International Journal of Trend in Research and Development, Volume 3(2), ISSN: 2394-9333
www.ijtrd.com

Using Medical Big-Data Platforms to Creating Stochatic Preference for Adaptive Crowd Sourcing

[1]Poornima Devi.L and [2]SathyaBalaji.N,
[1]M.E.Scholar, [2]Associate Professor,
[1,2]Department of Computer Science & Engineering,
J.K.K.Munirajah College of Technology, T.N.Palayam, Erode, TamilNadu, India

Abstract: This project proposes to store a huge value of data in medical platform. In this medical storage system use 60GHz wireless technologies for in hospital wireless network access. Crowdsourcing used for big data platforms, in which data are gathered by user participation. The datas stored on a centralized server via access point. The technology, which are used to process the data parallely on different cpu node. This technology provide storage for billions and trillions of unstructured data. While sending input data to centralized server via wireless access, the power allocation method make decision to allow the data to server and perform schedule operation when datas from multiple cpu node. In this project it storess amount of unstructured data which are stored on centralized server and the stored data be accessed secure with high availability.

Keywords: 60ghz, Crowd Sourcing, Centralized Server

I. INTRODUCTION

In recent years, the volume of medical data generated by large hospitals is becoming increasingly large due to technological advancements in medical devices.It includes high-resolution magnetic resonance imaging (MRI), motion MRI, ultrasound, and digital microscopy [1]. Furthermore, centralized storage of medical records is a common practice for sharing medical data among medical practitioners, as illustrated in Fig. 1 . Oftentimes, medical records are uploaded to the centralized medical record using modern movable equipments. Because of the sensitive nature of medical data, data aggregation, as shown in needs to be privacy preserving.Therefore, interconnecting medical storage platforms with external networks (such as the Internet) is not recommended.

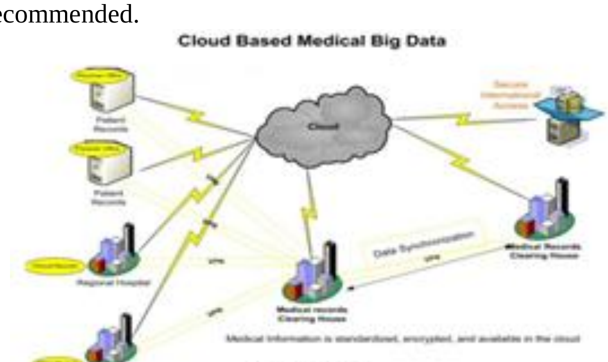


Figure 1 Medical Information Storage in Centralized server

Medical data in the proposed medical storage platform is often gathered and organized by using fixed users purposed by medical tablets, smartphones, computed tomography scanners, with the principle of crowdsourcing.In this proposed medical storage system, we consider using 60-GHz wireless technologies for in-hospital wireless network access.The choice of wireless technologies has been widely advocated and accepted in the literature because of high data rates achieved by ultrawide bandwidth. One of the many interesting design problems that need to be addressed to construct the centralized privacy preserving data.

II. DEFINING BIG DATA

Big data typically refers to the following types of data:

- Traditional enterprise data – includes customer information from CRM systems, transactional ERP data, web store transactions, general ledger data.
- Machine-generated /sensor data – includes Call Detail Records ("CDR"), weblogs, smart meters, manufacturing sensors, equipment logs (often referred to as digital exhaust).
- Social data – includes customer feedback streams, micro-blogging sites like Twitter, social media platforms.

In fact, there are three key characteristics that define big data:

- Volume is the amount of data generated by organizations or individuals.
- Velocity is the frequency and speed at which data is generated, captured and also shared. Consumers as well as businesses now generate more data and in much shorter cycles, from hours, minutes, seconds down to milliseconds.
- Variety is the access of new data types including those from social, machine and mobile sources. New types include content, location, hardware data points, log data, machine data, metrics, mobile, physical data points, process, radio frequency identification (RFID), search, sentiment, streaming data, social, text and web. They have variety includes traditional unstructured clinical data (i.e., free text)

Max-Weight Scheduling Algorithm

In this section we evaluate the Max-Weight scheduling policy, for the respect traffic flows while transaction. Informally speaking, the "weight" of a perfect schedule is the sum of the lengths of all queues included in it. The Max-Weight policy activates a feasible schedule with the maximum weight at any given time slot. More formally, under the Max-Weight policy, the scheduling vector S(t) belongs to the set:

$$\max : \sum_{aj \in A} \sum_{mi \in M} W(i,j)[t] \bullet x(i,j)$$

$$\text{subject to } \sum_{aj \in A} x(i,j) \leq 1, \forall mi \in M$$

$$\sum_{mi \in M} x(i,j) \leq NRFj , \forall aj \in A$$

$$\sum_{mi \in M} r(i,j) \bullet x(i,j) \leq Rj , \forall aj \in A$$

If this set includes multiple feasible schedules, then one of them is selected uniformly at random. This result is well-known.

III. BIG DATA IN HEALTH CARE

The types of data anticipated to be of use in BDA include:

1. Clinical data – up to 80 per cent of health data is unstructured as documents, images, clinical or prescribed notes;
2. Publications – clinical research and the medical reference material;
3. Clinical references – text-based practice plan and health product (e.g., drug information) data;
4. Genomic data – represents major amounts of new genetic material sequencing data;
5. Streamed data – home monitoring, tele health, Hand held and sensor-based wireless or smart devices are new data sources and their types;
6. snare and social networking data – consumer use of Internet – data from search engines and social networking sites;
7. Business, organizational and external data – clerical data such as billing and preparation and other non-health data.

BDA can mine volumes of medical literature and other unstructured data and integrate these results with the increasing volumes of distinct data captured in EHRs, EMRs and PHRs.

BDA can combine content analysis, evidence-based data and through natural language processing technology can understand, learn and then predict future the events. These analytics are then fed back to clinicians as considerations in their executive.

IV. BIG DATA CHALLENGES IN HEALTH CARE

- Leveraging the patient or the data correlations in longitudinal records.
- Understanding amorphous clinical notes in the right context.
- Competently handling large volumes of medical imaging data that extracting potentially useful information and biomarkers.
- Analyzing genomic data is a computationally intensive task that combining with standard clinical data and adds additional layers of complexity.
- Capturing the patient's behavioral data through several sensors; their various social interactions and communications in the BDA to get answers for their own conditions. Data could be presented back in a momentous way and persuade patient participation in their health care Plans and potentially reduce re-admissions or unfavourable outcomes.

V. SECURITY ISSUES

In cloud-based HIS, security should be the top precedence from day one. Patients' data should be protected with widespread physical security, data encryption, user authentication, and application security as well as the latest standard-setting security practices and certifications, and secure point-to-point data replication for data backup. These security issues have been broadly investigated for cloud computing in general. A major challenge to healthcare cloud is the security threats including cloud, loss of isolation of patient's information, and the unauthorized use of this information.

Hence, a number of security requirements should be satisfied by healthcare cloud computing systems. The main security and privacy requirements for healthcare clouds that are discussed below [3,4]:

- Authentication: in a healthcare cloud, both healthcare information offered by CSPs and identities of users (HPs, practitioners, and patients).

It should be established at the entry of every access using client names and passwords assigned to users by CSPs.

- Authorization: is an essential security obligation that is used to control access priorities, permissions and resource ownerships of the users on the cloud. Each cloud user is granted privileges based on his account.

The Patient that can allow or refuse to sharing their information with other healthcare practitioners or CDOs.

To implement patient permission in a healthcare system, patient may grant rights to users on the basis of a role or attributes held by the respective user.

- Non-repudiation: implies that one party of a transaction cannot deny having received a transaction nor the other party deny having sent a transaction.

In a healthcare system,technologies such as digital signatures, timestamps, confirmation receipt, and encryption can be used to create authenticity and non-repudiation for patients, CDOs, and practitioners.

- Integrity and Confidentiality: integrity means preserving the accuracy and consistency of data. In the healthcare system, it refers to the fact that EHRs have not been tampered by unauthorized use. Confidentiality is defined by the International Organization for Standardization (ISO) in ISO-17799 as "ensuring that information is accessible only to those authorized to have access".The confidentiality and integrity can be achieved by access control and encryption technology in EHR systems.
- Availability: For any EHR system to serve its purpose, the information must be available when it is needed. High availability systems aim to linger available at all times, preventing service disruptions due to power outages, hardware failures, and system upgrades. Ensuring the availability also involves preventing denial-of-service (DoS) attacks.

In conclusion, here is a brief example of how the transition from relational databases to big data is happening in the existent world. We, at Health Catalyst, are working with one of our large health system clients and Microsoft to create a massively parallel data warehouse in a Microsoft APS application that also includes a Hortonworks Hadoop Cluster. This means we can run a traditional relational database and a big data cluster in parallel. We can query both data stores simultaneously, which significantly improves our data processing power. Simultaneously, we are beginning to experiment with big data in important ways, such as performing natural language processing (NLP) with surgeon notes, predictive analytics, and other use cases.The progression from today's symmetric multiprocessing (SMP) relational databases to massively parallel processing (MPP) databases to big data in healthcare is underway.

VI. PERFORMANCE EVALUATION

For simulation geometry setting, we suppose that there are randomly deployed 1200 APs [22]. The channel states between AP $a_j \in A$ and MU $m_i \in M$, i.e., $h_{i,j}[t]$, are assumed to be additive white Gaussian noise. Moreover, NRF j= 1, ∀aj ∈ A for the simplicity of the simulations. In addition, the GTx i and GRx j are assumed to be 22 dBi . The possible video sources are presented in it. Each scheduled MU selects one medical imaging format among the given nine uniformly at random. In addition, the related frame rate (frames/s) is set to 60 frames/s. For the path-loss, a hospital specific model is considered. According to the randomness of the positions of APs and MUs, Monte Carlo simulations are performed with 100 iterations. In these criteria, the performance of the proposed algorithm is evaluated in terms of the scheduling performance and the buffering performance.

A. Simulation Result 1: Scheduling Performance

To verify the performance of the proposed max-weight scheduler,the corresponding simulations are performed with various numbers of MUs (from 1 to 5000). In addition, the proposed max-weight scheduler is compared with:

- SRM scheduler [i.e., W(i,j)[t] _ r(i,j) in (11)];
- Random scheduler [i.e., W(i,j)[t] is chosen uniformly at random in the range of [0, 1] as in.(from 1 to 5000) and y-axis shows the relative performance, which is denoted by α = (αCompare/αMax-Weight) where αMax-Weight isthe achievable rate with the proposed max-weight scheduling and αCompare is the achievable rate for the compared scheduler.

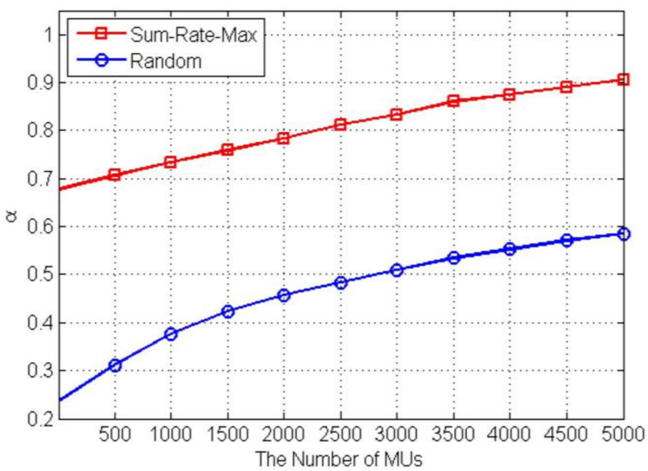
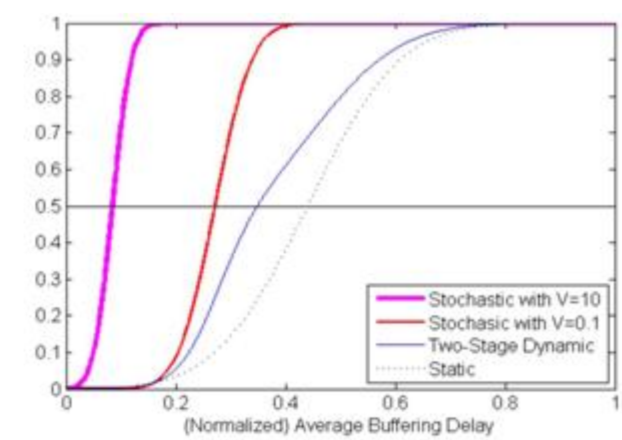


Fig. 2. Scheduling performance

B. Simulation Result 1: Buffering Performance



To verify the recital of the proposed stochastic buffering, the corresponding simulations are performed with various V values, where V ∈ {0.1, 10}. two-stage dynamic buffering (allocates PTxmin if the current buffer occupancy is less than half of the buffer size; otherwise it allocates PTxmax). The static power allocation that allocates the transmit power with PTxi [t] = 1/2(PTxmin+PTxmax).

CONCLUSION

This paper proposes a framework for secure Health Information Systems (HISs) based on big data analytics in mobile cloud computing environment. The framework provides a high level of integration, interoperability, and sharing of EHRs among healthcare providers, patients and practitioners. The proposed framework applies a set of security constraints and access control that secured integrity, confidentiality, and privacy of medical data. The eventual goal of the proposed framework is to introduce a new generation of HISs that are able to provide healthcare services of high quality and low cost to the patients using this combination of big data analytics, cloud computing and the mobile computing technologies. In the future we plan to design and implement HIS based on the proposed framework.

References

[1] Rasid Z. Big data infographic: What is big data? [Internet] 2013. [Cited March 17, 2014]. Available from: http://www.asigra.com/blog/big-data-infographic-what-big-data
[2] Dembosky A: "Data Prescription for Better Healthcare." Financial Times, December 12, 2012.
[3] R. Zhang and L. Liu, "Security Models and Requitements for Healthcare Application Clouds", IEEE 3rd International Conference on Cloud Computing, 2010.
[4] A. Youssef and M. Alageel, "Security Issues in Cloud Computing", the GSTF International Journal on Computing , Vol.1 No. 3, 2011.
[5] Ahmed E. Youssef and Manal Alageel, "A Framework for Secure Cloud Computing", International Journal of Computer Science Issues (IJCSI), Vol. 9, Issue 4, No 3, pp. 478-500, July 2012.
[6] J. Kim, Y. Tian, S. Mangold, and A. F. Molisch, "Joint scalable coding and routing for 60 GHz real-time live HD video streaming applications," IEEE Trans. Broadcast., vol. 59, no. 3, pp. 500–512, Sep. 2013.
[7] G. Chatzimilioudis, A. Konstantinidis, C. Laoudias, and D. Zeinalipour-Yazti, "Crowdsourcing with smartphones," IEEE Internet Comput., vol. 16, no. 5, pp. 36–44, Sep./Oct. 2012.
[8] J. Kim, Y. Tian, S. Mangold, and A. F. Molisch, "Quality-aware coding and relaying for 60 GHz real-time wireless video broadcasting," in Proc. IEEE Int. Conf. Commun. (ICC), Budapest, Hungary, 2013, pp. 5148–5152.
[9] M. Allahbakhsh et al., "Quality control in crowdsourcing systems: Issues and directions," IEEE Internet Comput., vol. 17, no. 2, pp. 76–81,Mar./Apr. 2013.
[10] A. Youssef and M. Alageel, "Security Issues in Cloud Computing", the GSTF International Journal on Computing , Vol.1 No. 3, 2011.