International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 13 Issue 3 –MARCH 2015.

AUTOMATIC INFORMATION FORWARDER FOR ELDERLY HEALTHCARE SUPPORT

Ms. Vidya G.Adike [#1] and Prof.Vivekanandreddy [*2]

[#1] *P.G Student, M.Tech CSE, VTU, PG Centre Belagavi, Karnataka, India.*

[*2] *Assistant Professor, Computer Science and Engineering, VTU ,PG Centre Belagavi, Karnataka, India.*

***Abstract*—Ubiquitous healthcare systems enabled by information and communication technology (ICT), can allow the elderly and chronically ill to stay at home while being constantly monitored. A big data pilot system is presented using wearable sensors capable of carrying out continuous monitoring of the patient at home and alerting the caretaker when necessary and to forward only pertinent information to the big data system for analysis. A challenge for such a solution is the development of context awareness through the multidimensional, dynamic and nonlinear sensor readings that have weak correlation with observable human behaviors and health conditions. To address this challenge, a wearable sensor system with intelligent data forwarder is discussed. The forwarder adopts a Hidden Markov Model (HMM) for human behavior recognition. Locality sensitive hashing is proposed as an effective mechanism to learn sensor patterns. The intelligent information forwarder can provide the remote sensors with context awareness. They transmit only important information to the big data server for analysis, when certain behaviors happen and avoid overwhelming communication and data storage. The system functions unobtrusively, by giving the user peace of mind that their safety is being monitored and analyzed.**

***Index Terms*—Big Data, Hidden Markov Model, locality sensitive hashing, wearable sensors.**

I. INTRODUCTION

According to the World Health Organization (WHO), the United States spends about 17.6% of its gross domestic product on healthcare, the highest level in the world and far higher than the percentage for other developed countries (9.3% on average). Nevertheless, the use of healthcare services in U.S. is far below that of comparable countries, reflecting greater inefficiency and higher prices for healthcare services in the United States. The skyrocketing medical expenditures and continuous aging of the world's population demand transformative technological innovations to provide more effective and affordable healthcare services, available to anyone at any time and in any place [1]. It is found that 17% of the elderly has less than weekly contact with their family members [2]. This shows that there is a risk for not being monitored from minor accidents or illness that causes immobility, and from other unforeseeable scenarios that will go undetected if no contact is made with the individual for a long time. Many assistive devices have been made available for a considerable period of time to monitor the patients in their residential environments after which many wearable sensor devices were used in order to interact with the user to ascertain their well-being and their physical health [3],[4]. The monitoring systems can be of two variations, one is autonomous problem determining another one is human problem determining. Autonomous problem determining require gathering of data to infer a belief about the user's state [5]. The human problem determining has the need of human involvement to access the status of the user. These applications gather readings related to the user by utilizing the sensors which are attached to human body or the environmentally located sensors [7][8]. Once the data are gathered they are uploaded to the server that are accessible by some healthcare professional or some other monitoring services, which could identify any issues being faced by the user. Such kind of systems have lower level of processing involved and requires large data throughput to the server and time consuming interpretation by healthcare professionals. These observations which are made is in raw form and the inference of the human behavior is made by human. These systems require low level of processing and heavier data throughput to the server. Acquiring and analyzing data from the distributed devices become a challenge to data communication and processing because such healthcare devices need to be deployed to a great amount of the elderly population for continuous monitoring. The data generated by the healthcare devices are often semi-structured or unstructured and have the 3v characteristics of big data like volume, velocity and variety [9]. In healthcare sectors only few data are considered to be of value for analysis. A big data pilot system is used in this paper which combines the two categories, i.e., autonomous problem determining and human problem determining, which covers the services like continuous behavior monitoring and long term analysis. The system consists of a wearable sensor node for collecting the information, a mobile phone for user interaction and remote access and a centralized big data system as a tool for health condition monitoring. There is a tradeoff between distributed processing in wearable sensor and the centralized analytics in the server cluster for managing such systems. Hence an intelligent data forwarder needs to be embedded in the mobile device to monitor the user's behavior continuously, alert caretakers in case of emergency and transmit only important information to the healthcare big data systems for analysis. Here ubiquitous information is forwarded anytime and anywhere for continuous monitoring. The intelligent information forwarder based on Hidden Markov Model (HMM) makes the distributed sensors context aware and reduce the communication load and data storage for a large scale system. The HMM is used in a broad spectrum of applications, with the ability to recover a hidden-state sequence from the visible observations. For example the model is ideal for gene prediction in bioscience field, where each state emits random DNA strings of random length, which are observable to determine the gene producing them and in the prediction of protein structure and genetic mapping [10] [11]. Utilization of the HMM benefit the cryptanalysis and cryptography [12]. In the measurement of Partial Discharge (PD),the sequential and the time varying properties lend themselves to be modeled with HMM such that the PD patterns can be classified to inform the insulation system defects [13]. Thus for a single observation the traditional HMM uses probability distributions or discrete probability values assigned to single observations. More detailed models take observations from a variety of sources in behavior task estimation in order to ascertain an intelligent estimation of the hidden state. If a number observation sources are reviewed, the hidden state can be determined with greater accuracy e.g., while using different sensors the fusion of inputs must be considered [14]-[16]. The fusion of multiple sensors can produce worst results when compared to single sensor, when taken into consideration. This is because; combining the inaccurate sensor reading with the one evaluated more accurate [15]. The high-dimensional and nonlinear issues of the sensor reading can also produce worst results [17], [18]. Here we use a sensor fusion scheme to estimate the observation probability of states for a HMM based user behavior detection utilizing the developed wearable sensors. To carry out Instance Based Learning (IBL) a locality-sensitive hashing (LSH) table is used. Experiments are conducted to compare the performance of the proposed method with the non-linear dimension reduction method [18], and the results show that the proposed scheme is more efficient for both learning and querying. Due to limited memory, computational power and communication bandwidth available on board, it is obvious that such intelligent processing embedded in mobile device should take resource-saving approach. The remaining session of this paper is organized as follows. The system architecture and software are described in section II, it includes details of operational processes and the signal processing for robust measurement. Section III presents the HMM- based state and anomaly identification that is the key component of the information forwarder. Session IV explains how LSH can be used as an efficient mechanism to estimate the users state from the captured multiple sensor signals using probabilistic modeling. And finally section V contains the conclusions.

II. BIG DATA SYSTEM IN HEALTHCARE

Due to rapidly growing aging population public healthcare is facing serious difficulties. Rather than relying on intrusive care and support every individual has the desire to live independently. There is also a high risk of suffering from illness, accidents and injuries in their day to day activities. There is also a need for a system that can be wearable to monitor the physiological parameters and check health conditions of the user. Managing a diverse user group is a challenging task for health service provider, since the users are dispersed in the whole country. Big data infrastructure is opening a new era to next generation healthcare. It provides individual users to access instant health service from big data system. The services provided can be daily health checks, medication reminders, first aid instructions, comparative effectiveness research, preventive care, and healthy lifestyle encouragement. To provide instant response to emergency situations, some applications can be downloaded from cloud to mobile device. It may be computationally intensive in order to analyze a huge amount of sensor data for a long-term healthcare service. Hence the design of big data system for healthcare should have a tradeoff between distributed intelligence and data analytics. Here the prototype of the big data system is used. Through the integration of distributed monitoring with centralized analytics, the long term care of the population can be improved also efficiency of healthcare can be increased. The system consists of three separate components: a wrist device, a mobile phone, and a big data cluster, as in Fig.3. Wireless measurement is linked with a centralized big data system.

A. *Hardware Device*

A wrist device is designed to include three sensors and it uses a Bluetooth low energy (BLE) technology for connecting with an Android mobile to form a personal area network. It is developed by using PIC16F877A as shown in Fig.1 and Fig.3 the hardware device include a temperature sensor to measure ambient temperature, a pressure sensor to measure the blood pressure and a heartbeat sensor to measure the heartbeat.

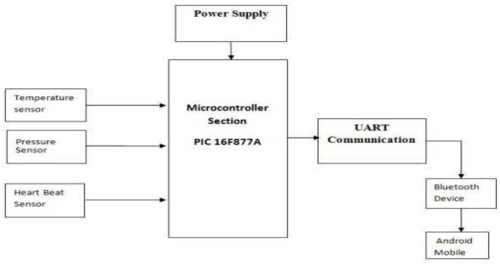

Fig. 1. Hardware architecture

B. *Mobile Application*

The measured parameters are sent to the mobile phone through Bluetooth low energy. In android mobile a mobile application is developed to process the gathered data and make subsequent decision. The mobile application enables intelligent behavior recognition and unobtrusive care. It recognizes the user state

Fig. 2. Schematic representation of wrist device

and forwards the important information to the server for analysis, it helps in alerting the caretaker in case of emergency and sharing of information in social media. Patients are able to share the diverse recorded elements of their personal health information through structured messages with their networked community, consisting of friends and relatives, caregivers, health professionals, and other patients. This information sharing enables the patients to obtain feedback or help (subject to their condition), receive emotional support, etc. patients can choose the exact receivers of the information from their networked community (e.g., specific persons or groups of persons), while they can also decide on the way of disseminating information, choosing between the spontaneous (user-proactive) and the event-driven mode (system reactive), i.e., information is sent automatically whenever a condition associated with the status descriptors occurs (e.g., the patient recorded nausea, the time is between 08:00 and 10:00). The user states include observable states and hidden states, the states are estimated using the accelerometer sensor, inbuilt in the android mobiles. The states are estimated using the axis. The states are no link, link, sleep, stand, sit, walk, run, Low Battery, call. It can be detected from the sensor and component readings directly. The hidden states, e.g., *Sleep*, *Sit*, *Stand*, *Walk*, *Run*, and *Abnormal* are estimations of the inferable behaviors of a user, which are not explicitly determinable from the sensor readings alone. A behavior classifier is developed for their detection.

C. Big Data Server

The states and the sensor readings are sent to the big data system for analytics, to improve and personalize the quality of care, guarantee efficient use of scarce health professional expertise. To ensure that patients know when and how medication should be adjusted, there is also a potential to reach rural patients without proper access to healthcare. It enhance the efficiency for large-scale unstructured data retrieval and analysis. Map Reduction is a software framework introduced by Google to support distributed computing on large data sets using a cluster of computers. It has been widely used as a standard model in big data systems. Several indexers in parallel and a reduction server for search are used in big data clusters. The data are send from mobile phones through a transmission control protocol (TCP) or user datagram protocol (UDP) ports. The UDP can be more appropriate than TCP for high velocity of data, if additional delivery checking is implemented. Data mining and pattern recognition algorithms can be developed to achieve context awareness from distributed information for historical behavior analysis, health condition prediction, and anomaly alerts.

To log information from a user using a mobile phone, a record of data stream is shown in JavaScript Object Notation as follows.

{ “userName”: “San” “deviceAdress”: [12, 42, 46, 68, 34, 12],
{
“time”:“09 : 20 : 112013/9/12 UK”, “eventType”: [Sit],
“accValue”: [45, 23,99], “accL1”: 167, “accAngle”: 1.5, “RSSI”: −72.4, “verityBattery”: 90, “phoneBattery”: 65, “bodyTemp”: 35.6,
“location”: [77.134235, −0.4354365], “callType”: [0, null]
“textType”: [1,“Hi, I am Verity. My friend, . . .”],
“PPG”: [12, 127, 0, 0, . . . , 127],
“HB”: 83, “SPO2”: 97,
“interface”: 0, “bleState”: 1
}

Each record has a unique 48bit address as the identity of a wearable sensor and a user name to identify its wearer, this supports in monitoring of many users. The record includes a timestamp to define the time series of information. The information includes details like readings from sensor, geolocation, call, alerts and so on. The whole time series of information is send to the server every 3s, the system is expected to manage 10,000 users with a replication factor of 3 To have data redundancy in the big data system. The incoming data for big data system can be compressed for storage and indexing, e.g., the compressed data file is approximately 10% of the incoming data, and the index file range in size is approximately 10% to 110% of the compressed raw data file in splunk. With the growth in population every year, we need to store PB of data, which may need more number of nodes in the clusters. Running such a cluster of servers can be very expensive, which requires significant power, cooling rack space, etc. To deal with the properties of 3Vs the big data system should avoid oversampling, particularly for the high velocity of data from distributed sensors. Here only valuable information is forwarded to the big data system and the other irrelevant data is ignored. HMM based hidden state estimation is used to schedule a data forwarder to achieve context aware communication.

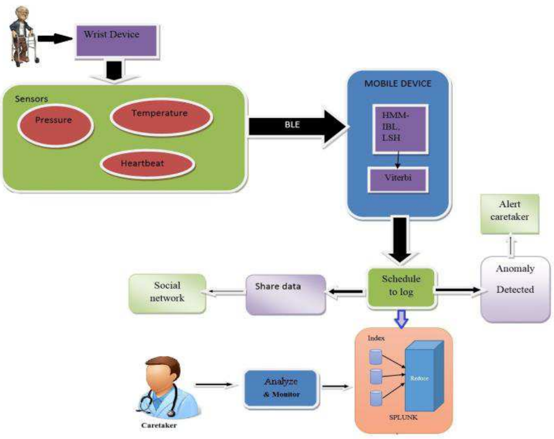

Fig. 3. System architecture

III. STATE-BASED DATA FORWARDER

The big data system needs to manage high volume, high velocity and variety of information assets, these data are often from wireless sensors, handhelds and websites. It is important to use a data forwarder to forward meaningful information to the system. The big data systems are goal or objective driven. For example, a big data system to manage the vital parameters of the elderly or chronically ill to understand general health conditions and exercise engagement through temporal and geographical statistics. The distributed data sources must be provided with intelligence to determine when and what to feed to the system. In each data source a data forwarder is embedded with context-aware capability. The Fig.4 describes such a state driven intelligent forwarder. A configurable schedule is developed in the data forwarder. The schedule consist of set of rules for logging data in to big data system. Users can specify the rules using meaningful states, e.g., “sensing sensor data when anomaly detected or any state transition”

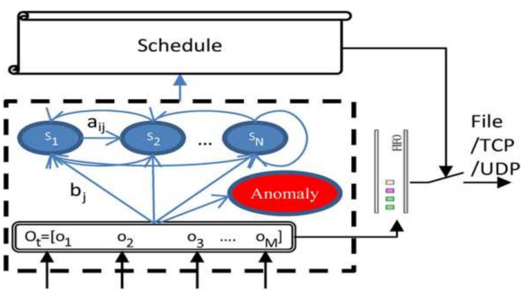

Fig. 4. State-driven information forwarder.

The context awareness of the forwarder is achieved by a HMM that is used to detect users hidden behavior’s, such as running and anomaly, from its sensor readings.

A. Optimal State Estimation using Viterbi Algorithm

The HMM in Fig. 4 has N hidden states $S = [S_1, S_2, \ldots, S_N]$, and M observations from sensors $O_t = [O_1, O_2, \ldots, O_M]$, $t = 1, \ldots, T$, where a_{ij} denotes the transition probability, i.e., $a_{ij} = P(q_{t+1} = S_i | q_t = S_i)$, and $b_j(O_t)$ represents the observation probability that particular sensor readings O_t are measured in the state j, $b_j(O_t) = P(O_t|q_t = S_j)$.
Given an observation sequence $O = [O_1, O_2, \ldots, O_T]$ and a model $\lambda = (a_{ij}, b_j, \pi_j)$, where $i, j = 1, \ldots, N$, and π_j is the initial probability of state j, the probability of the optimal state sequence $Q_* = q_{1*}, q_{2*}, \ldots, q_{T*}$ can be obtained by Viterbi algorithm [22]. Define $\delta_t(i) = \max P[q_1, q_2, \ldots, q_{t-1}, q_t = S_i, O_1, O_2, \ldots, O_t|\lambda]$ $q_1,q_2,\ldots,q_{t-1}$ where $\delta_t(i)$ is the highest probability along a single state sequence as calculated at time t, accounting for the first t observations and terminating with state S_i. The state sequence itself is given in array ψ, which is populated with the state maximizing that probability calculated by δ_t at each step.

1) Initialize:
$\delta_1(i) = \pi_i b_i(O_1), 1 \le i \le N$ $\psi_1(i) = 0, 1 \le i \le N.$
2) Recursion Step:
$\delta_t(j) = \max [\delta_{t-1}(i)a_{ij}]\, b_j(O_t)$
$\psi_t(j) = \text{argmax} [\delta_{t-1}(i)a_{ij}]$
$1 \le i \le N 2 \le t \le T ; 1 \le i \le N.$
3) Terminate:
$P_* = \max [\delta_T(i)]$
$1 \le i \le N$
$q_{T*} = \text{argmax} [\delta_T(i)]$.
$1 \le i \le N$
4) The backtracking procedure:
$q_{t*} = \psi_{t+1}\_q_{t*+1}\_$, $t = T - 1, T - 2, \ldots, 1.$
The resulting state sequence ψ is the most possible sequence that has emitted the observation at time T , given transitions from previous states.

B. Detection of Anomaly

The most likely state sequence based on observations is provided by HMM. For any state the probability returned not only provides

information about the certainty of the activation of the state but can be interpreted as a value that classifies its degree of anomalousness, deviation from the norm denote low probabilities[23],[24]. In case of anomaly, an alert is send to the care taker and current sensor readings to the big data system such information can be scheduled by the user.
Three types of anomalies are defined.

Type_1 Anomaly: Depending upon the certainty of winning state type_1 anomaly is defined. If the probability of the winning state occurring P* is close to other states probabilities. When the winning probability is close to the mean, the instance can be deemed uncertain as $\rho = |P_* - \mu| \leq \beta_1$ In this case where ρ falls within a specified threshold β_1, it indicates significant uncertainty of the identified state. The state is said to be illegible state which means a wrongly defined model that faces an un modeled state or need to be re-estimated using Baum-Welch algorithm.

Type_2 Anomaly: when the observation witnessed does not belong at all in the sequence, and then an equally likely scenario develops. To detect such errors the relevant observation probability has to be monitored. If the observation O_t is low, the inference is that the model has not seen such an observation before and therefore requires either reassessing or triggering an alert, i.e., An instance where this form of anomaly could occur is likely if not all of the possible observations and associated states were captured during the training phase or if the user exhibits a behavior typical of an unprogrammed state. e.g., a stroke or a heart attack indicated by an increase in temperatures and heart rates, the observation would trigger this type of anomaly due to the state not having been seen during training.

Type_3 Anomaly: A type_3 anomaly is a slight variant on the type_2 anomaly and can occur simply when the state at a time step differs for each state determining method within the HMM, e.g., the Viterbi state q_{t*} and the winning state according to pure observation probability $b_j(O_t)$ do not match significantly. For example, if the observation probability is highest for perhaps the state of *Running*, yet the determined state according to the Viterbi method q_{t*} returns *Sleeping* with much higher probability over its *Running* probability, this may in fact indicate a period of distress for the user such as in the instance of a heart attack or some other such observable problem. The schedule in this can be configured to select under which states or anomalies the sensor data should be sent to the big data system for analytics. In order to avoid missing important information when an event happens, a first-in–first-out buffer is used to hold a series of the latest information and will be sent to the big data system once fired by the schedule. The context awareness of the intelligent forwarder relies on correct behavior detection. In the case of an outdated Markov model, detected states could be wrong, and important infor-mation could be missed. It will cause an increasing number of abnormal behaviors to be detected, which may be due to health problems or due to outdated models.

IV. IBL OF OBSERVATION PROBABILITY

Two probabilities have been discussed so far , i.e., transition probabilities a_{ij} and observation probability $b_j(O_t)$ representing the probability that state j has observation O_t. it is possible to find the most probable state, using these probabilities, at a specific time step based on observations made at that point along with preceding states. A solid estimation of the most likely state sequence for an entire set of observations can be provided over a prolonged period. As the observation O_t includes readings from multiple sensors, determining the observation probability becomes more difficult. The dimensionality of the sensor readings can reach eight, which includes temperatures, heartbeat, pressure, and accelerations in three axes. Between data clusters there may be a considerable chance of nonlinearity, the classification is made based on the data distance. The lesser will be the ability to make sense of data if the number of data attributes is greater, due to the fact that with nonlinearity in a higher dimension, standard Euclidean distance functions lose their usefulness; thus, clustering with such methods becomes less accurate. To deal with nonlinear high dimension data, there are a multitude of techniques, with many sharing basic underlying principles to reach lower dimensional representation of a complex nonlinear data sets like Sammon's mapping[26], Isomap[27], and curvilinear component analysis[28],[29] all seek to repli-cate similar distances between points located in a high dimen-sion after placement in the lower dimension, by a means of gradient descent or iterative error reduction methods. A curvilinear distance analysis algorithm was presented in [18] for determining high dimensional and nonlinear space. the observation probability $b_j(O_t)$. The observation O_t may be in a high dimensional and nonlinear space. If it lies on a nonlinear manifold, Euclidean distance makes less sense for classification but has to be replaced by curvilinear distance to measure the distance along the manifold. The algorithm unfolds high-dimensional manifold data to a low-dimensional one by retaining topology, and it forces the clusters to be linearly separable. The algorithm's effectiveness was validated by experiments using the *Verity* platform; how-ever, it is quite time-consuming for the data unfolding because it involves intensive computation to project prototypes in high-dimensional space to a low-dimensional space and to maintain equivalent curvilinear distances. Sometimes, such equivalence may even not exist. IBL[30] is proposed in this paper as an alternative to facilitate learning of $b_j(O_t)$ from demonstration.

A. LSH for IBL

The IBL takes directly sampled data from any system at a known state and constructs a hypothesis regarding similarity without the need to generalize a model based on the often high-dimensional and nonlinear data. Through learning, data instances are stored in some form of memory. This is then ac-cessible for subsequent classification operations, where a query is submitted and compared with all trained values according to some distance metric in order to ascertain its membership to the encoded classes. IBL has multiple advantages over parametric and model-based algorithms, particularly in the storage of new unseen instances. Other algorithms would typically require a complete re-examination of the data set in order to be wholly inclusive of the new data points where IBL methods simply "insert" the new data instance without disrupting any earlier determined model. It is commonly accepted that the genus of and starting point where v is the M -dimensional vector to be hashed, and z_h is a random vector from a p-stable distribution, such as from a $N(0, 1)$ Gaussian distribution. Another random value b uniformly in the range $[0, \omega)$ is then added to the scalar projection, which is then quantized by ω. ω is the width of the bin in which a data point may fall into. _· is the floor operator. This paper presents an LSH-based IBL for obtaining the observation probability $b_j(O_t)$ from high-dimensional and non-linear sensor readings. It includes two stages, i.e., learning and querying.

B. Learning

The learning process is to sample typical sensor readings for different states and encode them into a hash table H with L independent LSH functions h_1, $h_2, \ldots, h_L$.

C. Querying

The retrieval method for any O_t is an LSH recall procedure with "bucket" checking. Different from the conventional LSH for k-NNs, we want to calculate the density of observations of a given state j near O_t in a given radius $R \in Z$ for the probability $b_j(O_t)$ estimation.

D. A Healthcare Big Data System

A prototype of the big data system has been developed by using Splunk Enterprise 6.0 for analytics of the behaviors of wearers. Splunk is a time-series engine that can collect, index, and analyze machine-generated data. It can support large-scale data collection and processing with paral-lelizing analytics via the MapReduction mechanism. Therefore, it can handle distributed information with the 3V characteristics from a great amount of wearable sensors very well. In this prototype system, we used the Dropbox system as a medium to transfer distributed user's information to Splunk engines via Wi-Fi or cellular networks. Each user's mobile phone was deployed with the intelligent forwarder that carries out HMM-based state detection continuously. The forwarder can be scheduled to log the records or alert a caregiver based on emergency. Because of the HMM-based state detection, the forwarder is aware of the wearer's behaviors, and only the records associated with certain events are saved to local files according to the schedule. The files are then synced with a folder in Dropbox by using an Android-synchronized appli-cation programming interface once communications becomes available. If the Dropbox folder is shared with the big data system, Splunk can monitor any changes in the folder and index the data for analytics. It is a concern that big data pose big privacy risks [33]. Therefore, the approach using personal Dropbox folders gives individual users the right to decide if they want to keep the collected information privately or share with someone they trust; for example, they can select to share the folder with caregive

VI. CONCLUSION

A big data healthcare system for elderly people and chronically ill is presented in this paper. The system connects with remote wrist sensors through mobile phones to monitor the wearer's well-being. Collecting real-time sensor information to the centralized servers becomes very expensive and difficult, due to a tremendous number of users involved. However, such a big data system can provide rich information to healthcare providers about individual's health conditions and their living environment. Therefore, this paper discuss about an information forwarder embedded in a mobile phone. It can be configured by a user to determine under which circumstance data should be logged to the system. It uses an HMM to estimate a wearer's behavior, which include an LSH table to determine the observation probability of a state. Considering nonlinear and high-dimensional aspects of the sensor observations, the LSH table is used to improve efficiency. It can be learned by inserting sample data and queried by checking their local density.

REFERENCES

[1] "Community care statistics 2009–10: Social services activity report," NHS Inf. Centre, Leeds, U.K., 2011.

[2] *Older People in the UK*, Age UK/Help the Aged, London, U.K., 2008.

[3] H. Yan, H. Huo, Y. Xu, and M. Gidlund, "Wireless sensor network based E-health system: Implementation and experimental results," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2288–2295, Nov. 2010.

[4] S. Patel, K. Lorincz, R. Hughes, N. Huggins, J. Growdon, D. Standaert, M. Akay, J. Dy, M. Welsh, and P. Bonato, "Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 864–873, Nov. 2009.

[5] F. Zhou, J. Jiao, S. Chen, and D. Zhang, "A case-driven ambient intelli-gence system for elderly in-home assistance applications," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 41, no. 2, pp. 179–189, Mar. 2011.

[6] U. Avci and A. Passerini, "Improving activity recognition by segmental pattern mining," in *Proc. IEEE Int. Conf. PERCOM Workshops*, Lugano, Switzerland, 2012, pp. 709–714.

[7] L. Ferreira and P. Ambrosio, "Towards an interoperable health-assistive environment: The eHealthCom platform," in *Proc. IEEE-EMBS Int. Conf. Biomed. Health Inf.*, 2012, pp. 930–932.

[8] V. Venkatesh, V. Vaithyanathan, M. P. Kumar, and P. Raj, "A secure Ambient Assisted Living (AAL) environment: An implementation view," in *Proc. Int. Conf. Comput. Commun. Inf.*, Coimbatore, India, 2012, pp. 1–7.

[9] M. A. Beyer and D. Laney, *The Importance of "Big Data": A Definition*. Stamford, CT, USA: Gartner, 2012.

[10] M. Stanke and S. Waack, "Gene prediction with a hidden Markov model and a new intron submodel," *Bioinformatics*, vol. 19, no. S2, pp. 215–225, Oct. 2003.

[11] V. D. Fonzo, F. Aluffi-Pentini, and V. Parisi, "Hidden Markov models in bioinformatics," *Curr. Bioinf.*, vol. 2, pp. 49–61, 2007.

[12] P. J. Green, R. Noad, and N. P. Smart, "Further hidden Markov model cryptanalysis," in *Proc. 7th Int. Conf. Cryptogr. Hardware Embedded Syst.*, Edinburgh, U.K., 2005, pp. 61–74.

[13] L. Satish and B. I. Gururaj, "Use of hidden Markov models for partial discharge pattern classification," *IEEE Trans. Elect. Insul.*, vol. 28, no. 2, pp. 172–182, Apr. 1993.

[14] H. Lee, K. Park, B. Lee, J. Choi, and R. Elmasri, "Issues in data fusion for healthcare monitoring," in *Proc. 1st Int. Conf. PErvasive Technol. Relat. Assist. Environ.*, Athens, Greece, 2003, p. 3.

[15] M. Dong and D. He, "Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis," *Eur. J. Oper. Res.*, vol. 178, no. 3, pp. 858–878, May 2007.

[16] L. Atallah, B. Lo, G.-Z. Yang, and F. Siegemund, "Wirelessly Accessible Sensor Populations (WASP) for elderly care monitoring," in *Proc. 2nd Int. Conf. Pervasive Comput. Technol. Healthcare*, Tampere,

Finland, 2008, pp. 2–7.

[17] J. Winkley and P. Jiang, "Adaptive probability scheme for behaviour mon-itoring of the elderly using a specialised ambient device," *Int. J. Mach. Learn. Cybern.*, Oct. 2012, DOI: 10.1007/s13042-012-0134-4.

[18] J. Winkley, P. Jiang, and W. Jiang, "Verity: An ambient assisted living platform," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 364–373, May 2012.

[19] H. S. Shin, C. Lee, and M. Lee, "Adaptive threshold method for the peak detection of photoplethysmographic waveform," *Comput. Biol. Med.*, vol. 39, no. 12, pp. 1145–1152, Dec. 2009.

[20] T. L. Rusch, R. Sankar, and J. E. Scharf, "Signal processing methods for pulse oximetry," *Comput. Biol. Med.*, vol. 26, no. 2, pp. 143–159, Mar. 1996.

[21] E. Sammer, *Hadoop Operations*. Sebastopol, CA, USA: O'Reilly Media, Inc., 2012.

[22] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 2,
pp. 260–269, Apr. 1967.

[23] N. Ye, "A Markov chain model of temporal behavior for anomaly detec-tion," in *Proc. IEEE Workshop Inf. Assurance Security*, West Point, NY, USA, 2000, pp. 171–174.

[24] J. Ying, T. Kirubarajan, K. R. Pattipati, and A. Patterson-Hine, "A hidden Markov model-based algorithm for fault diagnosis with partial and imper-fect tests," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 30, no. 4,
pp. 463–473, Nov. 2000.

[25] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization tech-nique occurring in the statistical analysis of probabilistic functions of Markov chains," *Ann. Math. Stat.*, vol. 41, no. 1, pp. 164–171, Feb. 1970.

[26] J. W. Sammon, "A nonlinear mapping for data structure analysis," *IEEE Trans. Comput.*, vol. C-18, no. 5, pp. 401–409, May 1969.

[27] J. Tenenbaum, V. D. Silva, and J. Langford, "A global geometric frame-work for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500,
pp. 2319–2323, Dec. 2000.

[28] J. A. Lee, A. Lendasse, and M. Verleysen, "Nonlinear projection with curvilinear distances: Isomap versus curvilinear distance analysis," *Neu-rocomputing*, vol. 57, pp. 49–76, Mar. 2004.

[29] P. Demartines and J. Herault, "Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 148–154, Jan. 1997.

[30] D. W. Aha, D. Kibler, and M. K. Albert, "Instance-based learning algo-rithms," *Mach. Learn.*, vol. 6, no. 1, pp. 37–66, Jan. 1991.

[31] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.

[32] J. Toyama, M. Kudo, and H. Imai, "Probably correct k-nearest neighbor search in high dimensions," *Pattern Recognit.*, vol. 43, no. 4, pp. 1361– 1372, Apr. 2010.