American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-5, Issue-8, pp-213-220
www.ajer.org
Research Paper Open Access

Intelligence based Outlier Disclosure for UDBMS at Sensor: Wireless Body Sensor Network

Chetna Laroiya[1], Dr. V. B. Aggarwal[2]
[1](Computer Science Department,Jaagn Institute of Management Studies affiliated to GGSIPU
[2](Dean Computer Science Department,Jagan Institute of Management Studies affiliated to GGSIPU, India

ABSTRACT: *Life style always has great impact oneveryone's health. Todays time demands for the use of wireless sensor technology to assist in patient monitoring at home or at hospital.Sensors are working in constraint environment majorly with respect to energy. Major research area for wireless sensor networks are energy consumption at sensor for computation or transmission, energy efficient communication protocol, , complex operation (outlier detection / data mining) algorithm, complex operation technology (web services or middleware) and place to support complex operations (at different locations i.e sensor/PDA/server), UDBMS architecture for complex operations (at sensor / PDA/ server). This paper views current implementation of complex operations at sensor data for UHealtCare domain. Our focus is on the implementation of outlier detection technique at sensor device for regular monitoring of sugar and BP level in the body to know weather the combination of the two leads to critical condition which could be notified as an alert. As per our finding UDBMS has been developed for deployment at the sensor to dynamically create and store databases at sensor and also to execute ad-hoc query at sensor itself. Work has to be done to ensure the viability of implementation of complex operation at sensor node.*

Keywords: *WBSN, Joint probability, Middleware, Web Services, UDBMS, Tight –coupling, Loose-coupling*

I. INTRODUCTION

Any Wireless Sensor Network (WSN) is a collection of nodes and gateway each oneequipped with radio transmitter,(radio module) processor (microprocessor module), sensors (sensor module) and memory (memory module) which is deployed for various applications i.e. in geographical monitoring (seismic activity), habitat monitoring (tracking of animal herds), transportation (traffic monitoring), medical diagnosis [4].

Introduction of technology in healthcare domain leads to ease in our life style in terms of self monitoring of physiological parameters of our body along with remote monitoring by the medical experts. One may get advice from experts anywhere on the physiological imbalance of one's body.

With the introduction of sensors in healthcare, a new term WBAN has come to existence which consists of a no of sensors working on RF technology with in a range of approximatelytwo meter as per the chosen WBANtechnologies. A PDA with in 10m – 100m range of sensors works as router to connect WBAN to WMAN through WPAN.

1.1 Some working models under Ubiquitous HealthCare Monitoring

Panasonic and Jewish home life carelaunch Telemedicine Pilot Program that usestelevisions as an interactive portal to connect patient and hospital staff for healthcare monitoring.

Patient's real time data stream isanalysed at the remote end. Infosys has a Service-Oriented Architecture (SOA) platform-based disease management and lifestyle management portal solution, Affordable Care Act ("The Affordable Care Act" or ACA). Amedisys (Bringing home the continum of care) an American organization working under Affordable Care Act ("The Affordable Care Act" or ACA) to improve the quality of healthcare at home.

1.2 Placeholders of Complex Operations

Complex operations are possible in following ways (i) distributed processing of sensor data, (ii) centralized processing of data from sensor, (iii) Hybrid approach. [7]

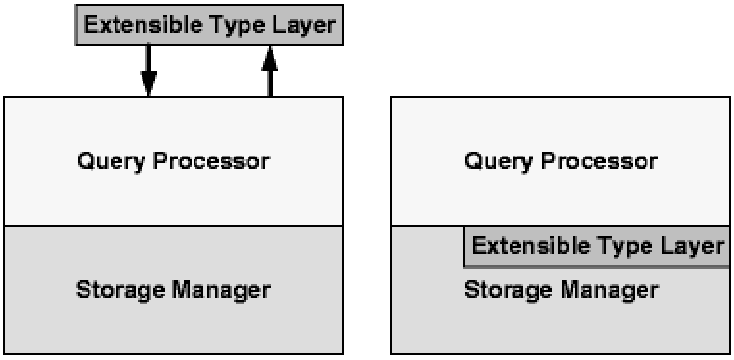
1.2.1 Centralized processing of sensor data

1.2.1.1 Through DBMS

Complex operations are supported by the UDBMS deployed at the sink/server node. Earlier UDBMS i.e. TinyDB or Cougar did not create database, so every sensor reading was to be transmitted to sink. TinyDB can filter sensor readings based on threshold. Cougar supports in-network data aggregation over sensor reading across a particular region.

Various UDBMS possible at mobile/PDA/ notepad or computer are: SQL anywhere, iAnywhere, Oracle Lite, Microsoft SQL Server Compact, SQLite, IBM DB2 Everyplace (D2e), POSTGRES. Without a provision for database at sensor,ad-hoc aggregation query over single sensor data like what is the maximum blood sugar level one had in last few hours.

could not be answered. Sensor readings are processed at server. DBMS at server might be loosely coupled or tightly coupled to support complex operations ex. Odysseus/ OpenGIS.Loosely coupled DBMS supports adding new data types or user defined functions using PL/SQL on the top of query processor as an extensible type layer. Tightly coupled DBMS support user defined functions as a part of storage manager.



(a) Loose-coupling. (b)Tight-coupling.

[5]

1.2.1.2 Through Middleware or Web Services

To perform complex operations at sink node, web services are created as separate abstract layer over existing DBMS architecture at the sink node and functionalities offered by the web service are deployed as services. Using (SOM) service oriented middleware over sensor network infrastructure, Ubiquitous applications get services of sensor network through service modules written in any programming language . These service modules are loosely coupled or tightly coupled with the DBMS at sink node (in each case under the limitations imposed by the coupling mechanism) as we see in Odysseue UDBMS.

The services offered through the service oriented middleware could be data aggregation, security, network reliability, and management services. The interaction between services is through the XML. Services can be developed using technologies i.e. CORBA (Common Object Request Broker Architecture), SOAP (service oriented architecture protocol), DCOM (*Distributed Component Object Model*), RPC (remote procedure call), web services. Services could be developed to perform real time operations on data i.e. outlier detection or non-functional requirements (OoS i.e. reliability of WSN, scalability of WSN)

VIS : B-VIS (Burapha Vehicle-Infrastructure System) is an SOM to use distributed RFID sensor networks for vehicle-based applications. It supports real-time vehicle fleet tracking and control and also supports road traffic monitoring systems[1] .

1.2.2 Distributed Processing of sensor data

Complex operations are implemented while sensor reading is moving through the wireless sensor network. Processing algorithm under this approach require the middleware or web services to work on flat file rather than the database in DBMS. Task of merging of sensor recordings from individual senor to ensure quality of Service(QoS), aggregation on sensors readings ,communicating the task from the sink to the sensors, troubleshootingto combat technical challenges raised due to the hardware's, security of communication between sensor& the sink are done by the middleware. WSN middleware is a software to process large amount of sensor records for analysis.

1.2.3 Sensor

No trace of implementation of outlier detection at sensor node is revealed from the literature review. Provision for adhoc query at sensor is mentioned. For adhoc queries non-SQL compliant language is proposed.

If a small footprint size DBMS which support ad-hoc query is installed at the sensor then they have dependency on external application to get query plan for adhoc query execution at sensor. The approach removes parser & optimizer from the sensor RDBMS in order to keep footprint size of suitable for constraint memory in sensor.[6]

II. RELATED WORK

2.1 Significance of Outliers:

Outlier is sensor reading with low probability of occurrence. Outliers if detected might be either supressed or be amplified. In case of an event and error, the readings of the sensor is significantly different from othersensor's readings but error never convey any significant evet rather leads to communication overhead. In Uhealthcare application rise or downfall in sugar & Blood pressure level might indicate some body anomalous condition. Outliers are to be supressed for data cleansing. Amplification of outlier will help to find rare patterns in the domain. Lot many of outlier detection algorithms are proposed till date.

2.2 Dealing with outliers

When working with sensor network, two most significant constraint exist are (i) the energy overhead incurred while sending sensor reading across the WSN (ii) the time delays incurred between sensor data collection and getting the processing outcome. Both the disadvantages increase in proportion to the average distance a data item needs to travel. In centralized approach to support complex operation the delays becomes two fold because of the packet collisions, as colliding packets are to be retransmitted thus increasing round trip delay and same me consuming extra bandwidth. Sensors transmit far more than the required number of packets therefore expend their energy comparatively faster. If sensor data could be stored at the sensor node by creating database then communication overhead from sensor node to sink node will be reduced. Further surveys show that cost of a query to select 100 tuples is less than the cost of single packet transmission over sensor network. For the physiological body sensors, critical body condition indicated by the outliers could be managed much quickly if outliers are interpreted at the sensor or cluster (PDA) node itself.

As per our findings from literature survey no UDBMS except Antelope, LittleD at sensor store databases.[6][9] Outlier can be local or global outlier. If some of the readings from the outliers are anomalous with respect t the rest of the readings from the sensor , the outlier is first order or local outlier. If some of the sensors reading is substantial different from the neighbouring sensors then the outlier is called global or high level outlier. [7] As per my finding no outlier detection algorithm has been proposed till date which work on the stream of data from one sensor, whereas Number of outlier detection algorithm with mathematical models are given which detect outlier over the reading from N number of homogeneous sensors deployed in a region[8].

For an outlier detection algorithm to work at sensor (a device with constraint on energy/power to be consumed both in transmission& computation, memory to provide space to complex SQL compiler/ PL/SQL compiler) we need to address following issues. (i) UDBMS to create & stored database at the sensor. (ii) either non-SQL compliant language where data-definition//data –manipulation queries are written in more explicit form over nested or co-related query or an inbuilt query optimizer & translator.(iii) Use of indexes where ever possible in the execution of query statement. with prime focus on joins.

III. PAPER CONTRIBUTION

We propose a UhealthCare model which works in preventive mode i.e. data from 2 physiological sensor (to measure blood sugar & blood pressure) is analysed at sensor/cluster head for any anomalous/critical body condition through a batch script .

3.1 Requisition of Outlier Detection at sensor:

The approach will accelerate the performance of outlier detection in three ways

a) Database at sensor will reduce transmission energy consumption which directly depends on underlying wireless technology (Zigbee or Blutooth low energy) and is also based on the distance sensor data has to travel. Fact could be revealed from the underlying table

	Max. Nodes per Master	Peak Current Consumption	Range	Data Rate	Topology	Relative Cost
Wi-Fi: IEEE 802.11b	32	~100mA	100m	54Mb/s	Star	Medium
Bluetooth	7	30 mA	10m	1Mb/s	Star	Low
ZigBee	100s	30mA	50m	250kb/s	Star, Mesh	Low
Bluetooth Low Energy	7+	15 mA	100m+	1 Mb/s	Star/Mesh Scatternet	Low

Figure 1: Energy requirement of popular wifi technologies

[3][6]
Energy = Power * Time (1)
Total Energy consumption at a sensor is a sum of
1. Transmission Energy consumed = Transmitting Current * Time* Voltage
2.Receiving Energy consumed = Receiving Current * Time* Voltage
3.Idle Energy consumed = Idle Current (mA) * Time* Voltage
4.Sleep Energy Consumed (if sensor is inactive for a while) = Sleep Current * Time *Voltage

b) Proposed approach will also reduce the need for continuous asynchronous data transfer to sink node. Outlier could be written back at cluster head/PDA in asynchronous mode. Sensor data is analysed at sensor in a sliding window of N sensor readings to be processed in N epoch duration.
c) At sensor outlier disclosure will reduce round trip delay. Sensor reading are to be recorded only if the sensor value of any of the 2 physiological sensor data (blood sugar sensor , blood pressure sensor) in the current sliding window is above threshold which is as per the standards. propagation delay term defines the time taken for the first bit to travel from the sender to the receiver as below.

Propagation time = Distance / propagation speed (2)
The propagation speed depends on the wireless medium of the communication.
Packet transmission time = Packet size / Bit rate (3)

3.2 Outlier Disclosure Algorithm

Algorithm pseudocode is here based on table 1.
Step 1: Set W = M
Step 2: Update the DS and DB from sensor reading at every epoch if (PiS <PiSU . and PiS > PiSL) and (PiB <PiBU . and PiB > PiBL)
Step 3: For each k ϵ DS Qk = ci/WC + cl/WW and For each k ϵ DB Qk = ci/WC + cl/WW
Step 4: Concatenate k-sugar and k-blood pressure at each t to get Jq = { (ci/Wc+cl/WW) (where k is sugar sensor reading) * (ci/Wc+cl/WW) (where k is blood pressure sensor reading) } and append to R.
Step 5: If (ci/Wc) != 0 then set flag =1 in R
Step 6: Data point in R with min Jq and flag =1 is outlier if ci (PiS) is not in $Prev_O$ and ci(PiB) is not in $Prev_O$
Step 7: Feedback from PDA to sensor if outlier is insignificant then [reading from R and Jq] ∪ $Prev_O$

Table I: Notation Table

DS	set of ∀ sugar sensor readings
DB	set of ∀ blood pressure reading
RS(.,.)	Probability calculation function of sugar sensor record X (first argumant) with reference to the Ds
RS(.,.)	Probability calculation function for blood pressure sensor record Y (first argumant) with reference to the Ds
PiS	ith sugar sensor reading
PiB	ith blood pressure sensor reading
PiSL	threshold for ith sugar sensor sugarreading
PiBB	threshold for ith blood sugar sensor reading
W	sliding window size
WW	no. Of sliding windows exemined till date
WC	current wimdow
ci	no of sugar / blood pressure readings in given range for W
cl	no of sugar / blood pressure readings till date in given range for WW
Qk	probability of sugar reading k
t	timestamp, sugar and blood pressure sensors are synchronized to read at same t
k	reading from sugar or blood pressure sensor
R	result set based on DS and DB for current W
Jq	Joint probability
$Prev_O$	Record set denoting earlier outliers detected but mentioned 'NA' in response from PDA to sensor

3.3 Asymptotic analysis of algorithm

Proposed outlier disclosure algorithm is implemented as sql script with following set of queries:

a) Insert query (2 times) based on corelated query to get M (window size) records
* O(M)
b) Update query(2) based on co-related query (i.e. select count(*) from table) to update of M fields of table
As count(*) has time complexity O(1) , complexity is O(1)+O(1) +. ....N times

Execution of a external query based on corelated query is fast provided the fields being referred by the correlated query has indexes, else the complexity for update will be (if update query set 1 record) 1*(x+y+z ....M) i.e table is scanned again & again for each row returned to external query

c) Insert query (1) based on equi join

Here equi-join can be based on hashing on key attribute of sensor table. Key for hash function is timestamp of the sugar records. Hashing based Equi join will create linked list of values with given timestamp. Complexity of join is O(M), O(1) for every M record where M is window size to be processed in N epoch duration.

3.3.1 Hashing based Join

t_{sec} : timestamp in seconds

h(): hashing function

e_{SIZE} : epoch size =M

$h(k : t_{SEC})= : t_{SEC} \% e_{SIZE}$

Hashing function *h()* is applied on t_{SEC}. Collision is handled through chaining.

h(k: t_{SEC})

Memory address to store sensor reading k

Record No.	*TimeStamp*	*SugarLevel*	Next Pointer	*BPLevel*
1				
2 .. W				

At sensor even if we can create table and databases sliding window size is very small so hashing join will increase performance. Over merge join, merge join with indexes and hash table join without the index on the join attribute .

d) Update query (N) based on correlated query for x=3 fields

update table set field_1=(Select field_T from T), field_2 =(select field_Q from Q), field_3=(select field_P from P) * (select field_O from O,field_L = const where condition_on_tableN * (O(1)+O(1)+O(1)+O(1))

e) Select query (1) based on criteria at one field of table O(M)

f) Delete query to delete whole table (3) 3 * O(M)

As per asymptotic time analysis total time complexity of the script is O(N) where M (window size)> N.

3.4 *Scenario Analysis*

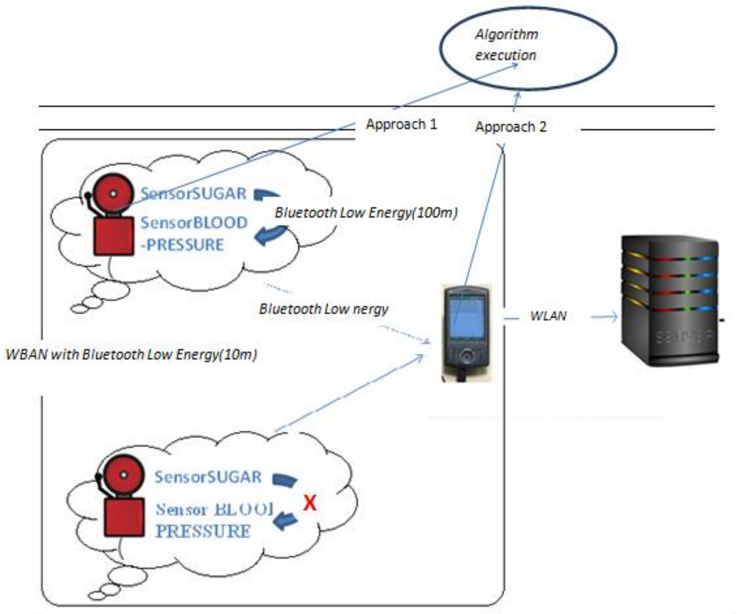


Figure 2: Approach 1has outlier disclosure at body sensor, Approach 2 has outlier disclosure at PDA beyond 100m renge

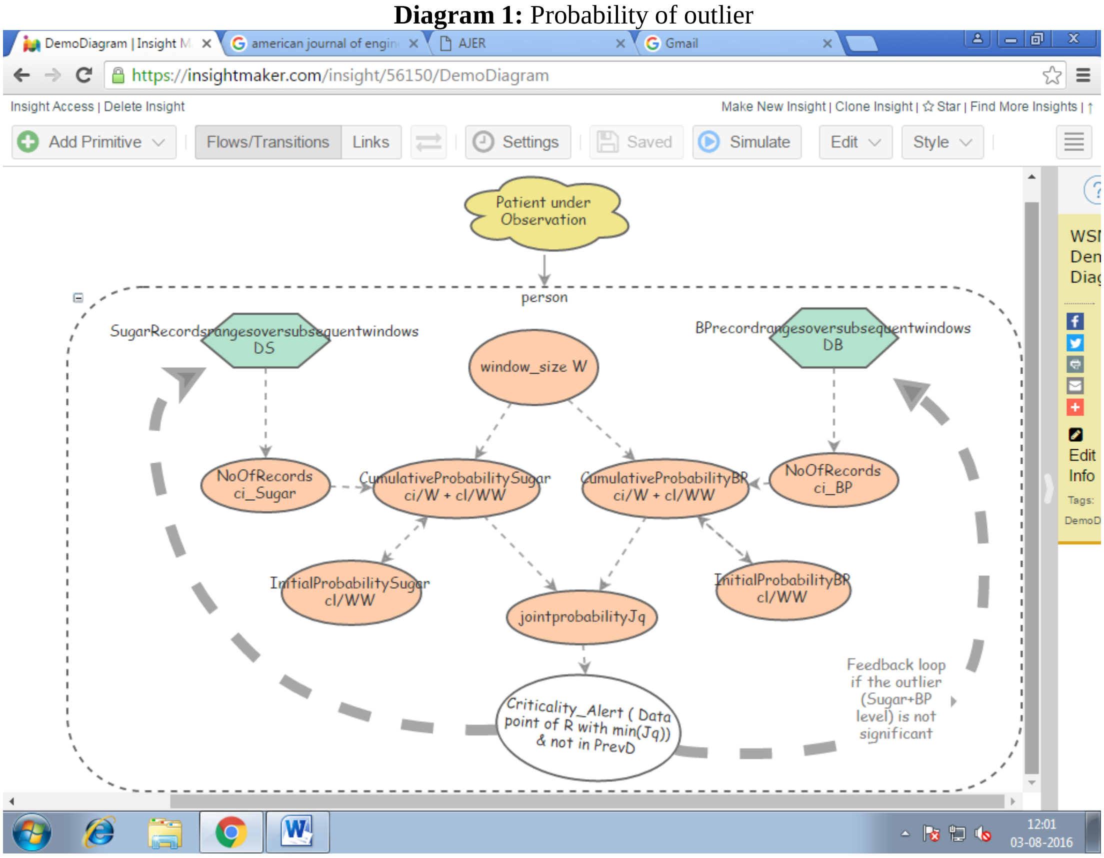
Table II: Comparison of outlier disclosure algorithmat two differnet locations in WBAN(provided PDA is in 100 m range as a component of WBAN) based on three parameters (cpu time, energy, footprint size)

Approach	1	2
CPU time	Time taken to call execute scriptsqlite.cmd at.sensor	Time to execute scriptsqlite.cmd at PDA
Energy	a) M reading from sensorSUGAR to sensorBLOODPRESSURE ver ***Bluetooth Low Energy*** b) Outlier BLOODPRESSUREsensor to PDA over ***Zigbee*** ***Energy= (15 mA** * t_{S_S}***V*M) + (30mA** * t_{S_PDA}*** V* 1 (equation 1))***	a) M reading from sensorSUGAR to PDA over ***Zigbee*** ***b)*** M reading from sensorBLOODPRESSURE to PDA over ***Zigbee30mA***if 100+ sensors are monitored by PDA ***or Bluetooth low power 15mA*** if 7+ sensors are monitored by PDA **E=30 mA * t_{S_S}*V *2M**
Foot ptint Of UDBMS	Algorithm is implemented in simple SQL commands. PL/SQL compiler not to be used.	
Asymptotic time analysis	Hashing based join is has asymptotic time as O(M or N) where M & N are no. of records in tables to be joined	

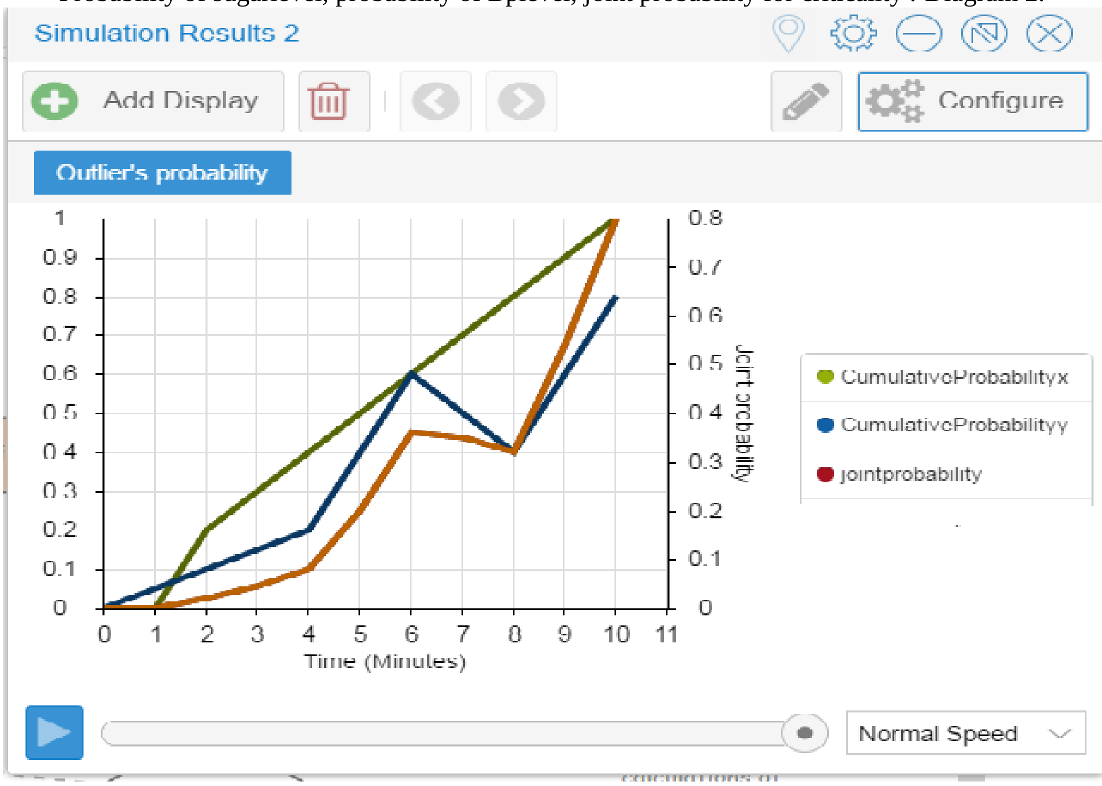
3.5 Simulation Results

Correctness of outlier detection strategy based on ’s joint probability theorem is verified from the result of simulation over Insight Maker.

Diagram 1: Probability of outlier



Probability of sugarlevel, probability of Bplevel, joint probability for criticality : Diagram 2.



Probability based outlier disclosure: Diagram 3.

Time	CumulativeProba...	CumulativeProba...	jointprobability
0	0	0	0
1	0	0.05	0
2	0.2	0.1	0.02
3	0.3	0.15	0.045
4	0.4	0.2	0.08
5	0.5	0.4	0.2
6	0.6	0.6	0.36
7	0.7	0.5	0.35

IV. CONCLUSION

Early detection of the outlier even at sensor device will clearly low down the energy consumption at sensor, round trip propagation delay in outlier detection and the size of asynchronous data transmitted from the sensor. This may also alert the patient and help him to escape critical body situations. Real time simulation results reveals that the algorithm proposed could be implemented at the energy constraint devices as time taken to execute outlier disclosure script on a window size of 50 records is time to ping a machine. Moreover outlier disclosure algorithm is implemented through simple SQL queries.

REFERENCES

[1]. Nader Mohamed, Jameela Al-Jaroodi, Service oriented middleware approaches for wireless sensor network, System Sciences (HICSS), 2011 44th Hawaii International Conference Kauai, HI on 4-7 Jan. 2011, 1530-1605, Pg 1-9

[2]. Levent Gürgen, Claudia Roncancio, Cyril Labbé, Shinichi Honiden, Data Management solutions in Networked Sensing system in Wireless Sensor Network Technologies for the Information Explosion EraVolume 278 of the series Studies in Computational Intelligence, Springer-Verlag Berlin Heidelberg, , 2010, Pg 111-137

[3]. Surthineni. Ashok, Dr. R.V.Krishnaiah, Overview and Evaluation of Bluetooth Low Energy: An Emerging Low-Power Wireless Technology, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 9, September 2013, ISSN: 2277 128X, Pg 737 - 740

[4]. Huasong Cao, Victor Leung, Cupid Chow, Henry Chan,Enabling technologies for Wireless Body area network: A survey and outlook,IEEE Communications Magazine (Volume:47 , Issue: 12), December 2009, 0163-6804, 0163-6804, Pg 84-93

[5]. Kyu-Young Whang, Jae-Gil Lee, Min-Soo Kim, Min-Jae Lee, Ki-Hoon Lee, Wook-Shin Han, Jun-Sung Kim,Tightly-coupled spatial database features in the Odysseus / OpenGIS DBMS for high-performance,GeoInformatica, October 2010, Volume 14, Issue 4, pp 425–446

[6]. Graeme Douglas,, Ramon Lawrence, , LittleD: A SQL Database for Sensor Nodes and Embedded Applications, SAC '14 Proceedings of the 29th Annual ACM Symposium on Applied Computing, March 24-18 2014, Pg 827-832

[7]. Ritika, Tarun Kumar, Amandeep Kaur, Outlier Detection in WSN- A Survey, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 7, July 2013 ISSN: 2277 128X, Pg 6-9 – 617

[8]. Dylan McDonald, Stewart Sanchez, Sanjay Madria , Fikret Ercal,, A Survey of Methods for Finding Outliers in WirelessSensor Networks, Journal of Network and Systems Management, January 2015, Volume 23, Issue 1, pp 163-182

[9]. Nicolas Tsiftes, Adam Dunkels, A DBMS in every sensor , SenSys'11, November 1–4, 2011, Pg 316-329