A Biometric Based Authentication and Encryption Framework for Sensor Health Data in Cloud

Surender Sharma and Venki Balasubramanian
School of Science, Information Technology and Engineering
University of Ballarat, University Drive, Mt Helen
Victoria, Australia
surendersharma@students.federation.edu.au, v.balasubramanian@federation.edu.au

***Abstract*— Use of remote healthcare monitoring application (HMA) can not only enable healthcare seeker to live a normal life while receiving treatment but also prevent critical healthcare situation through early intervention. For this to happen, the HMA have to provide continuous monitoring through sensors attached to the patient's body or in close proximity to the patient. Owing to elasticity nature of the cloud, recently, the implementation of HMA in cloud is of intense research. Although, cloud-based implementation provides scalability for implementation, the health data of patient is super-sensitive and requires high level of privacy and security for cloud-based shared storage. In addition, protection of real-time arrival of large volume of sensor data from continuous monitoring of patient poses bigger challenge. In this work, we propose a self-protective security framework for our cloud-based HMA. Our framework enable the sensor data in the cloud from (1) unauthorized access and (2) self-protect the data in case of breached access using biometrics. The framework is detailed in the paper using mathematical formulation and algorithms.**

Keywords—Healthcare; Data Protection; Cloud; Biosensor; Biometric; self-protective; Sensor data

I. INTRODUCTION

In recent years, healthcare has made its way from in-hospital care to remote care. Emergence of sensors in healthcare monitoring systems provides evident beneficial alternatives to both; care seeker and care provider. Biosensors[1] integrated with a Healthcare Application (HA) to form a HMA, can provide effective monitoring and diagnosis for a dispersed population, so that people can receive timely treatment and prevent diseases without frequent visits to hospitals or needing attendance of medical staff. In emergency situations, biosensors can deliver critical patient information to trauma centers by means of HMA, so that necessary treatment can be provided in timely manner. A new transparency market research report projects mobile healthcare monitoring market to be worth $8 Billion by 2019. This is a tremendous compound annual growth of 43.3% from 2013 to 2019 [1]. Both developed and emerging countries face three global megatrends the aging population, the healthcare cost and the healthcare worker shortage; these trends will have a crippling impact on their economies and societies if not addressed in the near future. The most significant trend is a rapidly aging world population. The proportion of the population over the age of 60 is projected to double from 11 % to 22 % – from 605 million to 2 billion – between 2000 and 2050, according to the World Health Organization (WHO). This global aging is already affecting the other two trends – spiraling healthcare costs and the healthcare worker shortages. While trend one is beyond an applicable technical solution; other two trends can be addressed with the emerging healthcare technologies such as telemedicine, healthcare monitoring system [8], High Definition 3D Telemedicine [9] and eTheraphy [10]. Most of these technologies rely on the sensor data either streamed in real-time or manually loaded from the sensor to the healthcare application. In either case, the volume of sensor data for analysis is enormous. In this paper we propose a framework to protect the sensors data on the fly to the cloud. The paper is organized as follows: A description about the background along with our newly build cloud-based HMA is presented in Section II. Brief overview about the sensor data is given in Section III. Section IV discusses related work in securing sensor data, while our work on security for our newly built cloud-based HMA is presented in Section V. Finally, Section VI will conclude this paper with a note on future work.

II. HEALTHCARE MONITORING APPLICATION

We developed an Assistive Maternity Care (AMC) a HA in [8]. This was designed so to assist care-staff in New South Wales (NSW), healthcare services, (such as doctors, nurses and midwives) and pregnant women to access the health records from beyond the physical location of a hospital zone, not only through desktop computers, but also through smart phones or PDAs. In the AMC application, a pregnant woman enters her blood pressure (BP) value regularly using a PDA or a desktop computer at home or at work, to allow her care-staff to monitor her condition. The AMC application was developed by establishing an Electronic Medical Record (EMR) system that contains EMRs of the patient to facilitate the clinical diagnosis. Furthermore, the AMC application is integrated

[1] Biosensors are the specific name given to those sensors that are used to measure health intrinsic data, such as cells, protein, nucleic acid or biomimetic polymers. In the context of this paper, both biosensors and sensors mean the same and are used interchangeably.

978-1-4799-5423-0/14/$31.00 ©2014 IEEE

with Short Messaging Service (SMS) gateway for sending an alert message. To monitor the BP without any manual data entry by the pregnant women, we extended this to a HMA by adding a Body Area Wireless Sensor Network (BAWSN). Therefore, in general, the HMA should be seen as two distinct parts, an HA (for example, the AMC application) and its monitoring component, the BAWSN. Because, it should be noted that the developed AMC (i.e. the HA) in the HMA should be modified considerably to accommodate the requirements of other disease such as heart attack, stroke and lymphoma[2].

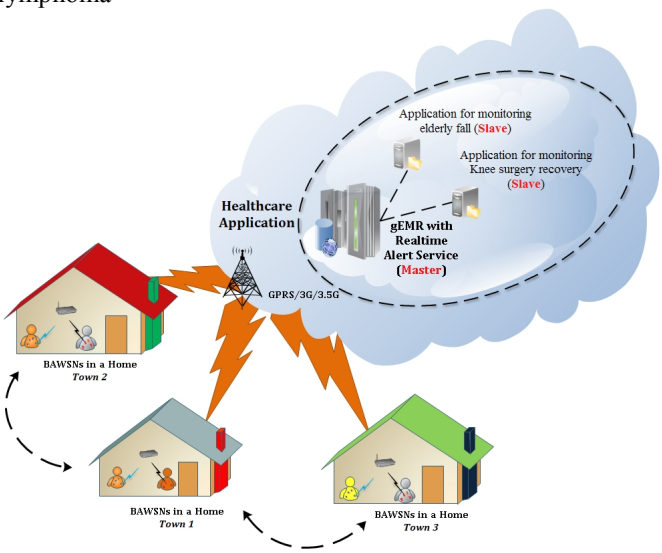

Fig. 1. A Cloud-based Healthcare Monitroing application

However, any HA would need health data only through its BAWSN. For instance, a HA is built to meet the requirements of specific diseases (say heart attack). It is infeasible to build a single HA that can meet the requirements of many acute and chronic diseases. Therefore, a cloud-based HMA was developed using Master and Slave pattern as shown in the Figure 1. A master cloud instance will have generic EMR, while the slave instances will have the functionalities specific to the medical condition or healthcare requirements – the Figure 1 shows two HAs, one to monitor the elderly fall and other HA monitors the knee surgery recovery. With a cloud-based platform, it is possible to run various instances of HAs to meet the requirements of most of the diseases that needs continuous monitoring. The following section describes the volume of sensor data from various BAWSNs.

III. PROLIFERATION OF SENSOR DATA FROM A BAWSN

In our initial AMC application we used crossbow mote sensors with MICAz processor that transferred data every 10 seconds. In this approach, a cloud-based HMA was designed and implemented using shimmer[3] sensors for BAWSN and Nectar[4] cloud for HA deployment. The Master and Slave instance are shown in Figure 2 and 3 respectively. The Salve instance monitors the elderly fall – by using three shimmer sensors Electromyogram (EMG) to detect the electrical activity produced by muscles, Electrocardiogram (ECG) to detect cardiac abnormalities by measuring the electrical activity generated by the heart and an Accelerometer to detect the movement of the patient.

Fig. 2. Master Instance shows the active slave instances

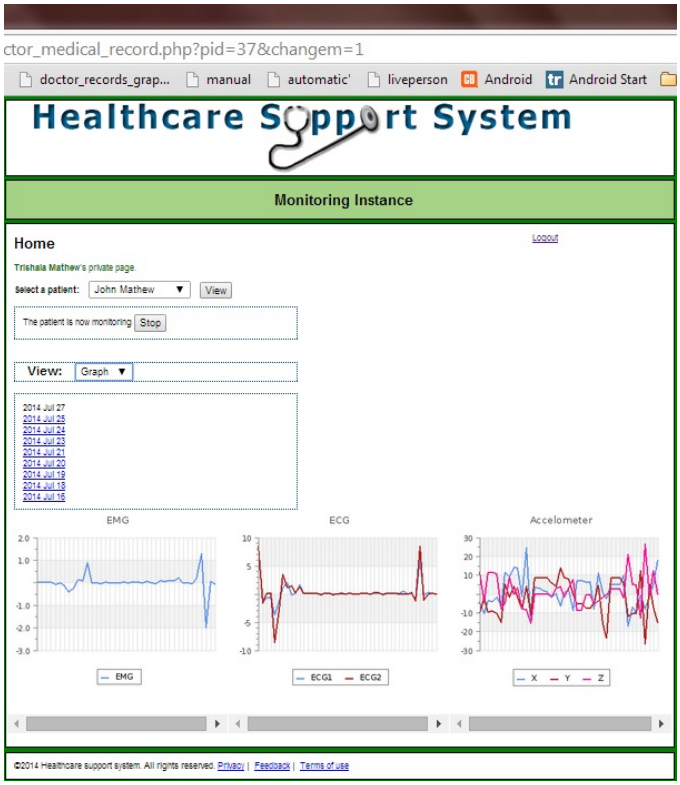

Fig. 3. Slave instance shows the EMG, ECG and Accelerometer graph for a patient

Both EMG and ECG sensors are capable of sending 8000 samples per second (SPS), while an accelerometer sensor is capable of sending 1000 SPS. At their maximum rate, of three sensors at once, the data can grow instantaneously. Table I.

[2] Lymphoma is a cancer in the lymphatic cells of the immune system

[3] http://www.shimmersensing.com/

[4] The National eResearch Collaboration Tools and Resources (NeCTAR) cloud funded by Victorian government for academic researchers.

demonstrates a sample calculation of data proliferation of sensor data.

TABLE I. DATA PROLIFERATION

1 Sec	1 Min	1 Hour	24 Hour	1 Week
5 MB	300 MB	17.57 GB	421.87 GB	2.88 TB

One can imagine the storage space required to store sensor data from continuous monitoring. Growing trend and evident benefits of cloud computing are compelling to healthcare industry to address the storage issue. But storage in the cloud has always been a concern for privacy and protection of sensitive data. There has been alarming increase in trends of data breach in last decade. Research from the Ponemon Institute's 2013[2] report analyzed data breach experiences from nearly 300 companies in nine countries and sixteen industries worldwide. The top three industries suffering the most cost were: Healthcare ($233 per lost record), Financial ($215), and Pharmaceuticals ($207). The following section discusses related work in securing the sensor data. Our approach is to address the security of sensitive healthcare data in shared cloud storage with use of biometric encryption.

IV. RELATED WORK

We have recognized that data proliferation and data protection remains two major challenges faced by healthcare industry. While shared storage and cloud computing can help to address the first issue, not much research is yet performed on self-securing sensitive data in cloud. Cheukuri et al [3] suggested to use biometric parameters to generate a key for data encryption with higher level of entropy e.g. blood glucose, blood pressure for secure communication of sensor data in a body area sensor network. Unfortunately, It is limited to body sensor networks and no further detailed work has been done. Kao Zhao et al [4] performed feasibility study of biometric encryption in mobile cloud computing suggesting use of Biometric Encryption (E_B) for identification and authentication. There are other works for architecture-based self-protection of software systems [5]. Software-Hardware architecture for self-protecting data [6] which uses hardware ID and event based monitoring for data protection. Data-centric framework for data protection by Chen and Hoang [7] uses Triggerable Data File Structure (TDFS) for self-defending data which is relevant to data self- protection but this approach does not use Biometrics. Existing methods of data protection rely on passwords, role based access control or multifactor authentication such as tokens, key cards and passwords. These methods work well until challenged by theft, loss and shared abuse. Organized crime such as a malware attack can also prove to be a big security threat for sensitive data. In next section we describe a Self-Protective Security Framework for Sensor Health Data in Cloud with use of biometric identity that makes the sensor data useless even after theft.

V. SECURED DATA ACCESS AND DATA SELF PROTECTION FOR BIOSENSOR DATA

We utilize biometric encryption for data protection. Biometric refers to human characteristics and traits that are unique to an individual. Example of biometric can be a fingerprint, retina, DNA, hand geometry etc. For our research we use fingerprints as biometric identity. A reason for us to choose fingerprint data for biometrics is to maximize contingency with multiple finger scans, should there be a bodily event-causing user to refrain from using registered finger. Finger print scans possess minimum risk and complexity to individual health as opposed to close counterparts such as retina or DNA Scan. Fingerprints are not only unique to an individual but also are unique to each finger of an individual. Benefit of using biometric information on the fly for data protection is that it cannot be faked or stolen. Only live biometric signature such as a live finger print scan in this research work can authenticate and authorize a user to access the data. Thus far biometrics have primarily been used for authentication and identification purpose i.e. to identify a user and prove that they really are, who they claim to be for identification and access control. We envisage biometric to be an enabler of privacy and protection of sensitive digital data in cloud, using biometric encryption. Unlike existing cryptographic techniques biometric encryption does not require to secure and or manage a decryption key, because individual's biometric characteristics can work as a decryption key. The sensitive data should be protected and be accessible only by authorized user even in shared or cloud storage.

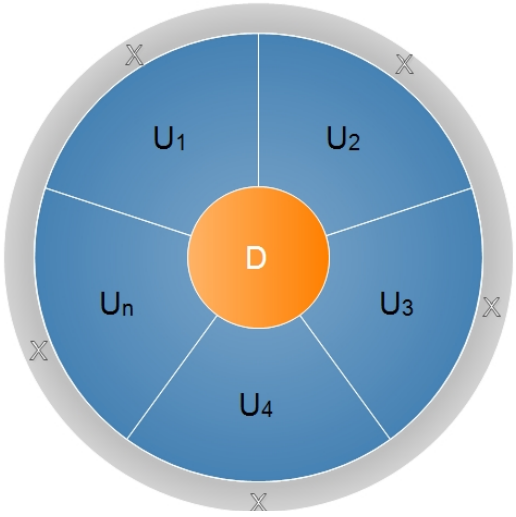

Fig. 4. Authorized data access model

In Figure 4, we assume that there are two entities data *D* and User U_n. There can be multiple users U_1, U_2…U_n, who need to have access to *D* and they are the only authorized entities to have access. *X* represents an entity that is not authorized to

access *D*. Therefore, *X* should have prohibited access and U_n should not be able to share *D* with *X*. Even if *X* get hold of *D*; the *D* should be unusable to *X*. The mathematical formulation that exemplifies the data protection using biometric is given below – the parameters used in the formulation are given in Table II.

TABLE II. PARAMETERS

U	User
D	Data
X	Unauthorized user
B_i	User's biometric identity, finger print.
B_t	Biometric template
R	Random string, R is never stored anywhere
H	Helper string, H is stored on BAESC
K^{+s}	Public Key
K^{-s}	Private Key
E_B	Biometric encryption
S	Secure Sketch
B_i'	Derivative of B_i

Let us assume that a user U_n holds the biometric identity B_i

Biometric identity B_i is not digital and noisy in nature therefore B_i' is derived from B_i, by removing noise from B_i, and close to B_i and represented as follows:

$$B_i' \sim B_i \quad (1)$$

The biometric based authentication and encryption server for cloud (BaesC) computes the secure sketch S and by using (1) the biometric template B_t is given by (2)

$$B_i' \sim B_i + S = B_t \quad (2)$$

By using (2) and the fuzzy extractor a helper string H and a random string R is generated.

Equation (2) is processed with R to generate K^{+S}

$$B_t + R = K^{+s} \quad (3)$$

The R in (3) is proportional to $B_i' \sim B_i$ for constant H. The value R is never stored and can only be regenerated with H and $B_i' \sim B_t$.

The generated K^{+S} is used to encrypt the cloud HMA data. To decrypt the data D before it can be used - only if there exists B_i' similar to B_t as given in (4)

$$\exists B_i' : B_t(B_i') \ni B_i' \sim B_t \quad (4)$$

BaesC will process B_i' with H to regenerate R as given in (5)

$$H + B_i' \sim B_t = R \quad (5)$$

By using (5) the decryption key is generated as follows:

$$B_i' \sim B_t + R = K^{-s} \quad (6)$$

It should be noted that under any given circumstances the user U_n holds unique biometric decryption key.

Our framework uses biometric based authentication and encryption server for cloud (BaesC). This can be a standalone server hosted in private or public cloud. It can also be hosted independently and integrated with cloud based HMA. Because biometric data is not digital and noisy in nature, we need to capture multiple variants of biometric identity for noise tolerance so that any scanned derivative is similar to user's live biometric identity ($B_i' \sim B_i$) and cannot be spoofed. Noise in fingerprint biometric data is a random variation of captured images that are used to generate a biometric template B_t. A biometric template is a digital representation of an individual's distinct characteristics that have been extracted from a biometric identity i.e. data obtained by a biometric system's capture device such as a finger print scanner. Biometric templates are used during the biometric authentication process. A secure sketch and B_i' makes it possible to reconstruct B_t given that B_i' is similar to B_i. Secure sketch is a collection of randomized fingerprint image capture procedures. Secure sketch makes it possible to reconstruct noisy input, so if the input is B_i and sketch is *S*. Given *S* and value B_i' close to B_i, it is possible to recover B_i. But sketch *S* doesn't give much information about B_i, so it is secure. Further we explain how our framework provides secure data access and enable data self-protection using biometric encryption.

Let us assume user U requests access to data D stored on cloud based HMA servers. U first needs to register his biometric identity (B_i) i.e. fingerprints using a finger print scanner. BaesC captures B_i' from finger print sensor, generates a secure sketch S, and processes a biometric template B_t. Here B_t is then processed to generate a random string R and helper string H. Helper string H and B_t can be stored publicly. For our research purpose all data is stored and processed on BaesC.

We utilize primitive fuzzy extractor scheme first introduced by Jules et al in [11, 12]. Fuzzy extractors are used to convert scanned fingerprint data into random strings R and helper string H that makes it possible to apply cryptographic techniques. Without fuzzy extractors biometrics can be used for authentication but not for encryption. B_t processed with random string R and secure hash algorithm-2 generates an asymmetric biometric encryption key K^{+s}. We chose SHA-2 from many of the available hash functions because it has evolved with high complexity and entropy that provides better protection against cryptographic attacks such as collision

attacks. A collision attack on a cryptographic hash tries to find two inputs producing the same hash value i.e. a hash collision. Secure hash algorithm also has a future version SHA-3 in development that would be relevant to our future research work and implementation. Successful attempts of hash collision have been made on both SHA-0 and SHA-1. Biham and Chen found near-collisions for SHA-0 — two messages that hash to nearly the same value; in this case, 142 out of the 160 bits are equal. They also found full collisions of SHA-0 reduced to 62 out of its 80 rounds. Subsequently, a collision for the full SHA-0 algorithm was announced [15]. This was done by using a generalization of the attack as in [14]. Finding the collision had complexity 2^{51} and took about 80,000 CPU hours on a supercomputer with 256 Itanium 2 processors. Preliminary results were announced in [14] about an attack on MD5, SHA-0 and other hash functions. The complexity of their attack on SHA-0 is $2^{40.}$ However, SHA-2 stands strong against attacks and thus far deemed one of the best available hash functions. As described above, Figure 5. Illustrates when a user requests access to data stored in cloud based healthcare monitoring application; HMA requests a check with integrated BaesC. If the user is already registered he or she needs to verify their identity by sending B_i' via a finger-print sensor integrated with BaesC. If $B_i' \sim B_t$ then B_i' is processed with H to regenerate R that is further processed with B_t to generate a decryption key. If the user is not registered then registration of secure sketch and biometric template needs to be done first.

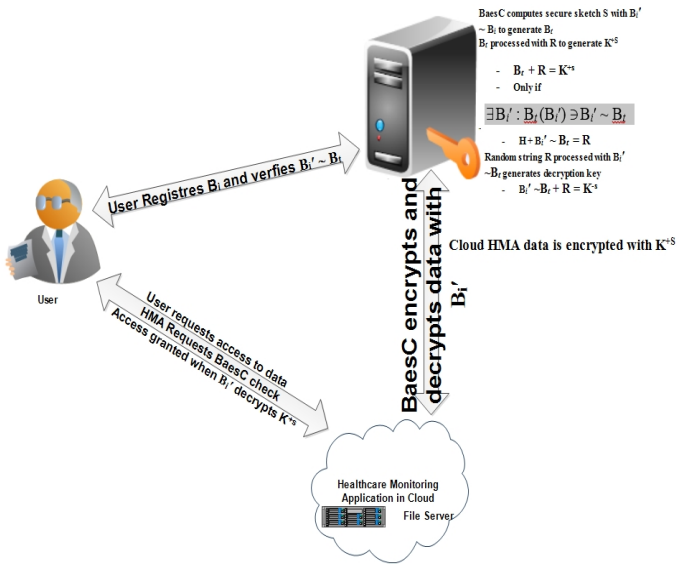

Fig.5. High Level Cloud Authentication and Encryption Framework

Since all data in the cloud HMA is encrypted using K^{+S}. Neither an unauthorized user nor an intruder can decrypt the data without a live biometric signature B_i'. Every authorized user requesting access to data first need to verify B_i' against B_t , only if $B_i' \sim B_t$ BaesC will process it with H to regenerate R. Even if H and B_t are stored publicly only the authorized user holds their biometric identity B_i that works as only decryption key and cannot be spoofed or stolen. To test the effectiveness of biometric system Matsumoto et al performed an experiment in 2002 [13] with an artificial "Gummy" finger to spoof fingerprint data with some success. Since then most biometric systems have been effectively implementing "liveness" detection mechanisms for determining if a living person is indeed presenting the sample. Variety of sensing technologies such as optical, solid state and others e.g. ultrasound are currently in use to make liveness detection more effective. These technologies utilize a combination of various biological characteristics of an individual such as [17] measurement of the pulse; reaction of skin to illumination with different wavelengths, variations of optical characteristics caused by pressure change etc. to differentiate between a fake and live figure-print. Spoofing of fingerprints these days is highly unlikely.

For affluence we represent an analogy of security alarm code system to simplify understanding of concept. As in on premises secure monitoring alarms each user is given a code to arm and disarm all or one particular zone of secure area, similar approach applies to our work of self-securing data with biometric authentication. We also optimize biometric encryption for data protection and unlike standard alarms or cryptosystems there is no key or code to remember

Thus far we have addressed secure access to data using biometric authentication and data protection via E_B. We further work through in process protection of data from unauthorized copy, deletion or manipulation. Figure 5, illustrates a very high level design of our data monitoring system in cloud. We propose an event monitoring scheme to protect the data from internal and authorized attacks. Every action in a file system generates a related event; these events are actively monitored and an alert is triggered for systems administrator if an unwanted event is sensed. Then the user session will be disconnected with a warning. We further assume a worst-case scenario that physical security system has been breached and hard disk where data is stored has been stolen or malware has been successful to relocate data. In such a scenario the data still remains encrypted and protected by using our BaesC framework that can only be used with live biometric signature. We represent an analogy of security alarm code system to simplify understanding of concept. As in on premises secure monitoring alarms each user is given a code to arm and disarm all or one particular zone of secure area, similar approach applies to our work of self-securing data with biometric authentication. We also optimize biometric encryption for data protection and unlike standard alarms or cryptosystems there is no key or code to remember that can be misused. Similar to alarm system our framework proposes real-time active monitoring.

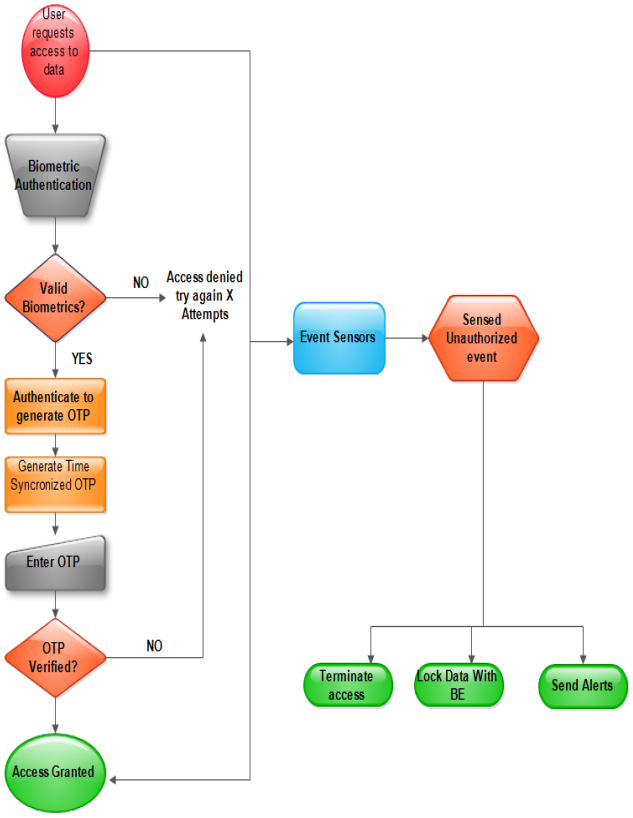

Fig.6. Event Monitoring Flow Chart for Data Monitoring in Cloud

VI. FUTURE WORK AND CONCLUSION

Super sensitive healthcare sensor data stored in shared cloud storage needs greater security and confidentiality. Current methods of secured access rely on what you know (password) and what you have (authentication device). Our framework provides who you are (Biometrics) factor to allow only authorized access to data and keep data protected in cloud shared storage all the times. Our framework achieves both secured access and data self- protection. As part of our ongoing research we are in the process of conducting cost evaluation and integration study of our framework with HMA. Future work includes integration and implementation within our cloud based health care monitoring application.

REFERENCES

[1] mHealth Monitoring and Diagnostic Medical Devices Market (Focus on medical devices with in-built plug-ins for connectivity with devices like smartphones and tablets) - Global Industry Analysis, Size, Share, Growth, Trends and Forecast, 2013 – 2019 http://www.digitaljournal.com/pr/1877980#ixzz35eeZPJi3

[2] Cost of Data Breach Study: Global Analysis Ponemon Institute, May 2013.

[3] Fen Miao, Lei Jiang, Ye Li and Yuan-Ting Zhang "A Novel Biometrics Based Security Solution for Body Sensor Networks", IEEE, 2009.

[4] Kao Zhao, Hai Jin, Deqing Zou, Gang Chen, Weiqi Dai, "Feasibility od Deploying Biometric Encryption in Mobile Cloud Computing" 8th Annual ChinaGrid Conference, IEEE, 2013.

[5] Eric Yuan, Sam Malek, Bradley Schmerl, David Garlan, Jeff Gennari, "Architecture-Based Self-Protecting Software Systems", George Mason University and Carnegie Mellon University, 2013.

[6] Chen, Y.; Jamkhedkar, P.A.; Lee, R.B. "A Software-Hardware Architecture for Self-Protecting Data" 19th ACM conference on Computer and Communications Security, Raleigh, NC, USA (2012)

[7] Lingfeng Chen, Doan B. Hoang "Active data-centric framework for data protection in cloud environment", 23rd Australasian Conference on Information Systems, 3-5 Dec 2012, Geelong

[8] V Balasubramanian, DB Hoang, NF Ahmad "SOAP based Assistive Care Loop using wireless sensor networks" IT in Medicine and Education, 2008. ITME 2008. IEEE, 2008.

[9] Andrew Stranieri , Richard Collman and Ann Borda, High Definition 3D Telemedicine: The Next Frontier? Stud Health Technol inform 182, 2012, 131-141, Published by IOS Pres.

[10] Stephen Vaughn, A Stranieri, Coalescing Medical Systems: A Challenge for health informatics in a global world. 159, Global Telehealth 2010, Published by IOS Press.

[11] A.Jules and M.Wattenberg. A Fuzzy Commitment Scheme. In Proc ACM Conference, Computer and communication security , page 28-36, 1999.

[12] A.Jules and M.Sudan, AFuzzy vault Scheme, In IEEE international Symposium on information theory, pages 408-421, 2002

[13] Tsutomu Matsumoto, H. Matsumoto, K. yamada, S.Hoshino: Impact of artificial "Gummy" fingers on fingerprint systems. Yokohama National university,2002.

[14] Stallings,William,Cryptography and Network Security, Prentice Hall, 1999.

[15] M.R.Verbauwhede. Secure Integrated Circuits and Systems 81, 2010, published by Springer Science plus business media, LLC, New York.

[16] M. Sandström, Liveness Detection in Fingerprint Recognition Systems [Diploma thesis], Linköping University, Linköping, Sweden, 2004.

[17] Martin Drahansky, Michal Dolezel, Jan Vana, Eva Brezinova, Jaegeol Yim, and Kyubark Shim, New Optical Methods for liveness detection on Fingers, BioMed research international, 2013.