Healthcare Event Aggregation Lab (HEAL), a knowledge sharing platform for anomaly detection and prediction

Alireza Manashty and Janet Light
Department of Computer Science and Applied Statistics
University of New Brunswick,
Saint John, Canada
a.manashty@unb.ca, jlight@unbsj.ca

Umang Yadav
Department of Electrical Engineering
S.V.National Institute of Technology,
Suran, India
umangyadav95@gmail.com

Abstract-**Due to the increase in elderly population, research in healthcare monitoring and ambient assisted living technology is crucial to provide improved care and at the same time contain the healthcare cost. Among existing systems, there is none robust system that can act as a bridge between different systems to facilitate knowledge sharing, so as to empower the detection and prediction capabilities of them. These systems cannot use the data and knowledge of other similar systems due to the complexity involved in sharing data between them. Storing the information is also a challenge due to a high volume of sensor data generated by each sensor. The proposed HEAL model is a platform that provides services to developers to leverage the previously processed similar data and the corresponding detected symptoms. The proposed architecture is cloud-based and provides services for input sensors, Internet of Things devices, and context providers. The ultimate goal of the system is to fill the gap between symptoms and diagnosis trend data in order to predict health anomalies accurately and quickly.**

Index Terms— **Sensor data, Healthcare knowledge sharing, Event Aggregation, Context-aware, Predict anomalies.**

I. INTRODUCTION

Population ageing, the phenomenon by which older people become a proportionally larger share of the total population, is being observed throughout the world. Globally, the share of older people (aged 60 years or older) increased from 9 per cent in 1994 to 12 per cent in 2014, and is expected to reach 21 per cent by 2050 [1]. Due to technological advancements, people in this generation are safer and live longer. Increase in ageing population will create many challenges for society and health care systems like increase in diseases, increase in healthcare costs, shortage of care givers, dependency, larger impact on society and many more. Systems that render their service in a sensitive and responsive way and are unobtrusively integrated into our daily environment are referred to as being ambient intelligent[2], [3]. Hence, ambient assisted living (AAL) is a popular topic of research in recent years. AAL tools such as medication management tools and medication reminders allow the older adults to take control of their heath conditions[4], [5]. Usually, an AAL system consists of smart sensors, user apps, actuators, wireless networks, wearable devices, software services which provides real time physical and medical information of the patient [6].

In AAL, most of the data are collected from sensors, video, cameras etc. at the low level. The result for processing systems is a very diverse collection of different types and formats of data. Processing and aggregation of these data is a major challenge especially when analyzing in real-time large streams of physiological data such as electrograph (EEG) and electrograph (ECG). An efficient system depends on improved hardware and software support [7]. Cloud computing and Internet of Things (IoT) devices are two endpoint technologies that can support the above challenge of remote healthcare and data processing.

IoT can address the problems of inter connectivity. AAL devices (such as laptops, smartphones, on board computers, medical sensors, medical belts and wristbands, household appliances, intelligent buildings, wireless sensor networks, ambient devices, RFID tagged objects) are identifiable, readable, recognizable, addressable and even controllable via the IoT [8]. The enormous amount of information produced by them, if processed and aggregated, can help in solving long-term problems and immediate prediction of emergencies.

On the other hand, patient's physiological data varies with different activities, age and it varies from one individual to another. In order to process such data and to aggregate it efficiently with other available data sources, a very large memory space and high computing power is required. A comprehensive system requires a complete knowledge repository and must remain context sensitive to satisfy different behavior profiles based on an individual's specialized needs. But performing such a massive task on a centralized model and location is failure prone and slow [9]. Cloud based and distributed frameworks are more easily scalable and accessible from anywhere specially when combined with IoT devices.

Several systems and middleware are proposed to address AAL data aggregation, processing, detection and even prediction [9]– [14]. Most of these systems are only tested in limited simulated areas and the data and techniques are not actually used and leveraged by the elderly in the way they require. Different systems have proposed different architecture of storing,

processing, aggregation and decision making. The problem identified in all of the above systems is the lack of a single platform that could act as a middle ware for such systems to provide services that all developers and healthcare systems can use to share trends, detection and prediction knowledge among each other. These systems have used a bottom up approach and lacked coordination at the central decision point.

In this paper, a top to bottom approach is used. A platform that can act as a middleware between partially processed raw data and trends and detection knowledge that is useful for not only the data provider itself, but also for other systems who might lack the necessary historical knowledge required to successfully detect and predict the anomalies from raw data is presented. The proposed HEAL model seeks to act as a bridge between different platforms. This platform provides services not only for sensors and third-parties, but also tools for developers to leverage previously processed similar data and the corresponding detected symptoms. The proposed architecture is based on cloud and provides services for input sensors, IoT devices, and context providers as well as Restful services for developers of other systems. Prototype of the model is implemented and tested on a Microsoft Azure Website.

II. MOTIVATION AND CHALLENGES

Existing models and systems mainly depends on local servers and local agent for the feature extractions, data storage, rules storages etc.[15][16] [17]. Lack of high speed processing of the complex algorithm and massive data encouraged us to use the cloud based platform[18]. In addition to the recent development in IoT technology persuaded us to focus more on cloud based IoT for AAL that helps to improve inter-operability among the sensing devices. The ultimate goal of the system is to fill the gap between symptoms and diagnosis for systems that lack sufficient training data in order to predict anomalies.

The main challenge is in the inter-operability capabilities of various sensing devices and converting all these data to high level structured data. Secondly, fusion of all these data and extraction of events from these complex data is also a difficult task. In addition, minimizing false positives alarms in real-time processing of vital signals, minimizing event detection response time, and handling large amount of patient data via aggregation and abstraction should also be addressed.

III. RELATED WORK

In the area of context awareness, many middleware centric solutions have been proposed. CoCaMAAL is one of the middleware centric cloud based solution [9] in which authors proposed different methods for processing various types of data i.e. vital signs, activity log, location log etc. Methods such as Hidden Markov models (HMM) and Holt's linear trend method for prediction and detection of anomalies are used [10]. This model lacks real time processing power, inter-operability of all the sensing devices, and security measures. Some European projects also aim at making scalable and smart home environment for AAL[19]. Most of these systems talk about a standalone application which focuses on one single context and are not much efficient for real-time processing. Apart from this Cubo Javier e. also proposed cloud based IoT platform for the AAL, but their model lacks prediction methods [8]. Many cloud based applications have also been developed like MoCAsH [20] but they lack complex event processing algorithms.

For the prediction and detection of the anomalies, a good attempt has been made by Forkan e.t. [10]. But their model lacks robust testing for real time analysis and it has no noise reduction mechanism. Other than these, Tim Van Kastern et al. have showed various methods for activity recognition such as Semi HMM, transfer learning theory, hybrid generative and discriminative methods for activity recognition [21]. This research has only been done for the activity recognition in the lower-level and it lacks the cloud based engine and its advantages [21]-[18]. All these attempts combining cloud based system and IoT for AAL [23]-[19] have focused on making an application rather following cloud based platform to take its advantage and are often focused on a specific activity recognition limiting their use.

IV. SYSTEM OVERVIEW AND METHODOLOGY

A context aware knowledge based framework is proposed here for any event anomaly detection and prediction. This framework makes it possible for third-party systems to provide high level monitoring data and obtain detection and prediction services from the system. User can work with the system to define their preferences regarding the input and output of the system. The proposed framework consists of the following 3 layers:

A. *Service Layer:* That allows other systems to connect to the framework using REST APIs and SPARQL endpoints allowing analytical systems and real-time data providers to access the system.

B. *Control Layer:* In this layer, the user can control and customize the system and define the relation of different inputs and outputs in a model that is created for a user.

C. *Distributed Cloud Based Data Providers:* Based on the data provided in the Control Layer and the data received by monitoring systems, different aggregators can prepare the data for higher level processing by context providers. Based on the input provided by context providers, newly proposed predictor components can retrieve the historical data, save it and be prepared to provide predictive parameters.
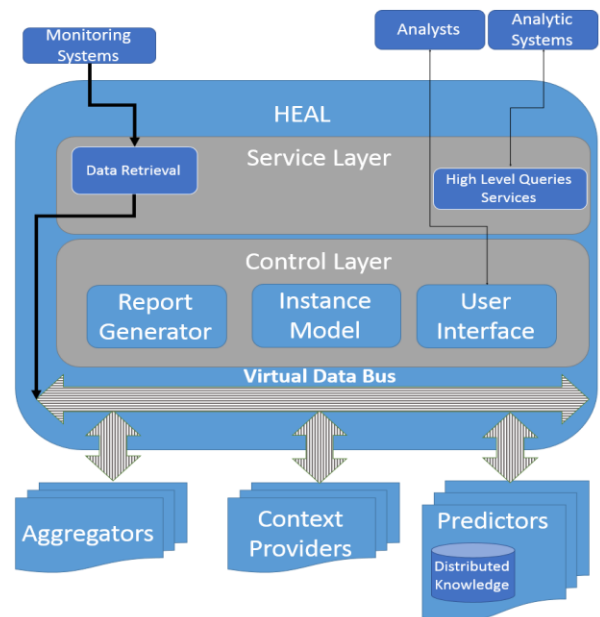

Fig.1. Proposed cloud based HEAL platform model

For the proposed distributed model in Fif.1, two new components of aggregator and predictor are added. Context providers are already defined in CoCaMAAL model [9].

1) Aggregators

Aggregators are the bridge between the real time streams of data from the monitoring systems or high level streams of data from other parts of the system, including context providers and predictors. As shown in Fig 2, these components will retrieve the data stream and using event processing language statements provided by the system user, create a different abstraction of the data, making it cleaner, more readable or more prepared for aggregation. In this level, many different formats of data are provided, a need for aggregation based on specific bit rates when required. The interpolator component will interpolate missing data to increase the data rate so that data stream can be aggregated with other streams of data easily. In the final step, the user gets another chance to define more specific data aggregation statements for final output of the component.

2) Predictors

Predictors are another novel components proposed for this model (Fig. 3). In these distributed cloud-based components, data from a specific duration of time or sequence is provided to the predictor as input. The predictor then stores the data in its data warehouse (which is managed by the predictor itself) and then using the prediction engine specified for its purpose will create a prediction model to interpolate or extrapolate the data. The system can then query the predictor to get future data, prediction error or possible trends. Having a separate distributed predictors help third parties and system analysts share different prediction engines and have specific data warehouse for their data. Some of the powerful current prediction engines are Google Prediction and PredictionIO.

The other components of the system are described here:

- Monitoring system: Monitoring system consists of various sensing devices; they collect raw data and send to the upper layer. Sensing devices can be EEG, ECG, electromyography (EMG), accelerometer, fall detector, magneto meter, gyroscope, motion sensor, blood pressure device, blood sugar sensing etc. These sensors together form a Body Sensor Network[25]. Each of these sensing devices works on low power and has capability to transmit the data wirelessly to upper layer in the cloud. Set up of the sensing devices varies individually. Sensors can easily be added or removed from the system without affecting the overall performance.

- Data Retrieval: In this layer raw data are directed to the specific aggregator for the event retrieval and more processing.

- Complex Event Processing: In this layer of the system, all the incoming real time high level signals are passed through the high level complex event processing language such as Esper and NEsper to detect anomalies in the high level data.

- Cloud based historical ware house: All the events, data and the information about the anomalies are saved in the data ware house for the future purpose. This data necessary for predictors to predict future trends and anomalies and for setting the threshold for the various vital signs for a person.

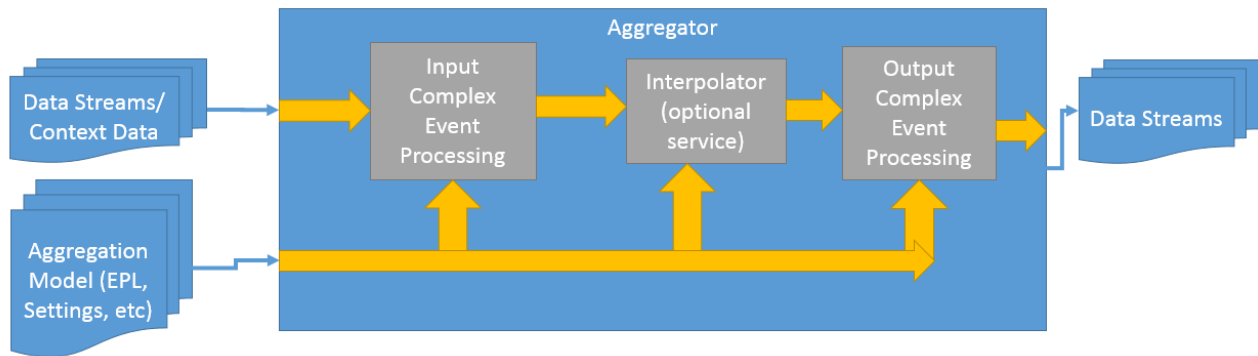- High level query services: Access endpoint for the analytic systems with REST and SPARQL endpoints.
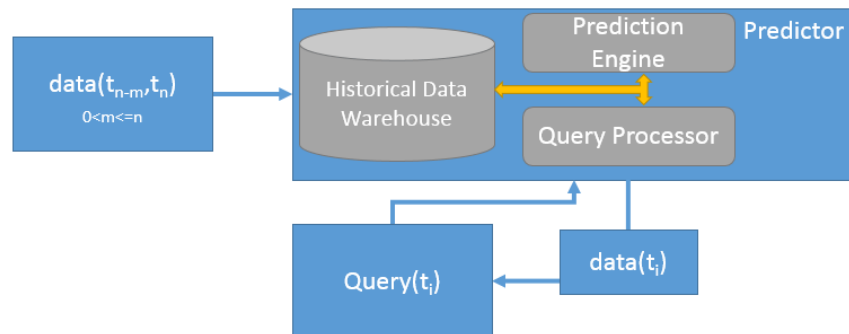

Fig. 2. Proposed aggregator model for HEAL.

Fig. 3. Proposed predictor model for HEAL platform

V. System Implementation and experimental results

Most subsystems of the proposed platform are implemented and tested with experimental data. A cloud infrastructure is designed and several applications have been deployed as separate modules on Microsoft Azure. A cloud based website is also used as the front end to test the platform (heal-platform.azurewebsites.net/preview/).

Also, rvent stream processing using NEsper is used to process real-time signals in predictors. The system itself is tested with a wink detection EEG dataset as a case study for real-time eye wink detection. This test is designed to evaluate the real time data transfer and processing performance of the system.

To send the recorded EEG data in real-time to the cloud based server, a Raspberry Pie 2 model B+ is used. Windows 10 IoT core is installed as OS to run the application on the Raspberry Pie 2, which is a good representative of biometric devices with limited processing power and resources. The system is then tested with 3 running applications on Rasbberry Pie 2, sending real-time EEG signals to the Microsoft Azure Event Hub every 100ms. Event Hub is a real-time event ingestor service that provides event and telemetry ingress to the cloud at massive scale (millions of events per second), with low latency and high reliability[26]. Each event hub partition can handle 1MB ingress and 2MB egress per second. Using default 16 partitions, our instance of Event Hub can handle 16,384 messages of size 1KB per second. The events are then consumed by an instance of Stream Analytics, which is a fully managed, real-time stream computation service hosted in Azure providing highly resilient, low-latency, and scalable complex event processing. It also helps developers to combine streams of data with historic records. Combined with Event Hubs, Stream Analytics is capable of handling high event throughput of up to 1GB/second [27]. The real-time system test indicated immediate transfer of information from Raspberry Pie to the Stream Analysis. The final analysis results and detected anomalies is then pushed to the web page Javascript client using SignalR instantly (~1s).

For anomaly detection we have used Tim Van Kastern's public datasets. It has data from three different houses [21] with recorded start time, end time of the activity and type of activity. Sensor Output is binary and represented in a feature space which is used by the model to recognize the activities performed. Using hidden Markov model to predict the anomalies based on Forkan's approach [10], the system could produce similar results in the above real-time context with the data being sent from the Raspbery Pie device.

VI. Conclusion and Future work

In this paper, HEAL-an event aggregation platform using a context aware knowledge based framework for behavioral anomaly detection and prediction- is presented. HEAL platform is an integrated system for high level behavior monitoring and supports many users and systems for long-term analysis, bridging the gap between many systems. Future works include testing multiple case studies to evaluate the performance of the system in complex heterogeneous scenarios with knowledge sharing.

VII. References

[1]   UN, "The World Population Situation in 2014," New York, 2014.

[2]   M. S. Emile Aarts, Rick Harwig, "invisible Future," *Ambient Intell.*, pp. 235–240, 2001.

[3]   E. Aarts, "Ambient Intelligence: A Multimedia Perspective," *IEEE Multimedia*, vol. 11, no. 1. pp. 12–14, 2004.

[4]   I. Qudah, P. Leijdekkers, and V. Gay, "Using mobile phones to improve medication compliance and awareness for cardiac patients," *Proc. 3rd Int. Conf. PErvasive Technol. Relat. to Assist. Environ. - PETRA '10*, p. 1, 2010.

[5]   K. a. Siek, D. U. Khan, S. E. Ross, L. M. Haverhals, J. Meyers, and S. R. Cali, "Designing a personal health application for older adults to manage medications: A comprehensive case study," *J. Med. Syst.*, vol. 35, no. 5, pp. 1099–1121, 2011.

[6]   F. Sufi, I. Khalil, and Z. Tari, "A cardiod based technique to identify Cardiovascular Diseases using mobile phones and body sensors," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*, 2010, pp. 5500–5503.

[7]   P. Remagnino and G. L. Foresti, "Ambient intelligence: A new multidisciplinary paradigm," *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, vol. 35, no. 1, pp. 1–6, 2005.

[8]   J. Cubo, A. Nieto, and E. Pimentel, *A Cloud-Based Internet of Things Platform for Ambient Assisted Living*, vol. 14, no. 8. 2014.

[9]   A. Forkan, I. Khalil, and Z. Tari, "CoCaMAAL: A cloud-oriented context-aware middleware in ambient assisted living," *Futur. Gener. Comput. Syst.*, vol. 35, pp. 114–127, 2014.

[10]   A. R. M. Forkan, I. Khalil, Z. Tari, S. Foufou, and A. Bouras, "A context-aware approach for long-term behavioural change detection and abnormality prediction in ambient assisted living," *Pattern Recognit.*, vol. 48, no. 3, pp. 628–641, 2014.

[11]   A. Copetti, J. C. B. Leite, O. Loques, and M. F. Neves, "A decision-making mechanism for context inference in pervasive healthcare environments," *Decis. Support Syst.*, vol. 55, no. 2, pp. 528–537, 2013.

[12]   WONGPATIKASEREE, "High Performance Activity Recognition Framework for Ambient Assisted Living in the Home Network Environment," 2013.

[13]   Y. Xu, P. Wolf, N. Stojanovic, and H.-J. Happel, "Semantic-based Complex Event Processing in the AAL Domain Semantic-based Event Processing in AAL," *9th Int. Semant. Web Conf.*, 2010.

[14]   A. Zafeiropoulos, N. Konstantinou, S. Arkoulis, D. E. Spanos, and N. Mitrou, "A semantic-based architecture for sensor data fusion," *Proc. - 2nd Int. Conf. Mob. Ubiquitous Comput. Syst. Serv. Technol. UBICOMM 2008*, pp. 116–121, 2008.

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

[16]   T. Taleb, D. Bottazzi, M. Guizani, and H. Nait-Charif, "Angelah: A framework for assisting elders at home," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 4, pp. 480–494, 2009.

[17]   F. Paganelli, E. Spinicci, and D. Giuli, "ERMHAN: A context-aware service platform to support continuous care networks for home-based assistance," *Int. J. Telemed. Appl.*, vol. 2008, 2008.

[18]   S.-H. Baek, E.-C. Choi, J.-D. Huh, and K.-R. Park, "Sensor Information Management Mechanism for Context-aware Service in Ubiquitous Home," *Consum. Electron. IEEE Trans.*, vol. 53, no. 4, pp. 1393–1400, 2007.

[19]   D. D. Vergados, "Service personalization for assistive living in a mobile ambient healthcare-networked environment," *Pers. Ubiquitous Comput.*, vol. 14, no. 6, pp. 575–590, 2010.

[20]   D. B. Hoang and L. Chen, "Mobile Cloud for Assistive Healthcare (MoCAsH)," in *Proceedings - 2010 IEEE Asia-Pacific Services Computing Conference, APSCC 2010*, 2010, pp. 325–332.

[21]   T. Van Kasteren, G. Englebienne, and B. Kröse, "Activity Recognition using semi-Markov Models on Real World Smart Home Data Sets," *J. Ambient Intell. Smart Environ.*, vol. 2, no. 0, 2010.

[22]   T. Van Kasteren and T. L. M. van Kasteren, "Human activity recognition from wireless sensor network data: Benchmark and software," *Act. Recognit. ...*, 2011.

[23]   C. Doukas and I. Maglogiannis, "Bringing IoT and cloud computing towards pervasive healthcare," in *Proceedings - 6th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, IMIS 2012*, 2012, pp. 922–926.

[24]   X. M. Zhang and N. Zhang, "An open, secure and flexible platform based on internet of things and cloud computing for ambient aiding living and telemedicine," in *2011 International Conference on Computer and Management, CAMAN 2011*, 2011.

[25]   G. Fortino, R. Giannantonio, R. Gravina, P. Kuryloski, and R. Jafari, "Enabling effective programming and flexible management of efficient body sensor network applications," *IEEE Trans. Human-Machine Syst.*, vol. 43, no. 1, pp. 115–133, 2013.

[26]   "Event   Hub."   [Online].   Available: http://azure.microsoft.com/en-us/services/event-hubs/.

[27]   "Exploring Microservices in Docker and Microsoft Azure." [Online].   Available: https://www.microsoftvirtualacademy.com/en-us/training-courses/exploring-microservices-in-docker-and-microsoft-azure-11796.