TEST Engineering & Management

March-April 2020
ISSN: 0193-4120 Page No. 2078 - 2087

A Novel Computing Approach for Data Analysis in Smart City Healthcare – A Case Study of Heart Disease

Swapnaja Hiray
Department Of Computer Science and Engineering, GITAM Visakhapattanam, India
srhiray.scoe@sinhgad.edu

Dr. Bbramaramba Ravi
Department of Information Technology GITAM Visakhapattanam, India bhramarambaravi@gmail.com

Article Info
Volume 83
Page Number: 2078 - 2087
Publication Issue:
March - April 2020

Article History
Article Received*: 24 July 2019*
Revised*: 12 September 2019*
Accepted*: 15 February 2020*
Publication*: 18 March 2020*

Abstract

Innovative idea of smart cities gained much significance as it facilitates the economic growth and increase in GDP. Development in ICT (Information and Communication Technology) is key role player in smart city concept.It is backbone of smart city environment. Basically it is considered as integrated systems and many life aspects are linked together to get the improved life style. Main domains integrated in the 'Smartcity' system are power, transportation, buildings, public securities, waste and water management ,city governance and health care. Smart city applications provide quality of life solutions related to these domains. Smart city applications basically involve pervasive and ubiquitous environment and Internet of Things can make it possible. Continuously sensing and actuating components produce the large amount of speedy ,real time data And hence this ecosystem comes under "BIGDATA" ecosystem.

In this paper mainly we discussed the architectural framework for any smart city application and tried to find out how much data traffic is there with traditional frame work. One prototype smart health application to conceptualize the problem has been developed. With this prototype we tried to get some traditional theoretical analysis related to the data traffic taking place and then wrote hypothesis for computational delay. Then we suggested the use of new improved architectural model taking advantage of edge computing technology with improved architectural framework . And tried to find out performance of data analytics algorithm at the edge device .

Keywords – *Big Data, Smart Healthcare applications, Internet of Things, , Smart City ,edge computing and machine learning algorithms.*

I. Introduction

The conceptualization of Smart city [13], very much context driven. It varies not only from city to city but also country to country. It depends on various factors such as city development, resources available, willing ness to change and reform and aspiration of citizens. In smart city mainly information and communication technology is used to solve all social problems and it involves residents in it. Though citizens are at focus , this idea later evolved to an urban space for good business opportunities[1] . As technology progresses this smart city environment becomes ubiquitous technology

March-April 2020
ISSN: 0193-4120 Page No. 2078 - 2087

implementation across the city.So,now it is part of every object in city.

By studying different Smart city applications related to analytics it can be characterized by following ways [4][8][9]

- 24* 7 either hardware or software sensors are collecting data and available for various type of analysis . This issue will allow us to characterize these applications as " Big Data applications"
- Automation wide spread is prime most requirement. These are sensing + actuation applications in the complete ecosystem. That is include both cyber plus physical component. These can be characterized as Cyber physical systems.
- Cities are geographically wide spread hence give rise to pervasive and ubiquitous applications.
- Pattern recognition is mainly used for decision making and actuation.
- Proactive measurements are equally important that of reactive like and these are carried out by prediction analysis.

The main focus areas for Smart city application domains are transport, energy, environment, agriculture, public culture, healthcare and retail. Whatever application domain we are using we can generalize the Smart city applications framework.

In our previous paper [1] we have discussed smart city application Framework by taking Smart healthcare as our case study and developed one application for the same.

In smart healthcare application data is collected from various data sources. It can be either a static data or real time data . It is mainly a multimedia data we are collecting in the smart healthcare. It is either a structured data like laboratory reports or unstructured data like X-Ray images from various Bio-sensors. Collection of data needs preprocessing and integration of data from heterogeneous environment is very important.

Generally in Smart-healthcare applications,first step is collection of data from different sensors or sources our task is to use this data for better healthcare operations. In healthcare operations there is wide spectrum starting from personal care, emergency management, hospital management, telemedicine to robotic surgery and clinical diagnosis. Data analysis is important part of the Smart-healthcare. The main obstacle in data analysis in smart-healthcare is this heterogeneity and volume of data . Also many times we require to perform real time data analytics and fast response to our analysis problem. We can easily prove that data generated and analyzed in smart healthcare is 'Bigdata' . If we apply traditional data processing and storing techniques for analysis purpose that may affect our quality of results . Due to advances in technology and wearables, ample amount of storage is available and possibility of collection of all kind of records . This will increase horizontal growth of data and this sensor collected data is called as 'Time series data'. And this data is significant for predictions. According to MarkWeiser vision healthcare industry is becoming ubiquitous in nature .Very low cost ,low power body sensors are playing important role in healthcare industry. These nodes are intelligent and playing important role in data analysis in healthcare industry[9].

In our research we have taken advantage of this wearables and tried to optimize the

March-April 2020
ISSN: 0193-4120 Page No. 2078 - 2087

framework by leveraging data analysis at edge of network instead in cloud.

II. Literature Survey

As this research work is in continuation of our previous research work we published in paper, we like to discuss some very important points regarding architectural frame work and want to take our survey to next level of studying performance of various machine learning algorithm on smart devices at the edge. We have two categories for edge computing .IoT enabled and Non-IoT usecases. In white paper [18] it is clearly mentioned. Smart city applications, as we discussed in our previous paper are IoT enabled use-cases. Device management, data analysis, AI and machine learning are major use-cases related to smart city applications at edges. In this we have focused more on data aggregation use case. In IoT enabled use cases devices are continuously sensing and collecting data. Due to large amount of data many replications of data will be there. We can take weather prediction system for example. In that sensors in same vicinity sending multiple temperature reading. Same usecase in vehicle monitoring system where vehicle on same platue sensing and giving lots of data. Its not necessary to send all these readings to centralized server. And edge computing is playing important role here. Edge device will select important data to send .It can be done with effective data aggregating or using certain algorithms.

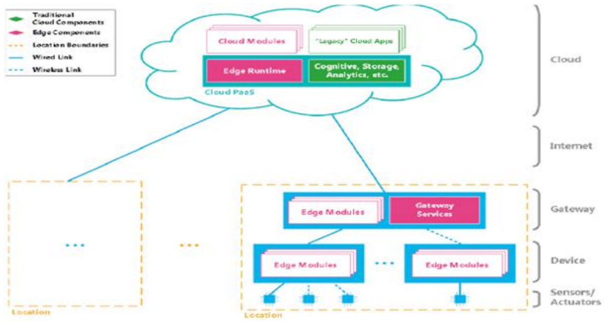

Fig 1. Architecture edge computing (Source:IBM)

As shown into Fig 1. Edge computing leverages the cloud environment at edges . Sensors/actuators, devices /nodes collecting data (e.g. Smart phones), gateways at particular locations are main components we have to focus mainly for developing edge computing algorithms. As technology advances cost of sensors reduced dramatically and hence cost of data collection is reduced dramatically.

In Smart healthcare system sensing generated data is Big Data and difficult to interpret. Need of special data analytics techniques needed for this purpose in smart healthcare[4]. These analytics techniques can be categorized like descriptive techniques, diagnostic analysis, predictive analysis, prescriptive analysis. The major use cases in the smart healthcare can be bigdata driven decision making, patient profile analytics, effective public health strategies, population management, cost reduction, social network for patient. scalable epidemiological studies, genomic analytics, evidence based medicines, improved remote patient monitoring,

Unstructured data analysis, supporting clinical decision system. To all these systems advanced analytics can be applied. These results can help us to identify the critical patients, whose health can be

March-April 2020
ISSN: 0193-4120 Page No. 2078 - 2087

improved by taking proactive care and lifestyle changes. The classic example is diabetes management By taking patients profile along with HER(Electronic Health Record) a classification model can be build and patient with higher risks can be identified and preventive care and lifestyle changes are suggested.

In traditional data analysis classification, regression, clustering, association rules these are main techniques we are using. Any ML model can be explained with three basic components :

Representation: This is how knowledge can be represented. For example, rulesets, decision trees, graphs or neural networks.

Evaluation: This is model performance component. This can be accuracy, precisions and recall, entropy, information gain, divergence .

Optimization: To improve the performance of model we have to use various optimization techniques. For example, convex optimization, combinatorial optimization, convex optimization, constrained optimization.

In any smartcity application we have basically cloud computing and internet of things technologies used. In this architecture actual data processing occurs far away from the data providing source. Fog computing is emerging paradigm which leverages some computing part to lower layer in given architecture. Concept of promoting edge computing is allowing collection, preprocessing and analysis of data near to sensors, instead of at cloud centers.[2]. In this type of architecture all machine learning activities are done in proximity to the source of data and thus it minimizes the overall communication and network delay. We can take a self driving car in smart city . All AI algorithms are used as a tool to complete mostly all human tasks.It has so many sensors as it generate 25 gigabyte of data per hour. This data has to be send to cloud server for further analysis and then given input to AI algorithm. In cloud architecture this excessive latency could is responsible to fail vehicle to react many sudden events. In edge computing this is carried out locally in side the car this will make decisions locally and avoid unnecessary delay. Now the problem is all machine learning and decision making algorithms are designed to work on highend servers. So all smart city applications heavy lifting of data to cloud is still carried out.To get the correctness of algorithm still we have to train the model with millions of record in the cloud and then deploy on the edge for the inference. In some domain like healthcare where decision emergency is time sensitive we are started using this new technologies . This is motivation to our research to develop such a new technique which can be able to take decision at edge independently without training model or with locally trained algorithms. This is far more cost effective require less bandwidth, storage cost, and computational delay. Swim is such type of one software developed by startup which promots this type of edge analytics[4].

Azure is another software mainly used in stream analytics , Azure Stream Analytics (ASA).It works on IoT Edge and help developer to deploy real time analytical algorithm along with AI near the sensing devices. It is designed for low latency, efficient use of bandwidth. Control logic can be deployed on the device and can be used for many industrial problems. Some of

March-April 2020
ISSN: 0193-4120 Page No. 2078 - 2087

the usecases tackled by Azure are as follows:
Control use cases : Manufacturing safety systems has to take immediate decision on operational data . Emergency medical services also. These are applications which demand low latency.
Cloud connectivity: Mission critical systems such as Mining equipment control, offshore drilling , nuclear reactor temperature control, these need continuous cloud connectivity for controlling the operations . So limited connectivity can be accommodated with these edge algorithms.
Bandwidth limitations: Automatic car driving, data produced by jet engines these are high volume data this will consume high bandwidth hence stream analytics at edge device is useful for this.
Compliance generation : Some daily compliance reports may require data to be aggregated locally these can be solved with edge computing
In Smart city which are mainly Cyber physical systems Bigdata is collected time to time and forwarded on cloud for further analysis. We are well-known about the 3 V's related to Bigdata hence dimension reduction ca
n be achieved at the edge. Our further paper explains the architecture we are suggesting for all sort of Smartcity applications.

III. Methodologies

In paper one we have explained the cloud computing architecture for Smart city applications.
It has three layers : Device or WSN , Routers/Gateways, Cloud Database .
Now in our proposed architecture Device /WSN layer is bit modified we have one intelligent edge device near the sensor network which is responsible to coordinate various sensors participating in a usecase of a particular application.
And then important data is forwarded to the next layer that is routers /Gateways and cloud. Where main machine learning model resides[10].

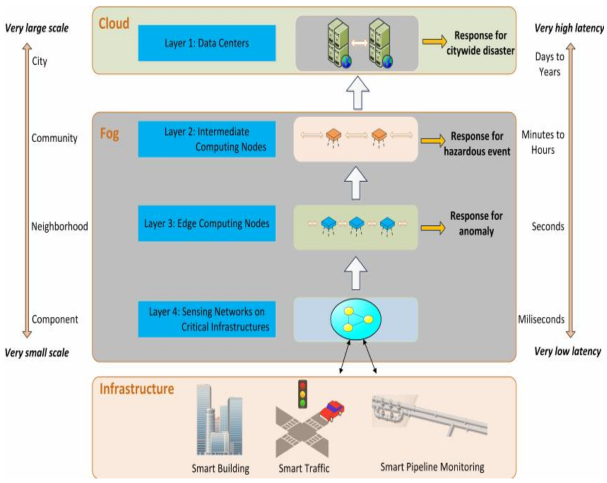

Fig.3.1 Fog Computing Architecture in Smartcity applications (Curtsy: Gerald Hefferman)

Our first objective of research is to develop prototype application with this suggested architecture. We have developed one Smart healthcare application which uses this concept for Emergency Management in Heart diseases. We have taken healthcare use case purposefully because latency in data transfer is very important in healthcare applications.

IV. EXPERIMENTATION AND HYPOTHESIS DEVELOPMENT

Use case : We are considering Smart healthcare Heart disease managing usecase.
In our usecase we are using Fitband as wearble device and smart phone as edge

March-April 2020
ISSN: 0193-4120 Page No. 2078 - 2087

device. First of all our model will, start collecting data from three categories of stakeholders Patient, Doctor and Lab technicians. EHR with unique patient ID will be stored on cloud. Initially we have developed one classifier in cloud which will give certain decision say, severity level of heart disease in this usecase. Now next part is on device. Most critical patients are handled by another module which run on mobile which will collect data from Fitband given to patients. In our case we are collecting Pulse rate, Heart bit, Sleep patterns for 24 hr. And Emergency alerts are developed. Near by Dr recommendations based on GPS locations are done .

Hypothesis:

We can state following postulate:

1. In any given smart city application data processing delay is directly proportional to total communication delay to transfer data from edge device to cloud server + Cloud server computational delay + Edge device computational delay. Suppose that D is total data processing time required , T_c time required to transfer m amount of data , T_d required for computation at edge device , T_{CC} time required for computation on cloud server.

$$D = T_c + T_d + T_{cc}$$

2. With appropriate data analytics methodologies we can have considerable reduction in computational overhead at edge device and amount of data transferred, so overall data processing speed can be increased.

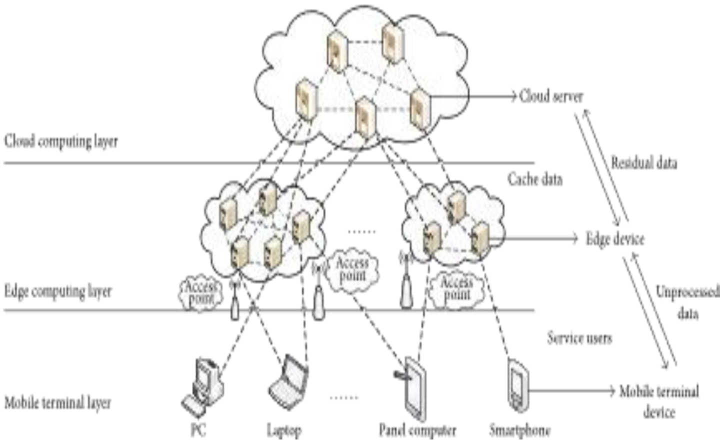

Fig 2. Edge Computing Concept (Courtesy: [3])

In smart city healthcare applications, for given usecase traditionally data processing delay or computational overhead mainly depend on following factors:

- Communication delay for data transfer from devices at edge to cloud server.
- Machine learning computational overhead at cloud server.

But introducing edge computation data processing time as stated above is function of

Delay at devices layer + delay at cloud layer + communication delay at edge devices to cloud server.

The major assumption we can make in this is

the delay at cloud layer and communication delay from edge devices to cloud server as a function of data size. Now we want to restrict our problem for data processing in above scenario can be formulated as follows:

$V=\{z_1,z_2,\ldots,z_i,\ldots,z_m\}$ (1)

is a vertex set, Where zi is a set of edge devices and m is the number of edge devices.

$E=\{e\ z1,z2,\ldots,ezi,zj,\ldots,ezm-1,zm\}$ (2)

TEST Engineering & Management

March-April 2020
ISSN: 0193-4120 Page No. 2078 - 2087

Is edge set e ez_i,z_j is the communication link between edge nodes z_i and z_j.$\tau z_i,z_j$ on the edge is the communication delay between edge nodes z_i and z_j,and the size of its value is randomly generated

Now as per our postulate we want to work on latency part of the data processing. To reduce communication delay between the edge layer to cloud layer we want to reduce the amount of data transferred from edge layer to cloud layer.

The computation delay at any device i can be described by the amount of computation x_i assigned to it. And the facts true for it , as amount of computations increases the computational delay will be increased at the device,

DEdge i = 1/ Vzi(aix^2 i)

where Vzi is the device computing capability of i, xi is device data be processed by i, and ai is a predetermined real number between 0 and 1.

Now delay in the cloud computing layer can be calculated, it is computational delay of server + communication network delay. For the cloud server j if the amount of data it needs to deal with is yj and the computing capability is Vj its computational delay can be expressed as :

Dj = yj / vj (j= 1, 2…. n)

To show the feasibility of the second postulate we have carried out the experimentation with given usecase.

V. EXPERIMENTATION

To implement emergency heart disease management we have prototyped our new edge computing architecture . We have two modules : 1. Machine learning module at Cloud layer to classify the severity level. After researching we have established decision tree classifier in cloud server. UCI data set for 680 patients is used for training and testing. Main attributes for the data set are as follows[6]:

Name	Type	Description
Age	Continuous	Age in years
Sex	Discrete	0 = female 1 = male
Cp	Discrete	Chest pain type: 1 = typical angina, 2 = atypical angina, 3 = non-anginal pain 4 =asymptom
Trestbps	Continuous	Resting blood pressure (in mm Hg)
Chol	Continuous	Serum cholesterol in mg/dl
Fbs	Discrete	Fasting blood sugar>120 mg/dl: 1-true 0=False
Exang continuous Maximum heart rate achieved	Discrete	Exercise induced angina: 1 = Yes 0 = No
Thalach	Continuous	Maximum heart rate achieved
Old peak ST	Continuous	Depression induced by exercise relative to rest
Slope	Discrete	The slope of the peak exercise segment : 1 = up sloping 2 = flat 3 = down sloping
Ca	Continuous	Number of major vessels colored by fluoroscopy that ranged between 0 and 3.
Thal	Discrete	3 = normal 6 = fixed defect 7= reversible defect
Class	Discrete	Diagnosis classes: 0 = No Presence 1=Least likely to have heart disease 2= >1 3= >2 4=More likely have heart disease

Table 1: Attribute Information

These are attributes required to predict heart disease.

2.After knowing the existence of heart disease emergency handle for critical patients is starting .and this module we shifted on mobile . By using any machine learning algorithm on mobile , by taking input from various sensor devices from fit bands wearables we can go for various predictions. As Fitband has its own APIs we can not collect data on device so we simulated it with Ardino kit acting as fitband.

Here we are using simple association rule for various predictions and as per requirement EHR on cloud is going to change.

We are achieving following objectives with this experimentation [11]:

March-April 2020
ISSN: 0193-4120 Page No. 2078 - 2087

1. Checking feasibility of new suggested edge computing architecture for Smarthealthcare application.
2. Effect of scalability on data analytics algorithm(Decision tree as case study)
3. Calculating computational complexity of various data analysis techniques on mobile as edge device.

Implementation of Association rule generation on edge Device :

We have Fitband/Arduino kit sensing the data and stored at mobile locally. we are collecting data for 24 hrs. to generate the association rules and then these rules are used for various recommendations either Machine learning algorithms like K-means used in collaborative filtering in cloud or simple alarm generation locally. Following Itemset we are considering which is event based. Above threshold value by particular sensor event will be generated. We have used Apriori algorithm due to its simplicity.

I1 -> When heart rate above threshold

I2 -> When sleep pattern not accurate

I3 -> When B.P. is above threshold

I4 -> When body temperature level above required level.

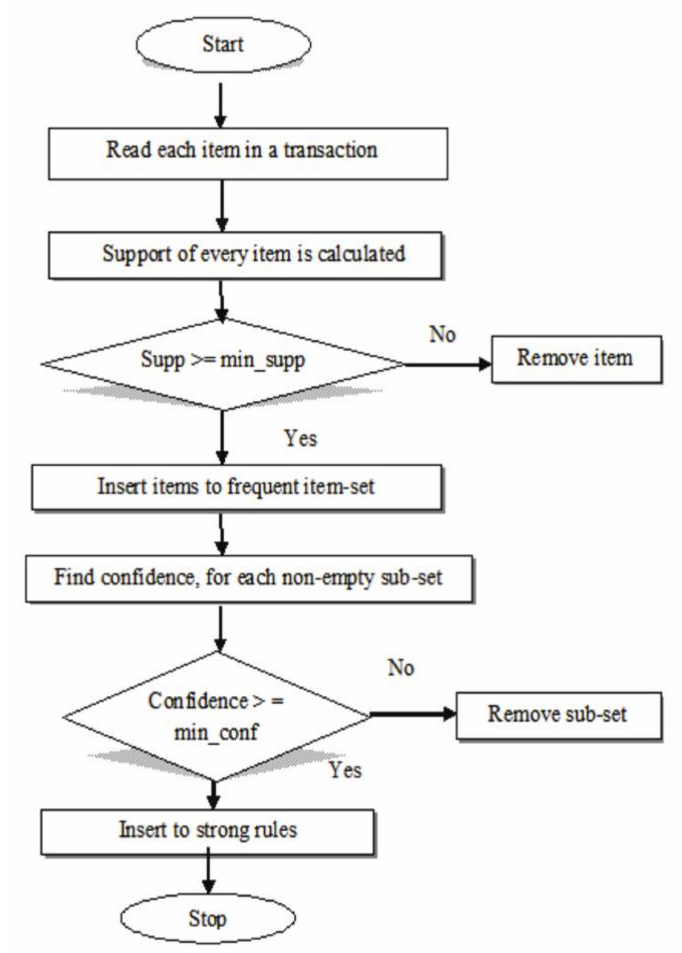

Fig3: Association Rule Mining Flowchart

VI. RESULTS AND DISCUSSIONS

Following are implementation :

- Collection of data
- Predicting severity
- Report generation
- Start fit bit model for severe patient
- Collect data of 24 hours by taking half hour sensors
- Collect data for severe patients by monitoring them for 24 hours
- Store data locally on phone memory
- Patient with less severity to be recommended best doctors
- Patient with more severity to be recommended nearest best doctors and also a alert to be sent to their relatives

March-April 2020
ISSN: 0193-4120 Page No. 2078 - 2087

In line of our Objectives following are few results :

1. Able to generate EHR by creating patient ID and storing it on cloud.
2. Generations of association rules for particular patient.

On Analysis part computational complexity for this Apriori algorithm is as follows:

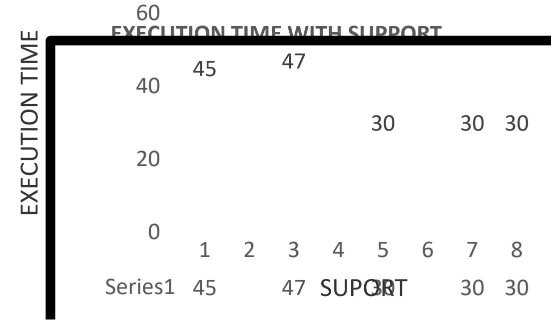

Fig 4 : Execution Time with Support

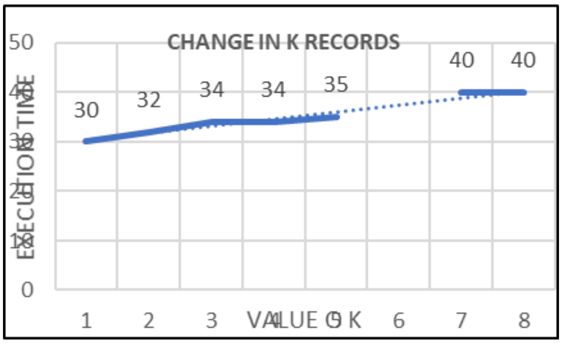

Fig 5 : Execution Time with K Transactions

This is time complexity . space Complexity will depend how many sensors you are taking.

This will help us to calculate by putting this algorithm as a part of mobile app to execute at device edge that is at client edge we avoid unnecessary data traffic and transaction to cloud server with very little computational complexity.

VII. CONCLUSION

By leveraging data analytics work at edges in Smart city applications we have many benefits. Firstly unnecessary traffic over the communication line can be avoided. Efficient use of bandwidth will take place due to this type of architecture. If internet is not available due to poor connectivity the emergency decision can be handled by the devices.

Our next research objective is in cooperating new Complex event detection methodologies at the device level in this architecture and making more efficient architectural framework for Smart city applications.

VIII. REFERENCES

[1] Andriopoulou F, Dagiuklas T and Orphanoudakis T (2017) Integrating IoT and Fog Computing for Healthcare Service Delivery, In: Keramidas, G., Voros, N., Hübner, M. (eds.) Components and Services for IoT Platforms, Springer International Publishing Switzerland

[2] Bresnick J (n.d) How Fog Computing May Power the Healthcare Internet of Things, Healthanalytics. Available at: http://healthitanalytics.com/features/how-fog-computing-may-power-the-healthcare-internet-of-things. Accessed on May 16, 2017

[3] Bonomi F, Milito R, Natarajan P and Zhu J (2014) Fog Computing: A Platform for Internet of Things and Analytics, Bessis, N., Dobre, C. (eds.), Big Data and Internet of Things: A Roadmap for Smart Environments, Studies in Computational Intelligence, Springer International Publishing Switzerland

[4] Chen M., Hao Y., Hwang K., Wang L., Wang L.Disease prediction by machine learning over big data from healthcare communities IEEE Access, 5 (2017), pp. 326-337

[5] Jason Burke. Health Analytics: Gaining the Insights to Transform Health Care. Wiley, 1 edition, July 2013

[6] Keyue Ding and Kent R Bailey et al, " Genotypeinformed estimation of risk of coronary heart disease based on genome-wideassociation data linked to the

electronic medical record", International journal of BMC cardiovascular Disorders, Vol.11,2011

[7] Leonidas G. Anthopoulos. 2017. Understanding Smart Cities: A Tool for Smart Government or an Industrial Trick?. Springer International Publishing. DOI: 10.1007/978–3–319–57015–0

[8] Maksimovic M and Vujovic V (2017) Internet of Things based e-health systems: ideas, expectations and concerns, In: Khan, S., Zomaya, A., Abbas, A. (eds.) Handbook of Large-Scale Distributed Computing in Smart Healthcare, Scalable Computing and Communications. Springer, Cham

[9] P.Shanmuga Sundari, Dr.M.Subaji, Big Data Analytics In Healthcare System For Diverse Perspectives, International Journal of Pharmacy & Technology, ISSN: 0975-766X

[10] Stojmenovic and S. Wen, "The Fog Computing Paradigm: Scenarios and Security Issues," in Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, 2014, vol. 2, pp. 1–8.

[11] Shou-En Lu and Gloria L Beckleset al, "Evaluation of risk equations for prediction of short-term coronary heart disease events inpatients with long-standing type 2 diabetes: the Translating Research into Action for Diabetes", International J

[12] S.DonandM.Dugki. Medicalcyberphysicalsystemsandbigdataplatforms. IntheMedical Cyber Physical SystemsWorkshop, 2013.

[13] Talari, S.; Shafie-Khah, M.; Siano, P.; Loia, V.; Tommasetti, A.; Catalão, J. A review of smart cities based on the internet of things concept. Energies 2017, 10, 421. [CrossRef]

[14] Tanwar, S.; Tyagi, S.; Kumar, S. The Role of Internet of Things and Smart Grid for the Development of a Smart City. In Intelligent Communication and Computational Technologies; Springer: Singapore, 2018; pp. 23–33.

[15] V. Stantchev et al., "Smart Items, Fog and Cloud Computing as Enablers of Servitization in Healthcare," J. Sensors & Transducers, vol. 185, no. 2, 2015, pp. 121–128.

[16] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile Edge Computing A key technology towards 5G," ETSI White Pap., no. 11, 2015.

[17] GSMA White paper , "Opportunities and Use Cases for EdgE Computing in the IoT"