for each citizen and each activity. In this way we labeled each time point with cluster (behavioral pattern or state) assignment. When scoring HMM models, probabilities that time point originates from cluster distributions are identified and largest probabilities are stored for anomaly detection purposes. Experimental setup is implemented in Python. Hmmlearn library is used for building HMM models while Pandas DataFrame is used for data manipulation. All experiments are conducted on a testing cloud comprising three servers with quad-core Intel Xeon class CPU each, 8 GB of RAM combined for data storage processes and up to 252 GB of RAM combined at disposal for data analytics and applicative processes.

6.3. Results and discussion

In this section we will analyze and discuss experimental results from the aspects of identification of adequate model selector, behavioral pattern recognition, behavioral change (transition) recognition and anomaly detection.

6.3.1. Identification of adequate model selector

Since there is no consensus about the best HMM model selection and evaluation metric in unsupervised setting, our first objective was to identify well suited metric for data at hand. Good metric should enable automated identification of parsimonious solutions: ones with high performance but as less complex as possible. For that purpose, we inspected general behavior of AIC, BIC over all experiments (care recipients and activity measures) and correlated these values with log likelihood performances. Log likelihood measures how probable is model given the series data. It is intuitively clear that models with maximum possible log

http://dx.doi.org/10.5772/intechopen.75203

likelihood are desired, however in general, likelihood monotonically increases with increase of model complexity. This means that larger number of clusters will almost always be preferred by log likelihood criteria. The aim of the first analyses was to inspect how AIC and BIC measurements capture degree of changes (slope) in likelihood values of the model. Distributions of average values of log likelihood, AIC and BIC over different model complexities (numbers of states) are shown on **Figure 6**.

On X-axis numbers of clusters are showed and on Y-axis average AIC, BIC and log likelihood values (over all experiments), respectively. It can be seen on figure below that AIC values follow adequately identify steep growth of log likelihood on log likelihood curve. Meaning that average AIC shows better model performance while log likelihood performance increases in large steps.

Optimal number of clusters (in average over all experiments) according to AIC measure is 5 where "elbow" in AIC curve is detected. This point corresponds to transition from higher growth (for number of clusters 2–5) of log likelihood to Lower growth (for number of clusters 6–10). On the other side, BIC model selector, ignores steep increase of log likelihood and identifies three as optimal number of clusters.

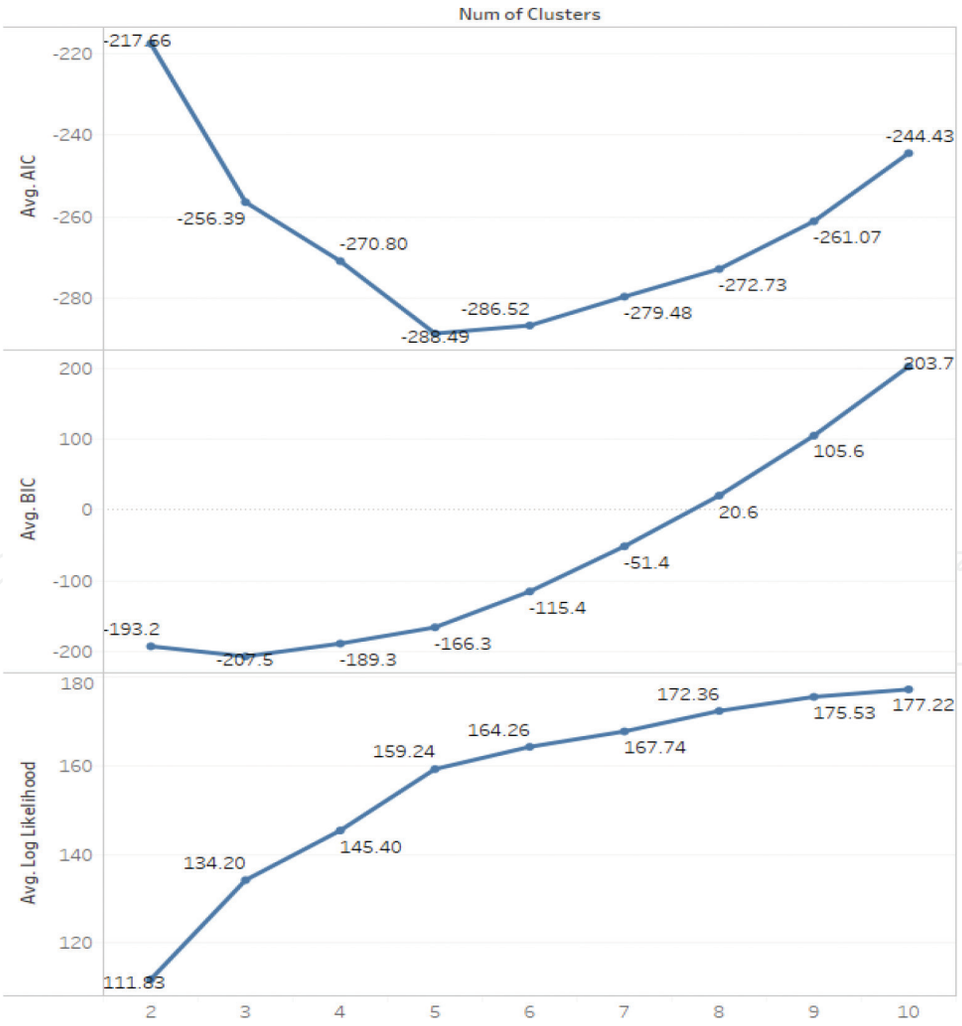

Figure 6. Distribution of average values of log likelihood, AIC, and BIC over different model complexities.

After this point, BIC curve grows super linearly meaning that it does not prefer models with higher number of clusters than 2 or 3. Deeper inspection of AIC, BIC and log likelihood curves for each care recipient and each activity showed consistent behavior with ones described on **Figure 6**. Thus we selected AIC as measure of choice for HMM model selection. Based on previous discussion we took AIC as measure of choice for model selection and identification of optimal number of behavioral state for each care recipient and each activity.

However, it is very important to emphasize that insights presented in previous text cannot be considered as conclusive and cannot generalize over all problems. This is because cluster performance is dependent on data distributions that are different for each dataset, but also because depends on the context of analyses.

6.3.2. Choosing optimal number of clusters

Based on previous insights, we used AIC measure to analyze quality of the models with respect to number of clusters. Results for each activity for one care recipient are shown on **Figure 7**. It can be seen from **Figure 7** that different activities have different "optimal" number of clusters. In this analyses term "optimal" have to be considered very loosely because, in many cases difference in AIC performance is very similar for different number of clusters. This means that for behavior analysis purposes adequate model can be selected in range of models with good and similar AIC performance. Most often, parsimonious solution is applied: model with satisfying performance and the least number of cluster is selected. On the other hand, in case of existence of global saddle point model selection is clearer process. Saddle points have strict mathematical definition based on function derivatives, but in this case, saddle point can be descriptively defined as: point with property that all points from the left side (lower number of clusters) are larger and all points from the right side (higher number of clusters) are larger. In these situations, model selection is based on minimal (optimal value of AIC). Clear example of saddle point on **Figure 7** is labeled with k = 4 for physical_activity_calories activity measure.

6.3.3. Behavior characterization

Figure 8 depicts behavioral patterns for activity sleep_light_time for one care recipient identified by HMM. X-axis represents temporal dimension in day units is presented for the period and Y-axis represents cumulative duration of sleep_light time for each day.

It can be seen that HMM model based on AIC model selection criteria identified three different clusters (behavioral patterns) that can be characterized as following:

1. Behavior (purple line): medium values of sleep_light_time (between 8000 and 13000 s) with low deviations,
2. Behavior (green line): high values of sleep_light_time (between 13000 and 20000 s) with low deviations and
3. Behavior (red line): low values of sleep_light_time (between 0 and 15000 s) with high deviations.

http://dx.doi.org/10.5772/intechopen.75203

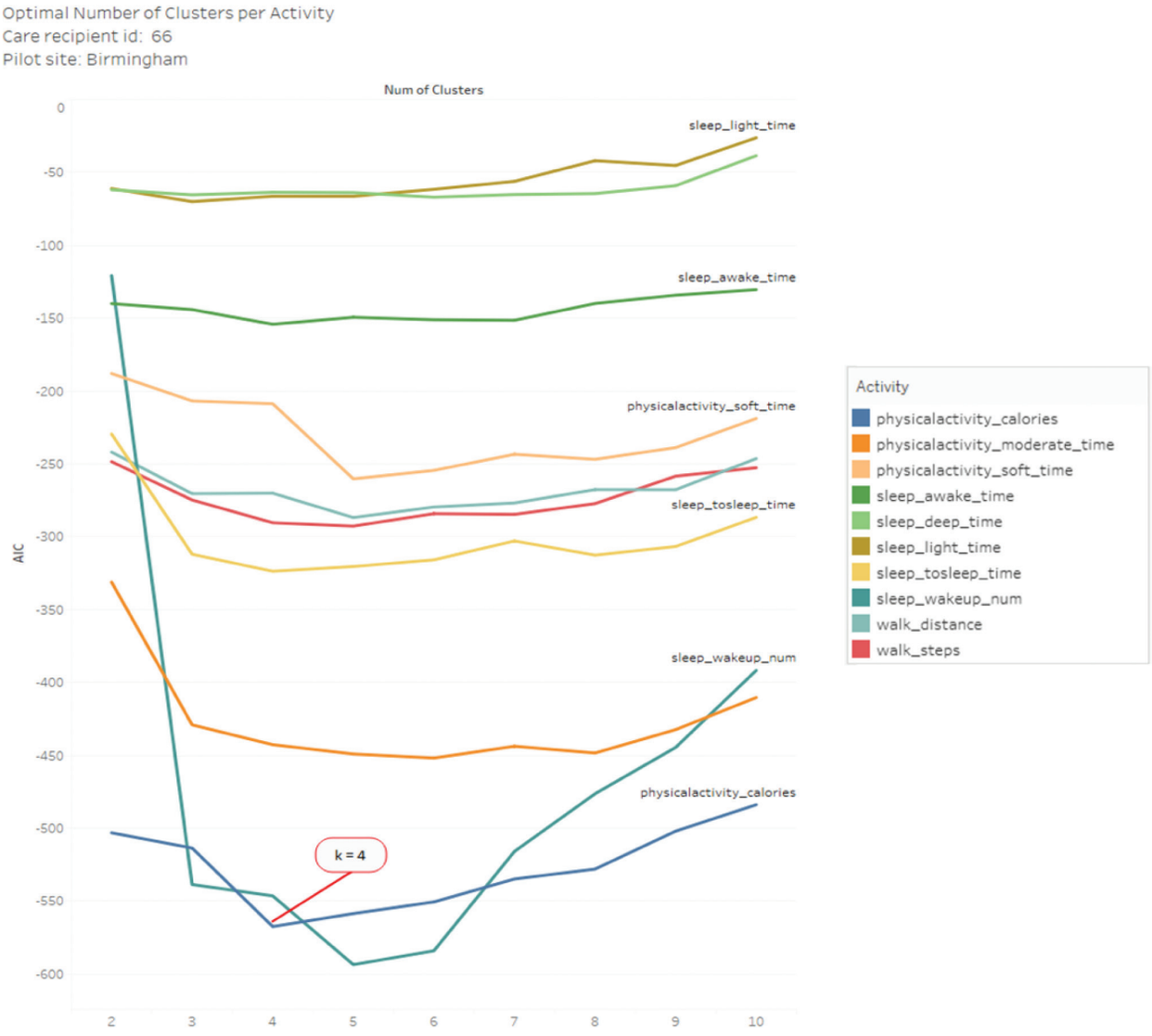

Figure 7. Selection of "optimal" number of behavioral patterns based on AIC values.

Normal sleeping process includes interchange of light sleep and deep sleep. First and second behaviors are considered desirable and such times of light sleep lead to mitigation of frailty risk. On the other hand lack of light sleep time and high variations are considered as negative behavior and could indicate increase of stress and chance of MCI/frailty risk development. Based on these observations behavioral patterns are quantified and ordered (e.g., 1—worst behavior, 2—medium behavior, and 3—good behavior) and pushed in further process of risk quantification through derivation of numerical indicators and grading (described in previous section).

6.3.4. Behavior variation change and anomaly detection

After characterization of behavioral patterns, we analyzed behavior (pattern) changes over time. Identification and characterization of behavior changes (transitions) over time is crucial step for building proactive systems and providing timely and preventive interventions. **Figure 9** describes transitions of behaviors identified in previous sub-section.

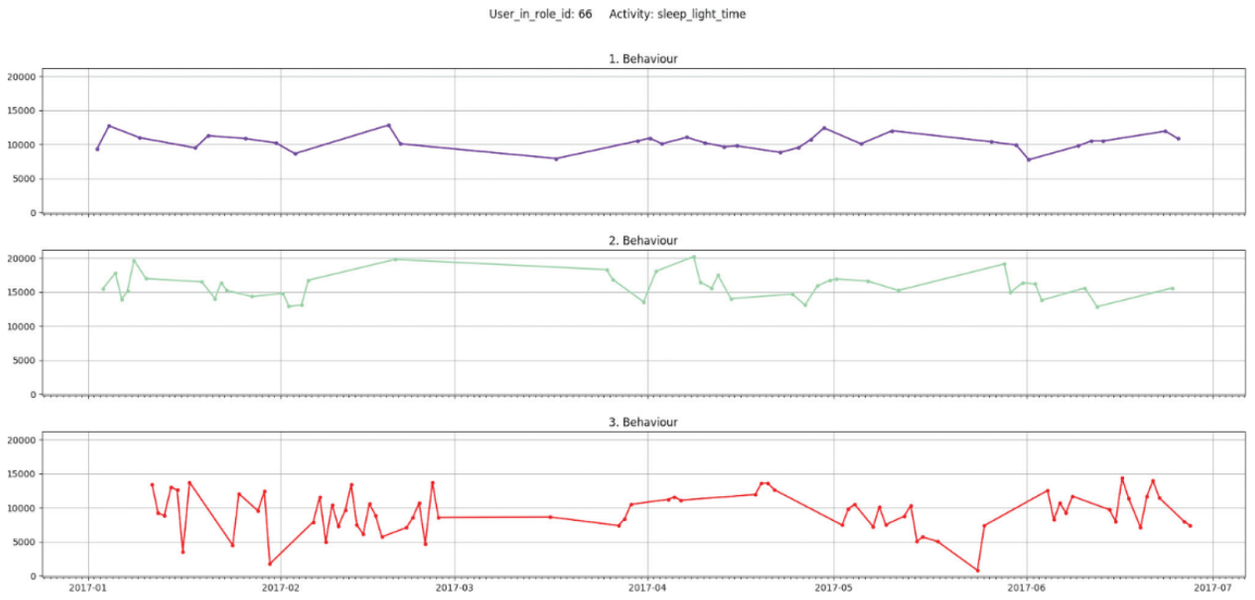

Figure 8. Behavioral patterns identified by HMM model.

Frequent pattern changes from **Figure 9** can be observed from green ("good" behavior) to red ("bad" behavior) lines. It can also be observed that red behavior appears more frequently than other two.

Finally, in most cases "medium" behavior (purple line) transitions to "good" behavior (green line). Based on this analysis it can be observed that after behavior improvement (from "medium" to "good") care recipients often have sudden worsening of behavior. Recognition of such transitional patterns enables predictive and preventive approach in risk prevention. Namely, HMM models, based on transitional probability matrices identify probabilities of

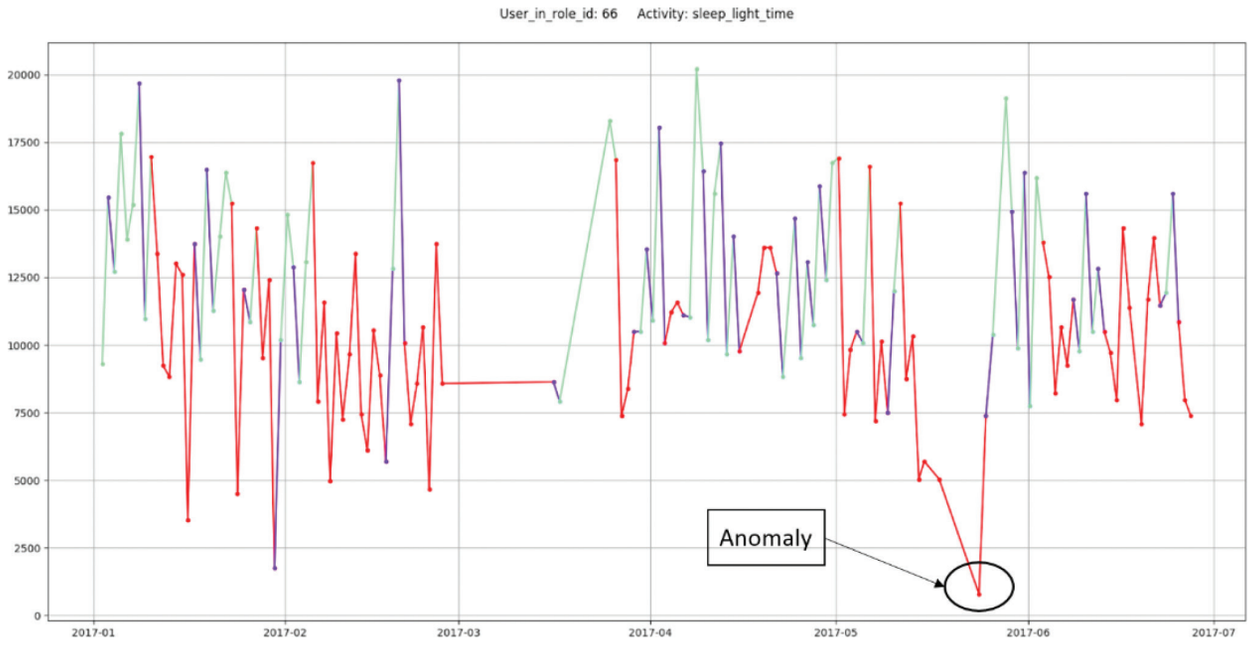

Figure 9. Behavior variations (transitions) and anomalous point.

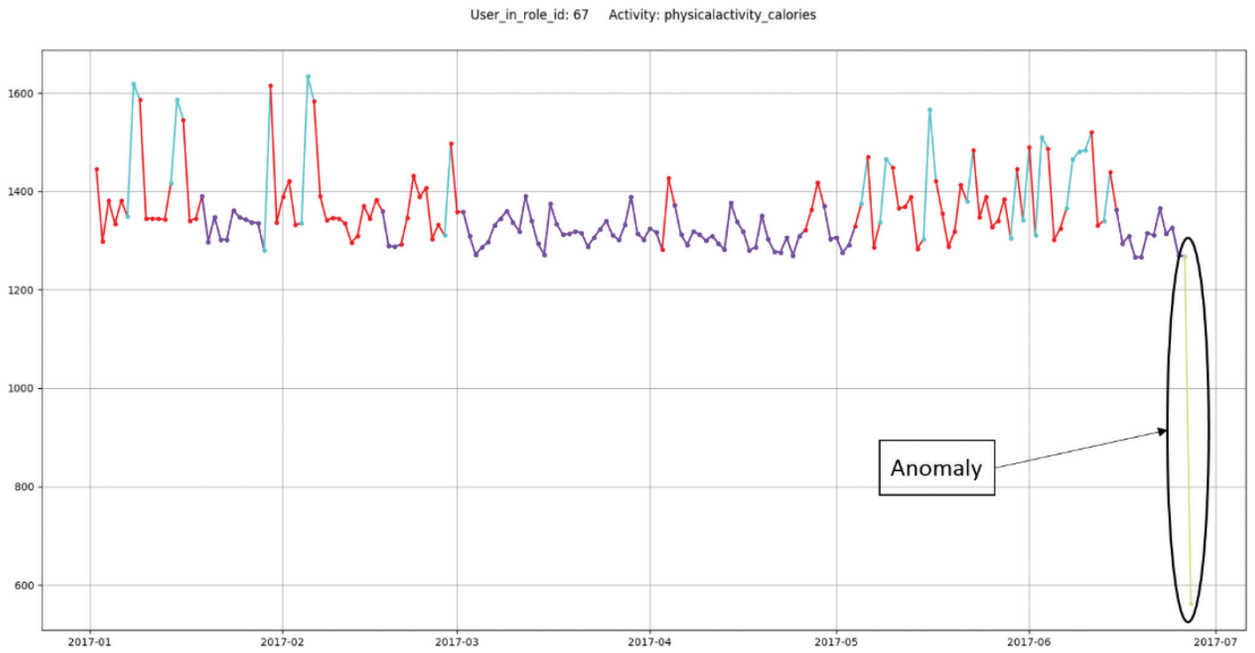

Figure 10. Anomalous state.

behavior transitions and if behaviors are characterized well, these probabilities can be used as early risk identification indicators. Furthermore, based on HMM, model anomalies can be automatically identified per user defined thresholds. For example, by manual labeling on behavioral series presented on **Figure 9**, the lowest point of bad behavior (red line between 2017-05 and 2017-06) is identified. This point is captured as anomalous based on probability threshold of 70%. This means that behavioral point (instance) has max. probability of belonging to any state less than 70%. Experiments on all other activities showed that optimal value of threshold should be between 65 and 75%. Similarly, anomalous states (behaviors) can be identified by setting threshold for minimum number of instances (behavioral measurements) that should constitute behavior (cluster). Since number of behavior measurements is variable for different users, activities and even periods of measurements, we define threshold as percentage of total number of measurements for selected period. In all our experiments series were constituted from 140 to 180 measurements. Experiments showed that good anomaly scoring is achieved by setting threshold to 3–5%. **Figure 10** illustrates situation where anomalous behavior is detected (last two measurements connected with yellow line).

7. Conclusion and future work

In this chapter we addressed the problem of behavioral pattern recognition, behavior change detection and anomaly detection based on IoT data in smart city environment. We proposed a framework for behavioral change detection that will be utilized in context of mild cognitive impairment (MCI) and frailty risk assessment and detection in the City4Age project. Behavioral modeling and risk assessment for MCI and Frailty are very challenging tasks because of the large variations between each specific personal case, and the practical lack

of universally agreed and adopted criteria in geriatric practice (in real-life environment, not controlled "lab" settings) on the referent thresholds or ranges of quantified risk factors or geriatric domain variables that actually denote certain MCI/frailty risk or potential onset.

Thus we developed data driven models based on HMMs that exploit IoT sensory data and allow automated behavior recognition, change and anomaly detection. Models are used for characterization of data that serves as an input for exploratory analytics through interactive dashboarding and/or enrichment of modeled Geriatric factors that quantify the specific behavior characterizations and risk levels for MCI and Frailty.

In future work, we will integrate results from this research in City4Age interactive monitoring dashboards and thus enable geriatricians to gain additional insights into care recipients behavior and potential risk. This will open the space for supervised behavioral scoring and risk prediction. Further, we will develop data driven behavioral models for multivariate IoT data series and explore mutual influence between series. Finally, we will evaluate more unsupervised models for behavioral modeling including deep learning models (e.g., recurrent neural networks) in the analyses.

Acknowledgements

Main body of this research is part of the European project City4Age that received funding from the Horizon 2020 research and innovation funding programme, under grant agreement number 689731.

Author details

Vladimir Urosevic[1], Ana Kovacevic[2], Firas Kaddachi[3] and Milan Vukicevic[4]*

*Address all correspondence to: vukicevicm@fon.bg.ac.rs

1 Belit Ltd., Belgrade, Serbia

2 Big Data Analytics, Belgrade, Serbia

3 Montpellier Laboratory of Informatics, Robotics and Microelectronics (LIRMM), Montpellier, France

4 Faculty of Organizational Sciences, University of Belgrade, Belgrade, Serbia

References

[1] Van Poucke S, Zhang Z, Schmitz M, Vukicevic M, Vander Laenen M, Celi LA, De Deyne C. Scalable predictive analysis in critically ill patients using a visual open data analysis platform. PloS One. 2016b;**11**(1):e0145791

[2] Van Poucke S, Thomeer M, Heath J, Vukicevic M. Are randomized controlled trials the (G) old Standard? From clinical intelligence to prescriptive analytics. Journal of Medical Internet Research. 2016;**18**(7):e185. DOI: 10.2196/jmir.5549. PMID: 27383622, PMCID: 4954919. http://www.jmir.org/2016/7/e185

[3] Kaddachi F, Aloulou H, Abdulrazak B, Fraisse P, Mokhtari M. Unobtrusive Technological Approach for Continuous Behavior Change Detection Toward Better Adaptation of Clinical Assessments and Interventions for Elderly People. In: International Conference on Smart Homes and Health Telematics. Cham: Springer; 2017, August. pp. 21-33

[4] Sprint G, Cook DJ, Schmitter-Edgecombe M. Unsupervised detection and analysis of changes in everyday physical activity data. Journal of Biomedical Informatics. 2016 July;**63**:54-65. ISSN: 1532-0464

[5] Gupta M, Gao J, Aggarwal CC, Han J. Outlier detection for temporal data: A survey. IEEE Transactions on Knowledge and Data Engineering. 2014;**26**(9):2250-2267

[6] Silva JA, Faria ER, Barros RC, Hruschka ER, de Carvalho AC, Gama J. Data stream clustering: A survey. ACM Computing Surveys (CSUR). 2013;**46**(1):13

[7] Aghabozorgi S, Shirkhorshidi AS, Wah TY. Time-series clustering–A decade review. Information Systems. 2015;**53**:16-38

[8] Copelli S, Mercalli M, Ricevuti G, Venturini L. City4Age frailty and MCI risk model, v2. City4Age Project Public Deliverable D. 2017 October;**2**(06)

[9] Azkune G, Almeida A, López-de-Ipiña D, Chen L. Extending knowledge-driven activity models through data-driven learning techniques. Expert Systems with Applications. 2015 April;**42**(6):3115-3128. ISSN: 0957-4174

[10] Bilmes JA. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. International Computer Science Institute. 1998;**4**(510):126

IntechOpen