Health Care System Using Cloud Computing

Keerthana.R[1], Aswin.J[2], Armaan Ahamed.S[3], Arvind Kumar.G[4], Harish.G[5]
[1]Assistant Professor
[1, 2, 3, 4, 5] Saranathan college of Engineering, Department of IT, Trichy.

Abstract- *One of the fastest growing fields is health care industries. The medical industries have huge amount of data set collections about patient details diagnosis and medications. To turn these data into useful pattern and to predicting coming up trends data mining approaches are used in health care industries. The health care industry collects huge amount of healthcare data which are not "mined" to find out hidden information. The medical industries come across with new treatment and medicine every day. This paper explores the good decision making technique to predict health related information. Experts and Researches believe that cloud computing technology can greatly improve the level of healthcare services. Currently, many experts have done fruitful research in this domain. It contains three major parts, portable medical devices, intelligent terminals and cloud services platform. The results of this study show that the technologies of cloud and data mining can be used to enhance the performance of the healthcare system so that humans can then enjoy various smart healthcare applications and services.*

Keywords- Cloud computing, Data Mining Techniques, Healthcare.

I. INTRODUCTION

HE purpose of data mining is to extract useful information from large databases or data warehouses. Data mining applications are used for commercial and scientific sides [1]. This study mainly discusses the Data Mining applications in the scientific side. Scientific data mining distinguishes itself in the sense that the nature of the datasets is often very different from traditional market driven data mining applications. Data mining algorithms applied in healthcare industry play a significant role in prediction and diagnosis of the diseases. There are huge number of data mining applications are found in the medical sector such as Medical device industry, Pharmaceutical Industry and Hospital Management. To find the useful and hidden knowledge from the database is the purpose behind the application of data mining. Popularly data mining called knowledge discovery from the data. The knowledge discovery is an interactive process, consisting by developing an understanding of the application domain, selecting and creating a data set, preprocessing, data transformation. Data Mining has been used in a variety of applications such as marketing, customer relationship management, engineering, and medicine analysis, expert prediction, web mining and mobile and mobile computing.

In health care institutions leak the appropriate information systems to produce reliable reports with respect to other information in purely financial and volume related statements. Data mining tools to answer the question that traditionally was a time consuming and too complex to resolve. They prepare databases for finding predictive information. Data mining tasks are Association Rule, Patterns, Classification and Prediction, Clustering. Most common modeling objectives are classification and prediction. The reason that attracted a great deal of attention in information technology for the discovery of useful information from large collections is due to the perception that we are data rich but information poor.Some the sample data mining applications are:

- Developing models to detect fraudulent phone or credit-card activity
- Predicting good and poor sales prospectus
- Predicting whether a heart attack is likely to recur among those with cardiac disease.
- Identifying factors that lead to defects in a manufacturing process.

Expanding the health coverage to as many people as possible, and providing financial assistance to help those with lower incomes purchase coverage [2]. Eliminating current health disparities would decrease the costs associated with the increased disease burden borne by certain population groups. Health administration or healthcare administration is the field relating to leadership, management, and administration of hospitals, hospital networks, and health care systems[1,3]. In the Healthcare sector Government spends more money.

- Proposal in draft NHP 2001 is timely that State¬ health expenditures be raised to 7% by 2015 and to 8% of State budgets thereafter [21].
- Health spending in India at 6% of GDP is among¬ the highest levels estimated for developing countries.
- Public spending on health in India has itself¬ declined after liberalization from 1.3% of GDP in

1990 to 0.9% in 1999. Central budget allocations for health have stagnated at 1.3%ofthe total Central budget. In the States it has declined from7.0% to 5.5% of the State health budget.

This paper mainly classifies and predict the set of medicated data using data mining techniques. The comparative study compares the accuracy level predicted by data mining applications in healthcare. Security and privacy is the key to be keep in mind, to provide security of data we have used AES Encryption for user's data and for more storage we have used cloud services. For any emergency issues a notification message is send to the neighbor using their emails.

II. PROPOSED SYSTEM

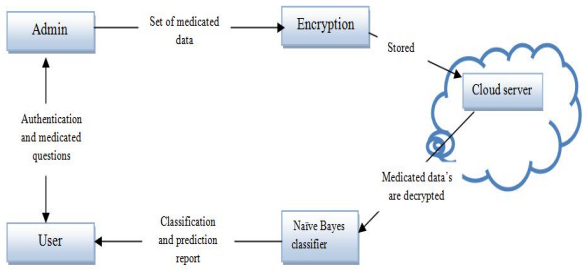


Fig 1.1 System Architecture

Data mining techniques are used for prediction and classification for predicting the set of medicated data's that are posted by the user's. SMTP(Simple Mail Transfer Protocol) used for any emergency issues. (i.e)For any critical cases the report that generated is send to the neighbor as a notification. AES(Advanced Encryption Standard) are used for security to protect clients data. The set of medicated data's are encrypted and stored in the cloud for privacy. Authentication is provided to user's for login. (See fig 1.1) for system architecture.

III. DATA MINING

Data mining is the non trivial process of identifying correct, paperback, likely useful, and ultimately understandable patterns in data. It is the process of classifying through huge data sets to identify patterns and inaugurate relationships to solve curious problems through analysis of data. The problem of effectively utilizing these massive volumes of data is becoming a major problem or all enterprises.

Definition: Data mining or knowledge discovery in database, as it is also known, is the non-trivial extraction of implicit, previously unknown and potentially useful information from the data. This holds a number of technical approaches, such as clustering, data summarization, classification, finding dependency networks, analyzing changes, and detecting anomalies (see fig 1.2) for data mining functions. The current assessment of data mining functions and products is the results of influence from many disciplines, including databases, information retrieval, statistics, algorithms, and machine learning.

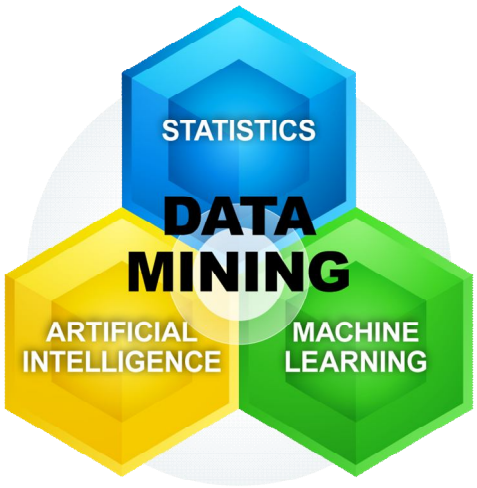


Fig 1.2 Data Mining Functions

History of Data Base and Data Mining

The data mining system has been started at the year of 1960s and earlier. In this, the data mining is simply on file processing. The next stage its Database management Systems to be started at the year of 1970s early to 1980s. In this OLTP, Data modeling tools and Query processing are worked. Database management system contains three broad categories to be worked. First one is Advanced Database Systems, this evaluated year of Mid-1980s to present in this Data models and Application oriented process are worked. The Second phase is Data Mining and Warehousing worked from the year of the late 1980s to present. The third phase is Web based Database Systems this started from 1990s to present and in Web mining and XML based database systems are included. These three broad categories are joined and create the new process that's called New generation of the Integrated Information system is started in 2000.

Data Mining Application Areas

Data mining is driven in part by new applications which require new capabilities that are not currently being supplied by today's technology. These up to the minute applications can be spontaneously into two broad categories.

- Business and E-Commerce

- Scientific, Engineering and Health Care Data

Data Mining Tasks

Data mining tasks are principally classified into two broad categories: (see fig 1.3) for types of data mining tasks.

- Predictive model
- Descriptive model

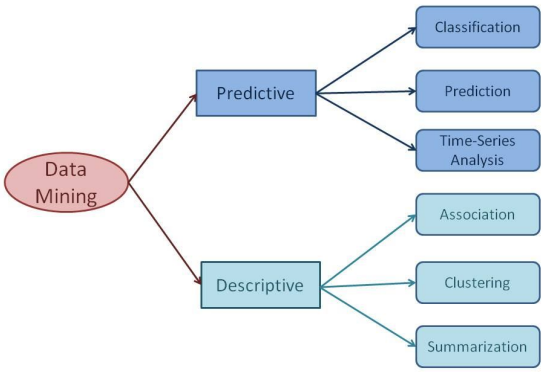


Fig 1.3 Types of data mining tasks

IV. CLOUD COMPUTING

Cloud computing is known as *on demand delivery of IT resources* through the internet with pay up costing. It bestow a solution of IT infrastructure in low cost. Actually, Small as well as some huge IT companies walk behind the customary methods to bestow the IT infrastructure. It means that **for any IT company, we need a Server Room that is the basic need of IT companies**. In server room, there should be a *database server*, *mailserver*, *networking*, *firewalls*, *routers*, *modem*, *swit ches*, *QPS* (Query Per Second) , configurable system, *high net speed* and the *maintenance engineers*.(see fig 1.4 for cloud computing architecture) To inaugurate such IT infrastructure, we need to disburse lots of money. To overcome all these mix-up and to diminish the IT infrastructure cost, Cloud Computing comes into extant.

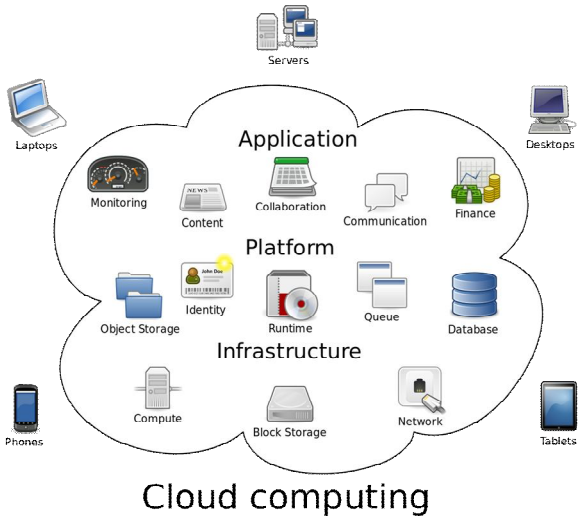


Fig 1.4 Cloud Computing Architecture

Types of cloud:

Public cloud:

A public cloud is based on the cloud services that are provided by third-party via the public internet, in which a service provider makes resources such as virtual machines, storage via the internet(see fig 1.5 for public cloud). Public cloud services may be cost free or offered on a pay-up model. The main sake of using a public cloud service are:

- For small and medium size of organization in general, Scalability is pay up.
- It enables scalability to meet workload and user demands;
- There are many wasted resources because users only pay for the resources they use.

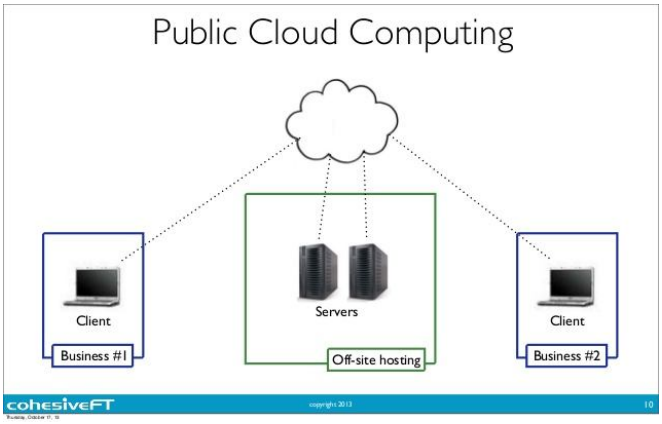


Fig 1.5 Public cloud computing

Private cloud:

A private cloud is the term of cloud computing service that provided through the internet or network of private to use only to the selected users as a replacement of general public(see fig 1.6 for private cloud computing). It is a unique model of cloud computing service that involves a secure cloud based service in which only the framed client can operate.

The main benefits and advantage of using private cloud:

- Costs
- Efficiency and control
- Customization
- Security and privacy
- Compliance
- Ensuring business continuity
- Geographic availability

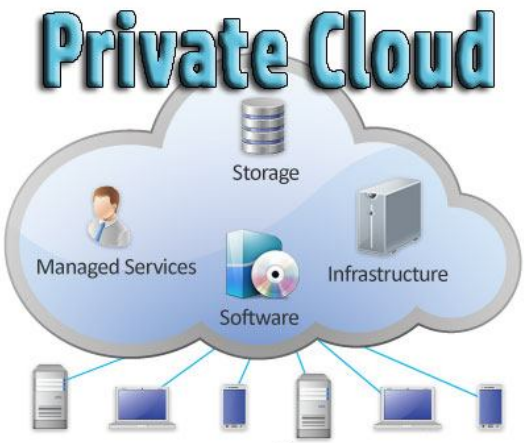


Fig 1.6 Private cloud computing

Hybrid cloud:

Hybrid cloud is a cloud computing environment which uses a mix of private cloud service and third-party, public cloud services with set up between the two platforms(see fig 1.7 for hybrid cloud computing). By allowing workloads between private and public clouds as computing needs and costs change, this service will gives businesses greater pliability and more data organizations.
The main benefits of hybrid cloud:

- Business continuity
- More opportunity for innovation
- Scalability
- Increased speed in market
- Risk management
- Secure systems
- Improved connectivity

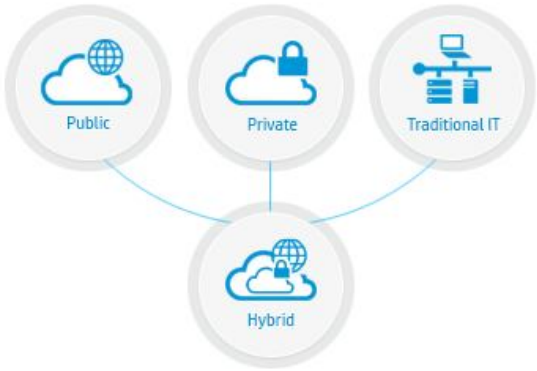


Fig 1.7 Hybrid cloud computing

V. NAÏVE BAYES CLASSIFIER

Naive Bayes classifies that‟ Rule is the major part for many machine-learning and data mining techniques. The rule is used to create models with predictive competence. It provides new ways of examine and understanding data. It learns from the “attestation” by calculating the tie-up between the target and other variables. Neural Networks consists of three layers: input, hidden and output units (variables). Connection between input units and hidden and output units are based on relevance of the assigned value (weight) of that particular input unit. The higher the weight the more important it is. Neural Network algorithms use Linear and sigmoid transfer functions. Neural Networks are suitable for training large amounts of data with few inputs. It is used when other techniques are unsatisfactory.

A. Analyzing the Data Set

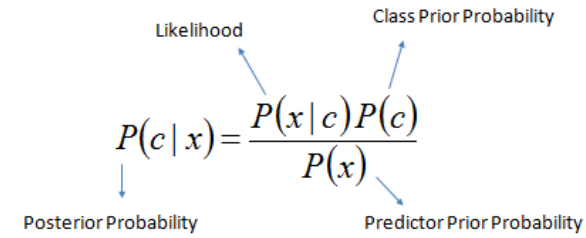
A dataset is a collection of data, normally represented in tabular form. Each column contains a particular variable. Each row represents a given member of the data set in question. It displays values for each of the variables, such as height and weight of an object or values of random numbers. Each value is called as a datum. The data set may contain data for one or more variables, corresponding to the number of rows.

The values are represented as numbers, such as real or integers, for example representing the height of a person in centimeters, but may also be ostensible data, for example representing a person's incite. Universally, values may be of any of the kinds distinguish as a level of measurement. For each variable, the values will be of the same kind. However, there may also be "missing values", which need to be indicated in some way.

A bunch of 500 records with 15 attributes as medical (factors) were obtained from the Heart Disease database lists. These records were split equally into two datasets: training (455 records) and testing dataset (454 records). To avoid prejudice, the records for each set were selected arbitarily. The attribute "Diagnosis" was identified as the predictable attribute with value "1" that represent the patients with heart disease and value "0" that represent the patients with no heart disease. The rest are input attributes. It is presumed that problems such as missing, inconsistent and duplicate data are resolved. Here in our project we get a data set from (.dat) file as our file reader program will get the data from them for the input of Naïve Bayes mining process.

B. Naïve Bayes Theorem

Bayes theorem provides a way of calculating the posterior probability, $P(c|x)$, from $P(c)$, $P(x)$, and $P(x|c)$(see fig 1.8 for Naïve Bayes Theorem). Naive Bayes classifier assume that the effect of the value of a predictor (x) on a given class (c) is independent of the values of other predictors. This suspiction is called class conditional independence.



$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Fig 1.8 Naïve Bayes Theorem

- $P(c|x)$ is the posterior probability of *class*.
- $P(c)$ is the prior probability of *class*.
- $P(x|c)$ is the likelihood, it will detect the probability of *predictor* given *class*.
- $P(x)$ is the prior probability of *predictor*.

VI. AES ENCRYPTION

The most well liked and universally espouse symmetric encryption algorithm likely to come across these days is the Advanced Encryption Standard (AES). It is originated that least six time faster than triple DES.

A substitution for DES was needed as its key size was too small. Triple DES was founded to overcome this drawback but it was slow.

The qualities of AES are as follows –

- Symmetric key block cipher
- 128-bit data, 128/192/256-bit keys
- Stronger and faster than Triple-DES
- Provide full specification and design details
- Software implementable in C and Java

C. Operation of AES

AES is an repetitive rather than Feistel cipher. It is based on 'substitution–permutation model'. It encompass a series of linked operations, some of which involve replacing inputs by specific outputs and others involve shuffling bits.

Interestingly, AES works all its operations on bytes rather than bits. Hence, AES will represent the 128 bits of a plaintext block as 16 bytes. These 16 bytes are arranged in four columns and four rows for processing as a matrix. AES works on 10 rounds for 128-bit keys, 12 rounds for 192-bit keys and 14 rounds for 256-bit keys. Each of these rounds uses a different 128-bit round key, which is computed from the original AES key.

See fig 1.9 for the overall AES structure is given in the following illustration–

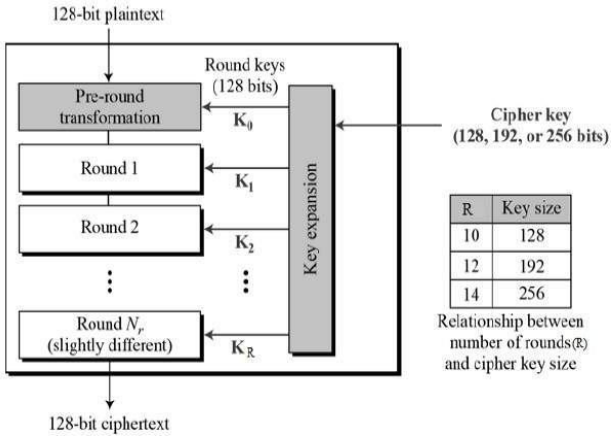


Fig 1.9 Overall AES Structure

D. Encryption Process

In Encryption, we limited to description of a typical round of AES encryption(see fig 1.10 for encryption process). Each one of the round encompasses of four sub-processes. The first phase round process is pictured below –

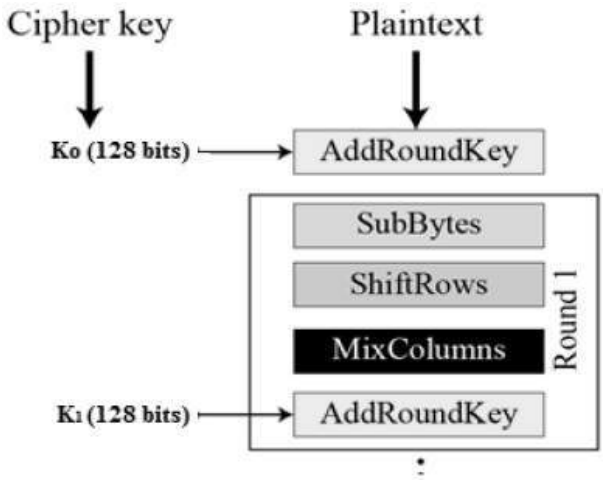


Fig 1.10 Encryption Process

E. *Byte Substitution (Sub-Bytes)*

The 16 bit input is substituted based on the fixed table given in design. The result is represented as four cross four matrix.

F. *Shift rows*

Rows of matrix are shifted to left. Incase of new rows that "fall off" are re-inserted on the right side of row.
Shift operations are as follows –

- First row is not shifted.
- Second row is shifted one position to the left.
- Third row is shifted two positions to the left.
- Fourth row is shifted three positions to the left.

The result of shift rows is a new matrix consisting of the same 16 bytes but shifted with respect to each other.

G. *Mix Columns*

Each column of four bytes is now computed using a special function based on mathematics . This mathematical function takes input from the four bytes of one column and outputs the new four bytes, instead of the original column. The output that generated is another new matrix contains 16 new bytes. It is noted that this step is not performed in the last round.

H. *Add round key*

The new 16 bytes of the matrix are now represented as 128 bits and they are XOR-ed to the 128 bits of the round key. If it is considered as last round then the output is the encrypted text. Otherwise, the resulting 128 bits are computed as 16 bytes and we recursively compute another similar round.

I. *Decryption Process*

` The decryption process in AES, it is similar to the encryption process in vice-versa. Each round contains four process that is illustrated in the reverse order –

- Add round key
- Mix columns
- Shift rows
- Byte substitution

Since all the sub-processes in each round are vice-versa, the encryption and decryption algorithms are computed separately , although they are related very closely.

VII. CONCLUSION

Nowadays, space and time are no longer a stumbling stone for modern healthcare by using more powerful analysis technologies. Medical diagnosis is evolving to patient-centric prevention, prediction and treatment. The cloud computing technology have been developed gradually and will be used everywhere. Consequently, healthcare will also enter the cloud computing era. This paper aimed to provide classified and predicted data for user to know their health status. The prediction of diseases using Data Mining techniques is a challenging task but it exponentially reduces the human effort and increases the diagnostic accuracy. Developing an effective data mining techniques for an application will reduce the cost and time restraint in terms of human resources and competence. Traversing knowledge from the medical dataset is such a risk task as the data found are noisy, irrelevant and massive too. In this case, data mining technique come in handy in exploring of knowledge of the medical dataset and it is quite interesting. This project has presented a smart healthcare system assisted by cloud computing and data mining techniques, which includes 1) To study about various data mining application in the healthcare sector 2) To provide health related information to users 3) To reduce the human effects and cost effective one. In the future, we will focus on developing various applications based on the Health care sector using cloud computing and big data to provide a better environment to humans.

REFERENCES

[1] J. Wan et al., "Cloud-enabled wireless body area networks for pervasivehealthcare," IEEE Netw., vol. 27, no. 5, pp. 56–61, Sep./Oct. 2013.

[2] G. Kim, S. Trimi, and J. Chung, "Big-data applications in the government
sector," ACM Commun., vol. 57, no. 3, pp. 78–85, 2014.
[3] K. Lee, T. T. Wan, and H. Kwon, "The relationship between healthcareinformation system and cost in hospital," Pers. Ubiquitous Comput.,vol. 17, no. 7, pp. 1395–1400, Oct. 2013.
[4] M. Chen, Y. Ma, S. Ullah, W. Cai, and E. Song, "ROCHAS: Robotics andcloud-assisted healthcare system for empty nester," in Proc. BodyNets,
Boston, MA, USA, Sep. 30–Oct. 2 2013, pp. 1–4.
[5] M. Chen, "NDNC-BAN: Supporting rich media healthcare services vianamed data networking in cloud-assisted wireless body area networks,"
Inf. Sci., vol. 284, pp. 142–156, Nov. 2014.
[6] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," Mobile Netw. Appl.,vol. 19, no. 2, pp. 171–209, Apr. 2014.
[7] M. Chen, Y. Hao, Y. Li, D. Wu, and D. Huang, "Demo: LIVES: Learn-ing through interactive video and emotion-aware system," in Proc. ACM
Mobihoc, Hangzhou, China, Jun. 22–25, 2015.
[8] H. Zhang, S. Mehotra, D. Liebovitz, C. Gunter, and B. Malin, "Mining de-viations from patient care pathways via electronic medical record system
audits," ACM TMIS, vol. 4, no. 4, p. 17, 2013.
[9] J. Wan, H. Yan, H. Suo, and F. Li, "Advances in cyber-physical systemsresearch," KSII TIIS, vol. 5, no. 11, pp. 1891–1908, 2011.
[10] J. Wan, D. Zhang, S. Zhao, L. T. Yang, and J. Lloret, "Context-awarevehicular cyber-physical systems with cloud supportArchitecture, chal-
lenges, and solutions," IEEE Commun. Mag., vol. 52, no. 8, pp. 106–113,Aug. 2014.
[11] M. Chen, Y. Hao, Y. Li, C. Lai, and D. Wu, "On the computation offload-ing at ad hoc cloudlet: Architecture and service models," IEEE Commun.
Mag., vol. 53, no. 3, pp. 1–7, Jun. 2015.
[12] C. Lin et al., "Temporal event tracing on big healthcare data analytics," inProc. IEEE Int. BigData Congress, 2014, pp. 281–287.
[13] L. Nie, T. Li, M. Akbari, J. Shen, and T. Chua, "Wenzher: Comprehensivevertical search for healthcare domain," in Proc. 37th Int. ACM SIGIR
Conf. Res. Develop. Inf. Retrieval, 2014, pp. 1245–1246.
[14] H. Takeuchi and N. Kodama, "Validity of association rules extracted byhealthcare-data-mining," in Proc. IEEE 36th Annu. Int. Conf. EMBC,
2014, pp. 4960–4963.
[15] V. Chandola, S. Sukumar, and J. Schryver, "Knowledge discovery frommassive healthcare claims data," in Proc. 19th ACM SIGKDD Int. Conf.
Knowl. Discovery Data Mining, 2013, pp. 1312–1320.
[16] M. Chen, S. Gonzalez, A. Vasilakos, H. Cao, and V. Leung, "Body areanetworks: A survey," ACM/Springer Mobile Netw. Appl., vol. 16, no. 2,pp. 171–193, Apr. 2011
[17] M. Chen, S. Gonzalez, Q. Zhang, M. Li, and V. Leung, "A 2G-RFIDbased E-healthcare system," IEEE Wireless Commun. Mag., vol. 17, no. 1,pp. 37–43, Feb. 2010
[18] M. Chen, Y. Zhang, Y. Li, S. Mao, and V. Leung, "EMC: Emotion-awaremobile cloud computing in 5G," IEEE Netw., vol. 29, no. 2, pp. 32–38,Mar. 2015.
[19] M. G. R. Alam, E. J. Cho, E. Huh, and C. S. Hong, "Cloud based mentalstate monitoring system for suicide risk reconnaissance using wearable
bio-sensors," in Proc. 8th Int. Conf. Ubiquitous InfManage. Commun.,2014, p. 56.
[20] C. He, X. Fan, and Y. Li, "Toward ubiquitous healthcare services with anovel efficient cloud platform," IEEE Trans. BiomedEng.,vol. 60, no. 1,pp. 30–234, Jan. 2013.
[21] O. Samuel, M. Omisore, B. Ojokoh, and E. Atajeromavwo, "Enhancedcloud based model for healthcare delivery organizations in developing
countries," Int. J. Comput. Appl., vol. 74, pp. 0975–8887, 2013.
[22] C. Yang et al., "Accessing medical image file with co-allocation HDFS incloud," Future Gener. Comput. Syst., vol. 43, pp. 61–73, Feb. 2015.
[23] M. Chen, Y. Zhang, Y. Li, M. Hassan, and A. Alamri, "AIWAC: Affectiveinteraction through wearable computing and cloud technology," IEEEWireless Commun. Mag., vol. 22, no. 1, pp. 20–27, Feb. 2015.
[24] M. Chen, Y. Zhang, L. Hu, T. Taleb, and Z. Sheng, "Cloud-based wirelessnetwork: Virtualized, reconfigurable, smart wireless network to enable 5G
technologies," ACM/Springer Mobile Netw. Appl., DOI: 10.1007/s11036-015-0590-7, 2015, to be published