International Journal of Emerging Engineering Research and Technology
Volume 3, Issue 1, January 2015, PP 63-69
ISSN 2349-4395 (Print) & ISSN 2349-4409 (Online)

Three Stratum Cloud Cadres for Concentric Segregation and Secured Access of Health Data

Aksha Mondal[1], Ashwini Shivaprasad[2,] Akanksha Singh[3]

Department of Computer Science and Engineering
SKN Sinhgad Institute of Technology and Science, Savitribai Phule Pune University, Pune, India
[1]aksha.m1993@gmail.com, [2]ashwini3101993@gmail.com, [3]akanksha810@gmail.com

Abstract: *The blistering development of cloud technology bears the users in distinctive fields with its lofty impact on the corollary they expect in a way that users can access and work anywhere, anytime. This astounding concept has incentivized us to propound a model for accessing data in secured mobile health care system. The information is segregated according to the result among the aplenty of stored relevant data, drifted to extract the better outcome of results and it is made available through the three stratum cloud cadre (TSCC) model. TSCC is an assimilation of public, private and hybrid cloud integrating facet medicinal value model, infer from a plenary assessment of medicines that is most reviewed respect to health problem. TSCC model brings the best corollary that people can approach for taking drugs and precautions through health pocket hospital remedies.*

Keywords: *Three stratum cloud cadre, segregation, secured mobile health care, hybrid cloud, cloud technology, facet medicinal value model.*

1. INTRODUCTION

In this fast growing technical era people are getting dependent more on web technologies like internet, preferring technologies that are efficient and time saving. Many user-centred platforms are now available for information sharing and user interaction, such as Epinion, Amazon, Facebook and Twitter. Nowadays when people are interested in a product or a service, they usually not only look for official information from product manufacturers or service providers, but from experienced and practical opinions of the customers' and users' points of view are also efficacious. As a result, more sophisticated aspect level opinion mining approaches have been proposed to extract and group aspects of a product or service and predict their sentiments or ratings [20], [21]–[23]. Recent state-of-the-art approaches such as frequency-based approach [22], relation-based approach [24], [25], supervised learning [26] and topic modelling [21], [27] showed that favourable results could be obtained. Online reviews, blogs and forums dedicated for different kinds of products are pervasive, and how to effectively analyse and exploit such immense online information source is a challenge. Fast access to medicinal review and prescribed health data enables better healthcare service provisioning, improves quality of life, and helps saving life by assisting timely treatment in medical emergencies. Anywhere-anytime-accessible electronic healthcare systems play a vital role in our daily life.

Unlike general products or services, drugs have a very limited number of kinds of aspects: price, ease of use, dosages, effectiveness, side effects and people's experiences. There are other more technical aspects such as chemical or molecular aspects, but they are almost not mentioned in drug reviews. A difficulty in dealing with drug reviews is that the wording in describing effectiveness, side effects and people's experiences are very diverse. In particular, a set of side effect symptoms for a drug is very unlikely applicable to another drug. This impedes some opinion mining approaches. More importantly, authors sometimes do not indicate which aspects they are describing, they just give descriptions of symptoms, feelings and comments.

Nevertheless, recent studies have shown that patient generated contents are useful and important [15]–[18], especially for chronic diseases and drugs with afflicting side effects. Many patients hope to get more information from other patients with similar conditions. They can also share their experience and propose practical ways to alleviate symptoms and side effects of drugs. These online communities were found to have positive impacts on patient health [19]–[21]. Apart from this there is a challenging issue of patients' prescriptions and report. Many times it happened that due to less information of

patients previous medical history doctors fails to understand and detect their ail. It generally happens due to patients' negligence in keeping their old medical reports. Observing all this issues we have planned to design an online system which will store the personal medical data and case history and will also provide consultant for health ail.

There are many models designed for data storage and also many companies provide this facility, but stakeholders are not able to trust them completely because of the privacy issues [8]. Also privacy for health data is becoming one of the important and concern part. A person health history plays an important role as it may responsible for his job, during education admission, for marriage proposal or for insurance [1] and many such related fields. If the health data is in public then it may also cause harassment in once life. Taking care of the importance of one's privacy we have proposed a more secured cloud structure combining three types of cloud i.e. private cloud, public cloud and hybrid cloud. Services supported by mobile devices, such as home care and remote monitoring, enable patients to retain their living style and cause minimal interruption to their daily activities.

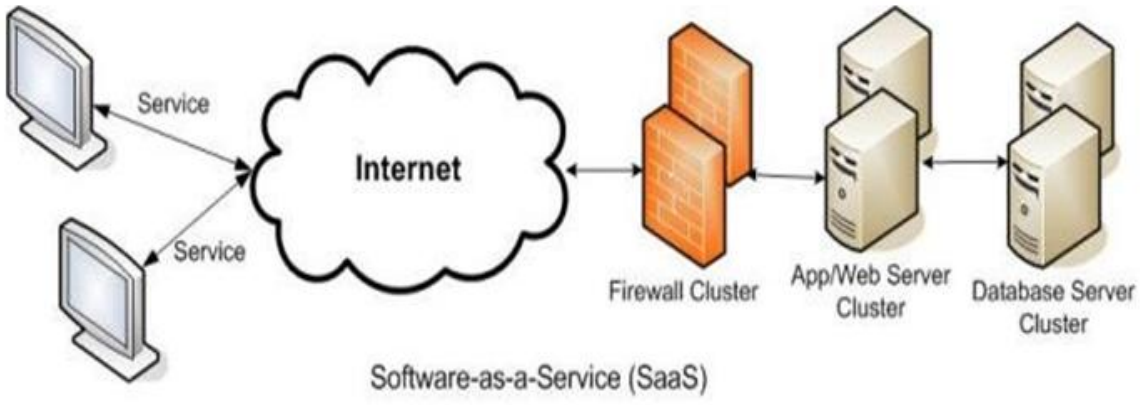

Fig1. *Software as a service*

Thus the system provides a secured health data storage facility with the medicinal review and has introduced an additional feature of medicinal drug review with health data storage [2]. This reviews and suggestions are checked and also reviewed by doctors and authorized people to maintain accuracy. The system is based on SAAS cloud structure fig.1 and is a mobile method which can be used anywhere anytime.

2. RELATED WORK

We note that there has been a great use of cloud technology for data storage. Related work has shown the use of private cloud and public cloud for storing medical data[1].

But its privacy concern is increasing day by day. To make it secure, many algorithms have been designed. Segregation of data is not done properly with the use of public and private cloud which somehow affects the security and also does not provide secured access control. The data on public cloud can be accessed by anyone while on the private cloud can only be accessed by the patients. It might result in leakage of information as very little information should be stored on public cloud in order to prevent information leakage. Moreover segregation of data is an important point which in turn influences the access control.

Keeping the security issues in mind, various security algorithms have been used. Identity based encryption(IBE) [12] is used [13] for simple cryptographic access control. Privacy preserving health data storage [15] let the patients encrypt their own data and stores it on remote server or third party server. Patient controlled encryption allow the data to be arranged in a hierarchy and provides symmetric keys. Searchable symmetric key(SSE) [15] algorithm was proposed by Goh and later improved by Curtmola [16]. It allows us to search through the encrypted data. Attribute-based encryption(ABE) was proposed by sahai and waters[9]. In ABE approach there are two classes Key-Policy ABE [10] and Cipher text-policy ABE [11]. The receiver after receiving the secret key and attributes of the sender can decrypt information if it has matching attributes.

There has been some algorithms designed for drug reviews which displays highly rated drugs [2] for particular diseases but this has not been used with cloud technology.

To ensure security, we base this study of security on ABE algorithm. Our proposed system aims to tackle these problems by introducing hybrid cloud for data segregation and providing access control with an additional feature of drug review using data mining.

3. PRELIMINARIES

3.1. Consonance Quest Encryption

Consonance quest encryption was edited and inspired by SSE which allows data owners to store encrypted documents on remote server, and simultaneously provides away to search over the encrypted documents. More importantly, neither the operation of outsourcing nor keyword searching would result in any information leakage to any party other than the data owner, thus achieving a sound guarantee of privacy.

Genkey: this function is used to generate keys. Parameters are taken as input and secret key K is given as the output

Indbuild (*D, K*): this function is used to build the indices, denoted by *I*, for a collection of document *D*. Secret key K is taken and D and I is the output. This enables searching for a document even if it is encrypted.

Trapdoor (*K,w*): This function is run to compute a trapdoor for a keyword *w*, enabling searching for this keyword. A trapdoor *Tw* can be considered as a proxy for *w* in order to hide the real meaning of *w.*. The function takes the secret key *K* and the keyword *w* and outputs the respective trapdoor *Tw.*

Search (*I, Tw)*: This function is executed by the remote server to search for documents containing the user defined keyword *w*. All documents in *D* are encrypted and stored in the remote servers. The index I consists of two data structures, namely an array A, for storing the nodes, and a look-up table T, for keeping information that enables the remote server to locate the elements in A. All nodes are encrypted with random generated keys (different from the keys for encrypting the document) and stored as entries in A "scrambled" in a random order. However to effectively organize the nodes, two measures are taken.

- All the nodes whose respective files containing the same key word *wi* are linked together in the linked list L*i*
- each node contains the index in A as well as the random generated encryption key of next node in L*i* .Obviously, with the information contained in the first node, one will be able to decrypt all the nodes in the same linked list L*i* , and, thus, access all the respective file identifiers of files containing keyword *wi* . However, because the first node in the linked list does not have a previous node, the first node's index in A and its decryption key are stored in the field value of an entry in T, which is defined as a map _address, value_. The field value is encrypted as it will be XOR-ed with an output of a pseudorandom permutation (PRP) function. The other field address is given by the output of a pseudorandom number generator to locate the first node. In other word, address serves as part of the trapdoor *Tw* to access the documents containing the respective keyword *w*. In fact, *Tw* consists of an output of a random number generator, for the purpose of locating entries in T, and an output of a PRP function, for the purpose of encrypting the entries, given the input *w* of pseudorandom algorithms. To set up SSE, the user runs IndBuild, which constructs A and T based on the documents *D* in clear texts in ways said above. To search document containing keyword *w*, the user run Search. Specifically, it uses Trapdoor to compute the respective trapdoor *Tw* and send the first part of *Tw* to the remote server. Upon receiving this information, the remote server uses it to locate and returns the respective encrypted entry in T. Then, the user uses the second part of *Tw* to decrypt the entry and get the information of the first node of the respective linked list. With that, the user can get all identifiers of wanted files, and, thus, retrieve and decrypt with the respective keys the encrypted files containing keyword *w*.

3.2. Attribute-Based Encryption

ABE has shown its promising future in fine-grained access control for outsourced sensitive data [31], [32], [33], [10], [34].

Typically, data are encrypted by the owner under a set of attributes. The parties accessing the data are assigned access structures by the owner and can decrypt the data only if the access structures match the data attributes.

4. PROPOSED ALGORITHM

The proposed work is an integrated model of [1] and [2] that provides easier and secured accessibility of health related data by authorised personnel. There are two things that needs to be considered in this model; security and concentric segregation of data. This model is based on attribute based encryption in which cipher text are labelled with a set of the patient's various attributes like country code, date of birth etc. Along with the attributes, the private key will be generated by using two languages that will correspond to the letters of the patient's attribute.

Since decryption consumes more resources than encryption, the process will be enhanced by using only a constant number of pairings to decrypt any cipher text. The data will be decrypted only if the cipher text attributes match with that of private key.

Setup (λ, U) $\rightarrow$ (P_k, M_k). Security parameter λ and a universe description U are taken as input. Public key P_k and master key M_k are obtained as the output.

Encrypt (Pk, M, A) $\rightarrow$ CT. The encryption algorithm takes as input the public parameters P_k, a message M and a set of attributes A and outputs a cipher text C associated with the attribute set.

LangChange (Pk, M, A): This algorithm takes public key P_k, message M and attribute A and converts the attribute to cipher text for private key generation using two local languages. Each letter of the attribute will be substituted by a letter taken from one of the two languages chosen earlier by the developer.

Substitution algorithm is as follows:

- Consider English alphabets. From some language L we will take 26 letters Z_1 $Z_2 \ldots Z_n$.
- For i=1 to 26

  O_i=Zi+7

  If O_i>26

  Then O_i=O_i-26

  Else Save O_i

KeyGen (M_k, S, O) $\rightarrow$ S_k.

The master secret key M_k, encrypted letter O_i and an access structure S are taken as input and outputs a private key S_k corresponding to the attributes.

Decrypt (S_k, C) $\rightarrow$ M.

The decryption algorithm takes private key S_k associated with access structure S and cipher text C associated with attribute set A and outputs message M to indicate if A satisfies S else an error message will be given. Thus the performance will depend on the number of attributes involved.

Now the next key point that is included is the way in which data is segregated to public, private and hybrid cloud. This can be done using any of the classifiers in data mining.

Given a set F of cases, an initial tree is built using divide and conquer algorithm:

- If all the cases in F belong to the same class or F is small, the tree is a leaf labelled with the most frequent class in F.
- Choose a test for each case in F with possible outcomes for each attributes in F. Partition F into sub sets F_1,$F_2 \ldots F_n$ and repeat the test cases for each of the subsets.

Attributes can be either numeric or nominal and this determines the format of the test outcomes. For a numeric attribute *A* they are $\{A \leq h, A > h\}$ where the threshold *h* is found by sorting *S* on the values of *A* and choosing the split between successive values that maximizes the criterion above.

5. SYSTEM AND THREAT MODELS

5.1. System Model

The main entities involved in our system are depicted in fig 2.

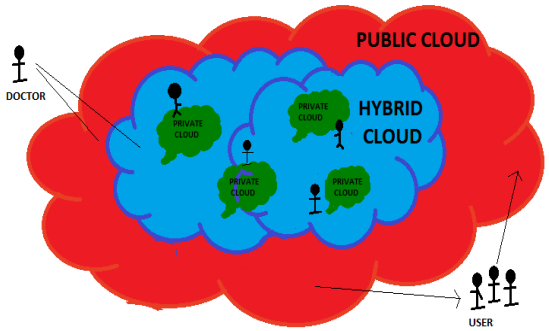

Fig2. *TSCC*

Users collect their health data through the monitoring devices worn or carried, e.g., electrocardiogram sensors and health tracking patches. Emergency medical technician (EMT) is a physician who performs emergency treatment. By user and EMT, were refer to the person and the associated computing facilities. The computing facilities are mainly mobile devices carried around such as smart phone, tablet, or personal digital assistant. Each user is associated with one private cloud. Multiple private clouds are supported on the same physical server. Private clouds are always online and available to handle health data on behalf of the users.. The private cloud will process the data to add security protection before it is stored on the public cloud. Public cloud is the infrastructure usually owned by the cloud providers such as Amazon and Google which offers massive storage and rich computational resource. We assume that at the boot strap phase, there is a secure channel between the user and his/her private cloud. After the bootstrap phase, the user will send health data over insecure network to the private cloud residing via the Internet backbone. Our focus is not on the location privacy of mobile users which can be leaked when sending health data to the private cloud. There is a large body of location privacy schemes [29], [30] in the literature.

5.2. Threat Model

The private cloud is fully trusted by the user to carry out health data-related computations. Public cloud will not delete or modify users' health data, but will attempt to compromise their privacy. Public cloud is not authorized to access any of the health data. The EMT is granted access rights to the data only pertinent to the treatment, and only when emergencies take place. The EMT will also attempt to compromise data privacy by accessing the data he/she is not authorized to. The EMT is assumed to be rational in the sense that he/she will not access the data beyond authorization if doing so is doomed to be caught .Finally, outside attackers will maliciously drop users' packets, and access users' data though they are unauthorized to.

5.3. Security Requirements

In this paper, we strive to meet the following main security requirements for practical privacy-preserving mobile healthcare systems.

- Storage Privacy: Storage on the public cloud is subject to 5 privacy requirements.
- Data confidentiality: unauthorized parties (e.g., public cloud and outside attackers) should not learn the content of the stored data.
- Anonymity: no particular user can be associated with the storage and retrieval process
- Unlink ability: unauthorized parties should not be able to link multiple data files to profile a user. It indicates that the file identifiers should appear random and leak no useful information.
- Keyword privacy: For searching, the keyword used should remain confidential because it may contain sensitive information, which will prevent the public cloud from searching for the desired data files.

- Search pattern privacy: whether the searches were for the same keyword or not, and the access pattern, i.e., the set of documents that contain a key- word [16], should not be revealed. This requirement is the most challenging and none of the existing efficient SSE [15]–[28] can satisfy it. It represents stronger privacy which is particularly needed for highly sensitive applications like health data net- works.
- Audit ability: In emergency data access, the users may be physically unable to grant data access or without the perfect knowledge to decide if the data requester is a legitimate EMT. We require authorization to be fine grained and authorized parties' access activities to leave crypto- graphic evidence.

6. CONCLUSION

This report proposed to build privacy into mobile health systems with the help of the private cloud. We provided a solution for privacy-preserving data storage by integrating levels of cloud, a search and access pattern hiding scheme based on redundancy, and a secure indexing method for privacy-preserving keyword search.

REFERENCES

[1] Yue Tong, Jinyuan Sun, Sherman S. M. Chow, and Pan Li*," Cloud-Assisted Mobile-Access of Health Data With Privacy and Auditability", March 2014

[2] Victor C. Cheng, C.H.C. Leung, Jiming Liu and Alfredo Milani," Probabilistic Aspect Mining Model for Drug Reviews", August 2014

[3] Mohamed Nabeel and Elisa Bertino," Privacy Preserving Delegated Access Control in Public Clouds", 2013

[4] Qin Liu, Guojun Wang and Jie Wu," Consistency as a Service: Auditing Cloud Consistency", 2013

[5] P.N. Polezhaev, A.E. Shukhman, U.A. Ushakov,**"** Mathematical model of cloud computing data center based on OpenFlow", 2010

[6] D. Sannella and Toon Calders," Item set frequency satisfiability: Complexity and axiomatization", 2010

[7] GuizhenYang and Ding-Zhu Du," Computational aspects of mining maximal frequent patterns"

[8] Sajjad Haider and Farhan Bashir Shaikh "Security Threats in Cloud Computing" 2011 Dec

[9] Sahai and Waters, "Fuzzy identity based encryption",Eurocrypt, ser. Lecture Notes in computer science, vol. 3494.Springer,pp. 457-473,2005

[10] V. Goyal, O. Pandey, A. Sahai, B. Waters, "Attribute-based encryption for fine grained access control of encrypted data", ACM Conference on Computerand Communications Security, pp. 89-98,2006

[11] J. Bethencourt, A. Sahaii and B. Waters, "Ciphertext-policy attribute-based encryption ", IEEE Symposium on Security and Privacy. , pp. 321-334, 2007

[12] M. C. Mont, P. Bramhall, and K. Harrison, " A flexible role-based secure messaging service:exploiting IBE technology for privacy in health care," Presented at the 14th Int. Workshop Database Expert Syst. Appl., Prague, Czech Republic, 2003

[13] D. Boneh and M. Franklin, "Identity-based encryption from the Weil Pairing. Extended abstract in CRYPTO 2001," SIAM J. Comput., vol. 32, no. 3, pp. 586-615, 2003

[14] J. Sun, X. Zhu, C. Zhang, andY. Fang, "HCPP: Cryptography based secure EHR system for patient privacy and emergency healthcare," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst.*, Jun. 2011, pp. 373–382.

[15] E.-J. Goh, "Secure indexes," *IACR Cryptology ePrint Archive*, vol. 2003, p. 216, 2003.

[16] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved definitions and efficient constructions," presented at the ACM Conf. Comput. Commun. Security, Alexandria, VA, USA, 2006.

[17] Susan Hohenberger and Brent Waters,"Attribute-Based Encryption with Fast Decryption", May 8, 2013

[18] Vipul Goyal, Omkant Pandey, Amit Sahai and Brent Waters ," Attribute-Based Encryption for Fine-Grained Access Control of Encrypted Data", October 30–November 3, 2006, Alexandria, Virginia, USA.

[19] Xindong Wu , Vipin Kumar , J. Ross Quinlan , Joydeep Ghosh ,Qiang Yang , Hiroshi Motoda , Geoffrey J. McLachlan, Angus Ng , Bing Liu Philip S, Yu Zhi-Hua Zhou , Michael Steinbach ,David J. Hand and Dan Steinberg," Top 10 algorithms in data mining", 4 December 2007, Springer-Verlag London Limited 2007

[20] M. Hu and B. Liu, "Mining and summarizing customer reviews," in Proc. 10th ACM SIGKDD Int. Conf. KDD, Washington, DC, USA, 2004, pp. 168–177.

[21] Q. Mei, X. Ling, M. Wondra, H. Su, and C. Zhai, "Topic sentiment mixture: Modeling facets and opinions in weblogs," in Proc. 16th Int. Conf. WWW, New York, NY, USA, 2007, pp. 171–180.

[22] A.-M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in Proc. Conf. Human Lang. Technol. Emp. Meth. NLP, Stroudsburg, PA, USA, 2005, pp. 339–346.

[23] I. Titov and R. McDonald, "A joint model of text and aspect rat- ings for sentiment summarization," in Proc. 46th Annu. Meeting ACL, 2008, pp. 308–316.

[24] B. Liu, M. Hu, and J. Cheng, "Opinion observer: Analyzing and comaring opinions on the web," in Proc. 14th Int. Conf. WWW, New York, NY, USA, 2005, pp. 342–351.

[25] S. Baccianella, A. Esuli, and F. Sebastiani, "Multi-facet rating of product reviews," in Proc. 31st ECIR , Berlin,, Germany, 2009, pp. 461–472.

[26] W. Jin, H. Ho, and R. Srihari, "Opinionminer: A novel machine learning system for web opinion mining and extraction," in Proc. 15th ACM SIGKDD Int. Conf. KDD, New York, NY, USA, 2009, pp. 1195–1204.

[27] Y. Jo and A. Oh, "Aspect and sentiment unification model for online review analysis," in Proc. 4th ACM Int. Conf. WSDM, New York, NY, USA, 2011, pp. 815–824.

[28] D. Song, D.Wagner, and A. Perrig, "Practical techniques for searching on encrypted data," in *Proc. IEEE Symp. Security Privacy*, 2000, pp. 44–55.

[29] A. Pingley,W. Yu, N. Zhang, X. Fu, andW. Zhao, "CAP: A context-aware privacy protection system for location-based services," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst.*, 2009, pp. 49–57.

[30] T. Xu and Y. Cai, "Location cloaking for safety protection of ad hoc networks," in *Proc. IEEE Conf. Comput. Commun.*, 2009, pp. 1944–1952.

[31] M. Li, S. Yu, Y. Zheng, K. Ren, andW. Lou, "Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 1, pp. 131–143,Jan. 2013.

[32] M. Chase and S. S. M. Chow, "Improving privacy and security in multiauthority attribute-based encryption," in *Proc. ACM Conf. Comput. Commun. Security*, 2009, pp. 121–130.

[33] S. S. M. Chow, "New privacy-preserving architectures for identity -/attribute-based encryption" Ph.D. dissertation, Courant Inst. Math. Sci., New York University, New York, NY, USA, 2010.

[34] S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving secure, scalable, and fine-grained data access control in cloud computing," presented at the IEEE Conf. Comput. Commun., San Diego, CA, USA, Mar. 2010.