Temporal Clustering and Anomaly Detection in Elderly-friendly Smart Cities Based on Hidden Markov Models

Vladimir Urosevic*, Ana Kovacevic**, Firas Kaddachi***, Milan Vukicevic****
* Belit Ltd, Belgrade, Serbia
** Big Data Analytics, Belgrade, Serbia
*** Montpellier Laboratory of Informatics, Robotics and Microelectronics, Montpellier, France
**** University of Belgrade, Faculty of Organizational Sciences, Belgrade, Serbia

vladimir.urosevic@belit.co.rs, ananp.90@gmail.com, firas.kaddachi@lirmm.fr, vukicevicm@fon.bg.ac.rs

***Abstract*—In this paper, we propose a methodology for behavior variation and anomaly detection from acquired sensory data, based on temporal clustering models. Data are collected from smart cities 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. Our results show that Hidden Markov Models (HMMs) allow efficient (1) recognition of significant behavioral variation patterns and (2) early identification of pattern changes.**

I. Introduction

Frailty and Mild Cognitive Impairment (MCI) are common and inevitable conditions in the elderly citizen population defined as critical intermediate (but reversible) presursor statuses of 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. The 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 at multi-monthly, or semi-annual intervals) and thus do not support predictive and preventive interventions. Finally, data collected via such methods are often subjective or incomplete.

The framework for collection and analysis of elderly citizens IoT data, developed within the City4Age project (www.city4ageproject.eu), aims to overcome the stated problems. The main tasks of City4Age analytic framework are recognition of behavioral patterns, behavior changes (transitions) in time and anomaly detection that can be used for early identification of frailty/MCI risks. Consequently, the main research challenges of this study are:

(1) Can meaningful behavioral patterns from temporal series of IoT data can be automatically recognized and characterized?

(2) Is it possible to predict behavioral pattern changes and detect the potential risk of MCI/frailty?

Given the unsupervised nature of the research questions, temporal clustering techniques impose as a natural solution. In this paper we propose a method for characterization of behavioral time series, identification of temporal pattern changes and anomaly detection. Method is based on Hidden Markov Models because they provide framework for: (1) time series characterization, (2) prediction of behavioral state transitions and (3) anomaly detection.

II. City4Age Project

The European Commission (EC) extensively supports the developments and establishing of fully Ambient-assisted Age-friendly Cities. Important high-level initiatives that concentrate and channel innovation and stakeholder efforts are the:

- European Innovation Partnership on Smart Cities and Communities (EIP on SCC), involving about 400 committed cities and other partners with specialized initiatives and solutions.
- European Innovation Partnership on Active and Healthy Ageing (EIP on AHA), first established EIP, in 2011.

The City4Age pioneering cross-cutting project (funded under the EC Horizon 2020 programme) acts as a bridge between the mentioned EIP on SCC and EIP on AHA, contributing to specific and shared objectives from both Partnerships, with the primary aim to enable fully Age-friendly Cities through provision of ICT tools and services for unobtrusive early detection of MCI/frailty risks from heterogeneous (IoT and other) data sources at homes or on the move in the city, comprising the research and development work performed and results presented in this paper 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 final objective is to define a model ensuring sustainability and extensibility of the developed services and tools. The developed system and components are being validated through pilot deployments in 6 smart cities – Athens, Birmingham, Lecce, Madrid, Montpellier and Singapore.

III. State-Of-The-Art

Behavior recognition, change and anomaly detection can be modelled naturally with clustering algorithms, as noted above. Clustering techniques 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 [1].

Exhaustive and comprehensive reviews about temporal clustering algorithms and applications have recently been published [1][2][3], and we will focus here only on the review of recently published papers that are closest to our research and experimental evaluation.

HMMs [4] provide a widely used framework for identification of temporal patterns in time series data, as well as anomaly detection. Additionally, HMMs support building predictive models of identified behavioral patterns in future and this adds a predictive and preventive component to the analytic framework.

Smart-Hospital, an IoT-based platform for Hospital Information Systems (HIS) that helps in solving problems of fixed information point. or inflexible networking mode that prevents automation of processes, is proposed by [5]. They exploit dynamic probabilistic nature of Hidden Markov Models for development of dense sensing approach based on traces of activities in a smart hospital scheme settings. Hybrid HMM-IoT system continuously synchronizes itself with the current spectral analysis data of the bio-signals of the admitted patients acquired by non-invasive processes like electroencephalogram, electromyogram, electrocardiogram etc.

HMMs are used by [6] to address the problem of efficient power usage of IoT devices. This problem grows in importance since evolution of Internet of Things (IoT) demands interconnection of many autonomous and heterogeneous devices, and many of such devices have very limited power. HMMs have been employed to efficiently orchestrate the heartbeat duration of status of a device, the approach identifies the device anomalies with high accuracy and also saves the end device power, by intelligently transmitting heartbeats based on HMM analysis.

HMMs are also successfully used in DNA analyses and diagnosing cancer. [7] addressed the problem of the size of DNA sequences collected from next generation sequencer using clustering methods based on HMMs and Gaussian Mixture Models (GMMs), scalable enough to cope with such large sequences. This approach showed better results compared to Pruned Exact Linear Time method, binary segmentation method and segment neighborhood method.

HMMs are widely used spefically in Smart City projects. Fall detection method is proposed by [8]. They used accelerometer and gyroscope sensor to correctly predict the falls and reduce the false positives and false negatives and increase the accuracy. This method allows real time detection of falls considered as risky for the elderly people and requiring momentary prevention and interventions.

In [9], HMMs are used for direction estimation for a pedestrian monitoring system in smart cities. Parallel key frame extraction for surveillance video service in a smart city is developed in [10]. Reliable stochastic data-driven models applied to the energy saving in buildings based on HMMs are proposed in [11]. A method for inference of Activities of Daily Living (ADL) is presented in [12].
However, to the best of our knowledge, application of HMMs for *Behavior Characterization, Behavioral Transition Prediction* and *Behavioral Anomaly Detection* is not present in the literature. Our research is conducted on recently collected (2017/2018) activity data from the pilot sites of the City4Age project.

IV. Hidden Markov Models

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 to the analytic framework. Here we will consider first order HMMs where each temporal state depends only on one previous state. This is a strong assumption, but allows development of scalable models and realtime inference. Figure 1 describes first order Markov chain where each state x depends on previous state (x-1) and observed data (y).

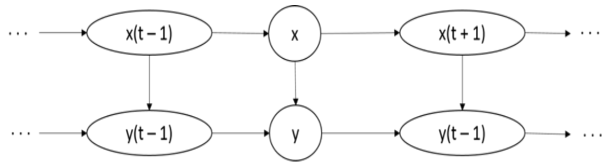

Figure 1. First order Markov chain

HMMs can be explained as total probability of X and Y by following formula:

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

where: $p(y_t|x_t)$ represents observation probability, while $p(y_{t'}|x_{t-1})$ represents transition probability. Details about HMMs can be found in [4]. In our case observations are series of IoT sensory data while hidden states represent categorized, homogenous series parts (that will be characterized as behavioural patterns or behaviours). This is why we use Gaussian HMMs that characterize states with Gaussian distributions. This is depicted on Figure 2.

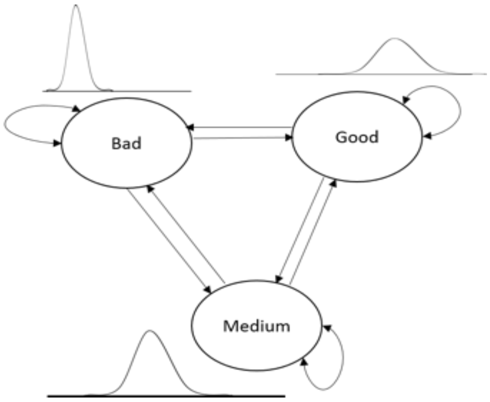

Figure 2. Behaviour modelling with Gaussian HMMs

Each HMM model is thus constituted from 3 elements:

1. Prior probability distribution of hidden states (vector π) that describes how frequently each state occurs in general.
2. Transition matrix (Ai,j) that describes 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 modelled and thus means and standard deviations are used for definiton of hidden state (behaviour) probability distribution. HMMs allow modelling of discrete data as well, but in that case state probability distributions are conditional.

V. EXPERIMENTAL EVALUATION

The main goal of our experiments was to show that HMM 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, identification of an optimal number of behavioral states for each observed elderly citizen (geriatric Care Recipient) and each Activity type, and finally the characterization of identified behaviors (clusters or behavioral patterns).

A. Data

Data used in experiments originate from the Birmingham Pilot of City4Age project. Data have been acquired by monitoring 3 Care recipients during a 6-month period (from January to July 2017, and ongoing). Sensory data have mostly been collected using Nokia Steel (ex Withings Activité) smartwatches.

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 were sets of Activity Measures for each citizen. Summary of observed Activity Measures is presented in the Table below:

TABLE I
OBSERVED ACTIVITY MEASURMENTS

Geriatric Sub-factor	Activity Measure	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

B. Experimental setup

Since HMM-type models cannot implicitly learn an optimal number of hidden states, we built HMM with a varying number of clusters (in the range 2-10) for each Care Recipient and each Activity type. Additionally, since there is no consensus for evaluation of cluster models in an unsupervised setting, each model was evaluated with Log Likelihood, Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) evaluation measures. We conducted 810 evaluations in total (3 Care Recipients * 10 activities * 9 variations of state numbers * 3 evaluation measures), applied to Activity Measure time series for each citizen and each Activity type. In this way, we labelled each time point with cluster (behavioral pattern or state) assignment. Probabilities that time point originates from specific cluster distributions are identified during *scoring* of HMM, and largest probabilities are stored for anomaly detection purposes. Experimental setup is implemented in Python. HMM learn library is used for building HMM while Pandas DataFrame is used for data manipulation.

C. Evaluation measure selection

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. The good metric should enable automated identification of parsimonious solutions: ones with high performance but minimal possible complexity. 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. Distributions of average values of Log Likelihood, AIC and BIC over different model complexities (numbers of states) are shown on Figure 3.

X-axis shows the numbers of clusters, and Y-axis the average AIC, BIC and Log Likelihood values (overall experiments), respectively.

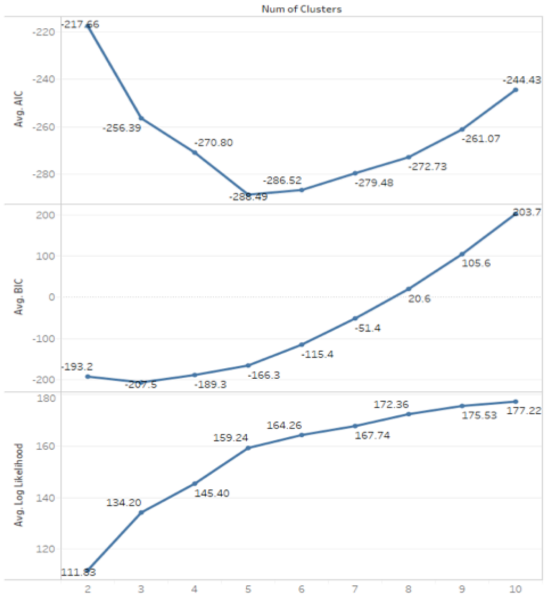

Figure 3. Distribution of average values of Log Likelihood, AIC and BIC over different model complexities

It can be seen on the Figure below that AIC values follow adequately the identified growth of log likelihood on Log Likelihood curve. Meaning that average AIC shows better model performance while Log Likelihood performance increases in large steps. However, it is very important to emphasize that insight presented in the 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.

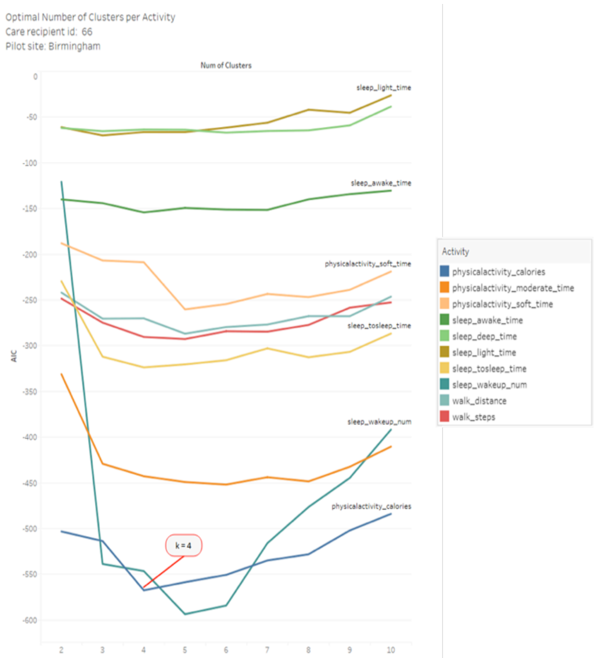

Figure 4. Selection of "optimal" number of behavioural patterns based on AIC values

Based on previous insights, we used AIC measure to analyze the quality of the models with respect to number of clusters. Results for each Activity type for one Care Recipient are shown on Figure 4. "Optimal" clusters identified by HMM have shown very promising performance in behavioral pattern recognition.

D. *Behaviour characterization*

Figure 5 depicts behavioral patterns identified by HMM for Activity Measure *SLEEP_LIGHT_TIME* for one Care Recipient. The X-axis represents temporal dimension in day units for the period and Y-axis represents cumulative duration of *SLEEP_LIGHT_TIME* for each day (24h).

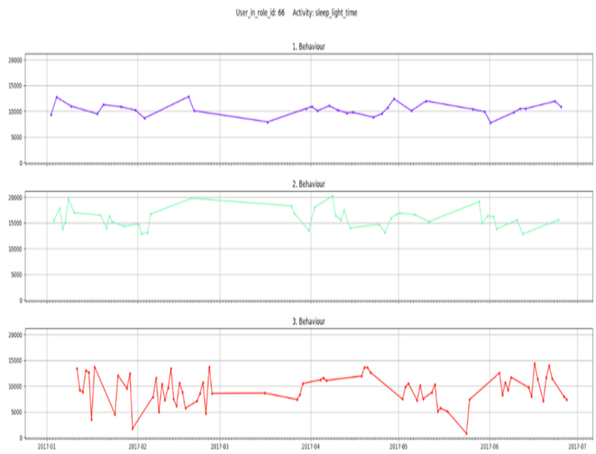

Figure 5. Behavioural patterns identified by HMM

It can be seen that HMM based on AIC model selection criteria identified 3 different clusters (behavioral patterns) that can be characterized as following: medium values of *SLEEP_LIGHT_TIME* (between 8000 and 13000 seconds) with low deviations (purple line), high values of *SLEEP_LIGHT_TIME* (between 13000 and 20000 seconds) with low deviations (light green line), low values of *SLEEP_LIGHT_TIME* (between 0 and 15000 seconds) with high deviations (red line).

The normal sleeping process includes interchange of light sleep and deep sleep. Medium and High values of *SLEEP_LIGHT_TIME* are considered desirable and generally lead to mitigation of frailty risk. On the other hand, lack of light sleep time and high variations are considered as non-favorable behavior and could indicate an increase of stress and chance of MCI/frailty risk development. Based on these observations, behavioral patterns are quantified and ordered (e.g. 1 – relatively "worst" behavior, 2 – "medium" behavior, and 3 – "good" behavior), and propagated to the further process of risk quantification.

E. Behavior pattern changes and Anomaly Detection

After characterization of behavioral patterns, we analyzed behavior (pattern) changes over time as a crucial step for building proactive systems and providing timely and preventive interventions. Figure 6 describes transitions of behaviors identified in the previous sub-section. Frequent pattern changes from Figure 6 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). 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, based on transitional probability matrices identify probabilities of behavior transitions and if behaviors are characterized well, these probabilities can be used as early risk identification indicators. Furthermore, anomalies can be automatically identified per user-defined thresholds. For example, by manual labeling on behavioral series presented on Figure 6, 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%.

We can conclude that data-driven models based on HMMs allow automated behavior recognition, change and anomaly detection from sensory IoT data. Additionally, they allow proactive behavior change detection and timely warnings of risk levels for MCI and frailty in Smart City environment.

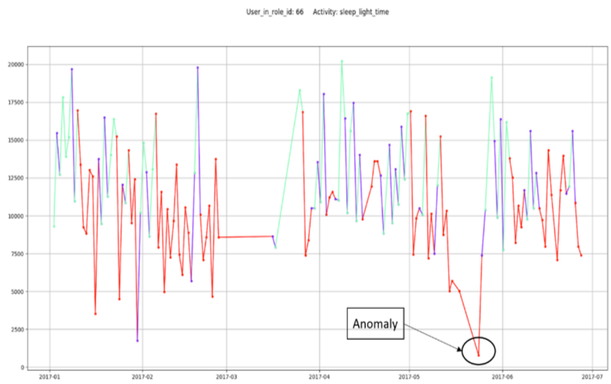

Figure 6. Behaviour variations (transitions) and anomalous point

Another clear example of anomaly that is detected can be seen on Figure 7. Closer inspection of the usage of data has shown that this specific anomalous value of *PHYSICALACTIVITY_CALORIES* had been due to an error on the data acquisition device.

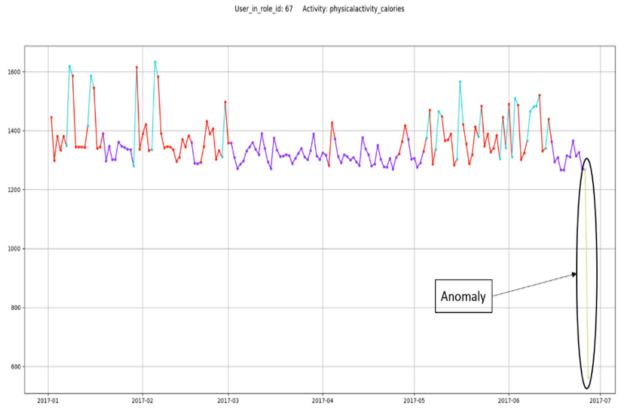

Figure 7. Anomalous state

VI. CONCLUSION AND FUTURE WORK

In this paper we addressed the problem of behavioral pattern recognition, behaviour change detection and anomaly detection from personal IoT data in Smart City environment. We proposed a framework for behavioural 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 assesment 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.

We have therefore developed data-driven models based on HMMs that exploit IoT sensory data and allow automated behaviour 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 modelled Geriatric Factors that quantify the specific behaviour characterizations and risk levels for MCI and frailty.

In future work we will integrate results from this research in City4Age interacive monitoring dashboards and thus enable geriatricians to gain additional insights into Care Recipients behaviour and potential risks. This should enable semi-supervised and/or supervised behavioural scoring and risk prediction. Further, we will develop data-driven behavioural models for multivariate IoT data series and explore mutual influence between series. Finally, additional and other unsupervised models for behavioural recognition will be evaluated in the analyses, including deep learning methods (e.g. recurrent neural networks).

ACKNOWLEDGMENT

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.

REFERENCES

[1] [Aghabozorgi, S., Shirkhorshidi, A. S., & Wah, T. Y. (2015). Time-series clustering–A decade review. Information Systems, 53, 16-38

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

[3] Silva, J. A., Faria, E. R., Barros, R. C., Hruschka, E. R., de Carvalho, A. C., & Gama, J. (2013). Data stream clustering: A survey. ACM Computing Surveys (CSUR), 46(1), 13.

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

[5] Seth, D., & Chakraborty, D. (2017, July). Hidden Markov model and Internet of Things hybrid driven smart hospital. In *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.

[6] Kumar, P., & D'Souza, M. (2017, January). Design a power aware methodology in IoT based on hidden markov model. In *Communication Systems and Networks (COMSNETS), 2017 9th International Conference on* (pp. 580-581). IEEE.

[7] Manogaran, G., Vijayakumar, V., Varatharajan, R., Kumar, P. M., Sundarasekar, R., & Hsu, C. H. (2017). Machine learning based big data processing framework for cancer diagnosis using hidden markov model and gm clustering. *Wireless Personal Communications*, 1-18.

[8] Chandra, I., Sivakumar, N., Gokulnath, C. B., & Parthasarathy, P. (2018). IoT based fall detection and ambient assisted system for the elderly. *Cluster Computing*, 1-9.

[9] Raman, R., Sa, P. K., Majhi, B., & Bakshi, S. (2016). Direction estimation for pedestrian monitoring system in smart cities: An HMM based approach. *IEEE Access*, *4*, 5788-5808.

[10] Zheng, R., Yao, C., Jin, H., Zhu, L., Zhang, Q., & Deng, W. (2015). Parallel key frame extraction for surveillance video service in a smart city. PloS one, 10(8), e0135694.

[11] Ridi, A., Zarkadis, N., Bovet, G., Morel, N., & Hennebert, J. (2013, September). Towards reliable stochastic data-driven models applied to the energy saving in buildings. In International Conference on Cleantech for Smart Cities & Buildings from Nano to Urban Scale (CISBAT 2013).

[12] Cheng, B. C., Tsai, Y. A., Liao, G. T., & Byeon, E. S. (2010). HMM machine learning and inference for Activities of Daily Living recognition. *The Journal of Supercomputing*, *54*(1), 29-42.