ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

Personal Healthcare Provider by Using Information Forwarder Front End In Big Data Systems

C.MerlinLagnes[*1], M.B.PrasanthYokesh[#2], D.Rajini Girinath[#3]

P.G Student, M.E, Dept of CSE, Anand institute of higher technology, Chennai, T.N, India[*1].

Asst. prof, Dept of CSE, AIHT, Chennai, T.N, India[#2]

prof, Dept of CSE, AIHT, Chennai, T.N, India[#3]..

ABSTRACT: Wearable sensors and mobile devices are capable of monitoring and alerting the caretaker when necessary, Pervasive healthcare systems enabled by information and communication technology (ICT), can allow the elderly and chronically ill to live an independent lifestyle without relying on intrusive care programs. A big data solution is presented using wearable sensors capable of carrying out continuous monitoring of the patient. A challenge of 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, locality sensitive bloom filter is used to increase the instance based learning efficiency for the front end data processing so that they transmit only important information to the big data server for analysis, when certain behaviors happen and avoid overwhelming communication and data storage.

KEYWORDS*:* Big Data, Instance based learning, locality sensitive bloom filter, wearable sensors.

I. INTRODUCTION

Big data in healthcare has become one of the prominent areas of research in order to address the challenges encountered in healthcare environment.Big data has many implications for patients, providers, researchers and other healthcare constituents. It will impact how these players engage with the healthcare ecosystem, especially when external data, regionalization, globalization, mobility and social networking are involved. The healthcare model is undergoing an inversion. In the old model, facilities and other providers were incented to keep patients in treatment — that is, more inpatient days translated to more revenue. The trend with new models, including accountable care organizations (ACO), is to incent and compensate providers to keep patients healthy. At the same time, patients are increasingly demanding information about their healthcare options so that they understand their choices and can participate in decisions about their care. Patients are also an important element in keeping healthcare costs down and improving outcomes. Providing patients with accurate and up-to-date information and guidance rather than just data will help them make better decisions and better adhere to treatment programs [2].

The healthcare model is undergoing an inversion. In the old model, facilities and other providers were incented to keep patients in treatment — that is, more inpatient days translated to more revenue. The trend with new models, including accountable care organizations (ACO), is to incent and compensate providers to keep patients healthy. At the same time, patients are increasingly demanding information about their healthcare options so that they understand their choices and can participate in decisions about their care. Patients are also an important element in keeping healthcare costs down and improving outcomes.

Recent u-Healthcare system can measure more complicated bio signal according to both built in sensor technology and personal carrying equipment's' appearance. The system has been developed to analyze bio signal anywhere and anytime[3]. 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.

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. To hold all data accurately a kNN algorithm is used resulting in the requirement of expansive data storage to hold all data accurately.There are number of instance based classification algorithm that in some part address the speed optimization issue for real-time operation.

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.2. 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, 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.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

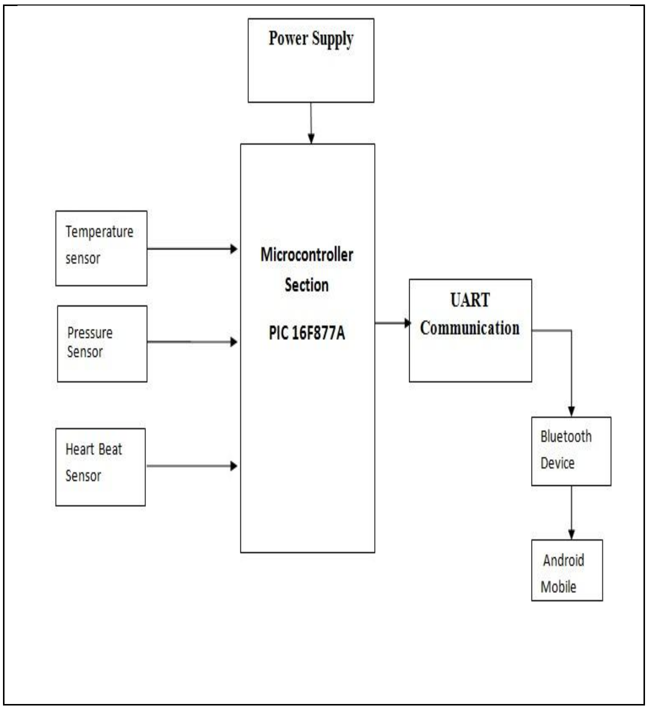

Fig. 1. Wearable sensor Hardware architecture

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 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 in this paper 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

}

Table 1. Patients log information

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.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

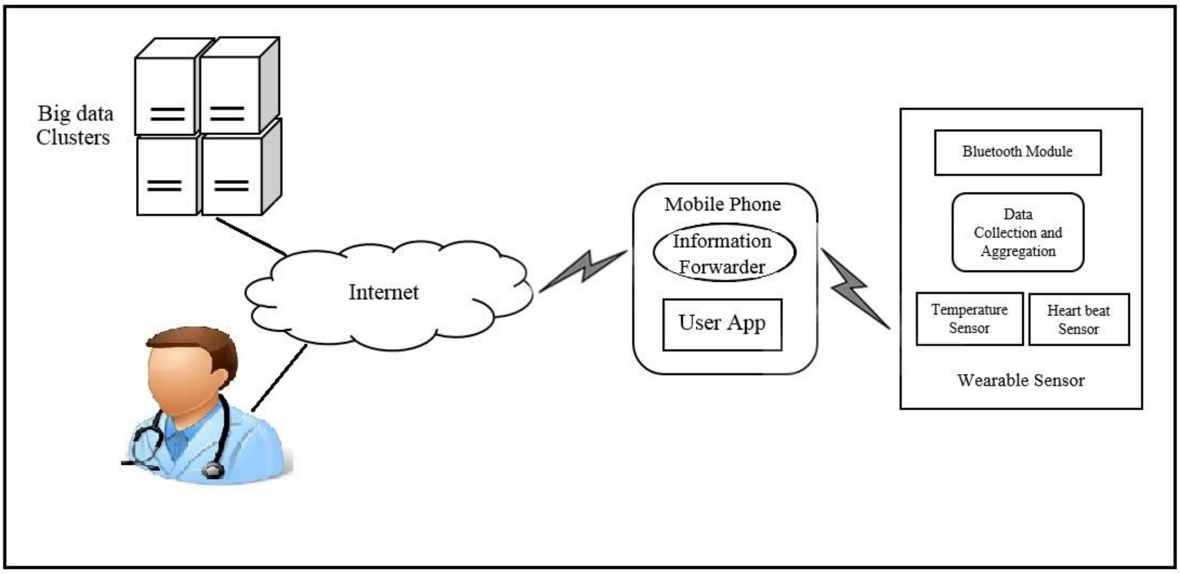

Fig. 2. Information Forwarder System architecture

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.3 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".

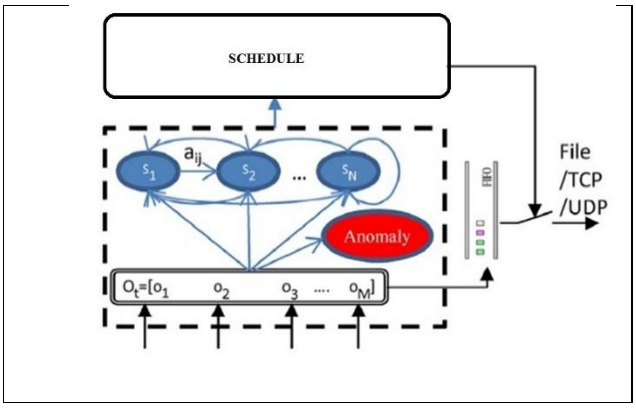

Fig. 3. State-driven information forwarder.

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.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

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:

$\delta_1(i) = \pi_i b_i(O_1),\ 1 \le i \le N \quad \psi_1(i) = 0,\ 1 \le i \le N.$

2) Recursion Step:

$\delta_t(j) =$ max $[\delta_{t-1}(i)a_{ij}]\ b_j(O_t), 1\le i\le N$

$\psi_t(j) =$ 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 $[\delta_T(i)]$

$1\le i\le N$

$q_T^*=$ 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

$$\rho = |P^* - \mu| \le \beta_1$$

In this case where ρ falls within a specified threshold $\beta_{1,}$ it indicates significant uncertainty of the identified state. The

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

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.,

$$\sum_{j=1}^{N} b_j(O_t) \leq \beta_2.$$

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 onthe 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 Fig. 3 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 information 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. 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.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

An ISO 3297: 2007 Certified Organization Vol.3, Special Issue 1, February 2015

National Conference on Computing and Communication (NC^3 2K15)

Organized by

Dept. of CSE, CARE Group of Institutions, Tiruchirapalli-620009, India on 27th February 2015

[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 EmbeddedSyst.*, 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.