RESEARCH ARTICLE OPEN ACCESS

A Fast Nearest Neighbour Search Scheme Over Outsourced Encrypted Medical Images

Mrs J.Sarada MCA, M.Tech, M.Phil, (Ph.D) [1], G.Akhil [2]
[1]Associate Professor, Department of Computer Applications
[2] Student, Department of Computer Applications
[1],[2] Chadalawada Ramanamma Engineering College (Autonomous)

ABSTRACT
The delicate nature of medical photographs mandates the use of stringent security and privacy measures. Medical pictures should be encrypted before being sent to a cloud-based medical system for Healthcare Industry 4.0. However, querying encrypted data without first decrypting it is difficult and impracticable at the moment. Over encrypted medical pictures, we offer a safe and fast method for finding the precise next-door neighbour. Because the mean and standard deviation are directly connected to the lower limit of Euclidean distance, we dismiss candidates by determining the lower bound of Euclidean distance. In contrast to most current systems, ours can provide the precise closest neighbour rather than an approximation of it. We next conduct an evaluation of our suggested method in order to prove its usefulness.
Keywords: - Cloud Computing, Efficiency, Medical Images,Nearest Neighbor Search, Privacy

I. INTRODUCTION

Our culture is moving toward LOUD computing [1], which allows data owners to outsourceCloud server with databases and management tools.The latter stores the databases and offers access methodsto query and administer the outsourced database. This permits dataowners to minimise data management expenditures and improvequality of service. However, the cloud may not be totally trustedbecause it may expose critical information to unauthorisedenterprises (e.g., compromised) or foreign government agency[2].The fast rise of cloud computing is transformingThe Entire Industry of e-Health 4.0 in the realm of healthcare.One of the most widely used is a cloud-based electronic healthcare system.application for Healthcare Industry 4.0. A well-designedelectronic healthcare system can certainly improve the qualityOf healthcare consumers' ability to obtain and utilise services. Recent years have seen a flurry of interest in Healthcare Industry 4.0 applications for cloud computing and big data computing[3, 4, 5]). In an electronic healthcare system, patients' medical images may be outsourced to a third-party [6], such as a professional community healthcare cloud server. There are various types of medical imaging in the healthcare business that are important for diagnosis and treatment quality, including MRI, ultrasound, computed tomography (CT), and computed radiography (CR). Additionally, earlier or archive data (such as those pertaining to disease outbreaks) may be a valuable resource for the healthcare business. When a new patient is examined, clinicians may rapidly and reliably establish a diagnosis and design the proper treatment regimens by identifying comparable instances in the database and analysing them. For example, hospitals or associated medical organisations may save patients' medical images in a professional and secure database of a realistic electronic healthcare system. If clinicians get medical images from a new patient, they may identify comparable images in the outsourced database to be used as a reference. This is especially essential when dealing with symptoms fromsuch as those on the https://globalgenes.org/rarelist/ list of rare illnesses Access to this page was last updated on January 16, 2018. Hence, a search for the nearest neighbour may be employed in this situation. However, the sensitivity and privacy of medical images demand that the security and privacy of such images be secured and safeguarded.

Encrypting data by the data owner is a naïve method to secure privacy [7], although it ensures the secrecy of the outsourced data from the cloud and unwanted users. Additionally, to preserve query privacy, allowed users should submit their queries to the cloud for review after encryption. However, by analysing the data access patterns, the cloud (or a malicious insider) may extract private information about the genuine data items even while the data and queries are encrypted [8],[9].

In particular, the access pattern covers not only the content of the data block accessible by the user but

also the method how the user accesses the data block, such as frequency, location, order, habit and so on. Data mining, statistical analysis, and other methods allow the cloud server to determine the kind of user, their interests, and the frequency with which they access various types of material. For example, traffic analysis technology may obtain certain sensitive information regarding access pattern. When using a search engine like Google, a user's search patterns, including their identity, may be seen in their history. Also, the frequency of the search may leak the popularity of the obtained data. A correlation between subsequent accesses may also be established by the cloud server. In other words, we need to assure secrecy of the outsourced data and a user's query record in secure query processing, as well as masking data access patterns. The proposed scheme must reduce the end-calculation user's costs in order to be useful and efficient.

The nearest neighbour search is an important operation in data mining, machine learning, and information retrieval, and more recently the healthcare business as well. The need for a fast and space-efficient nearest neighbour search algorithm has recently grown in relevance due to the rise of high-dimensional medical images. Generally, it is a challenging effort to handle encrypted data without first completing the decryption procedure. The challenge is how a cloud server can perform the queries over encrypted medical images. An efficient and successful method for finding the precise next-door neighbour using outsourced encrypted medical images has been proposed in response to this problem.

Specifically, in the study, we examine the topic of precise nearest neighbour search on encrypted medical images and propose a safe and efficient solution. Our scheme permits dynamic updating. It enables data users to simply add or remove medical images as required.

Nearest neighbour search will be discussed in the next section. In Section III, we look through the system model and the project's overall aims. The proposed scheme is outlined in Section IV. Section V summarises our security and performance findings, and Section VI wraps up the rest of the study.

II. RELATED WORK

Research on Knuth's [10] 1973 study of nearest neighbour search has continued ever since. In the literature, there are a great number of approaches linked to the query process over encrypted data, such as searchable encryption [11]. For the purpose of resisting keyword assaults, Curtmola and coworkers in 2006 presented a novel searchable symmetric encryption (SSE) structure. In order to maximise dynamic operations, Kamara et al. proposed a dynamic SSE scheme and a novel security architecture [13]. In 2013, Kamara et al. proposed a novel dynamic SSE scheme, which is designed to facilitate parallelizable search [14]. However, these techniques rely on keyword search, rather than the nearest neighbour search. For an overview of current SSE initiatives, we recommend that interested readers see [15].

One might also employ sequential scan (brute-force search), which successively calculates and compares the Euclidean distance between the query item and every record in the encrypted database. This scheme, however, requires large amounts of time and space, both of which are directly related to the volume and number of the data. As a result, despite the availability of strategies to decrease expenses, such a scheme is not ideal for dealing with large-scale and high-dimensional data. In 2009, Wong et al. [16] recommended looking for nearest

Using the asymmetric scalar product preserving encryption (ASPE) scheme, neighbours may communicate with each other through encrypted data. However, employing the scheme demands linear search time according to the number of data records. Hu et al. [17] introduced a scheme using tree-based data structures and ASPE two years later in 2011, resulting in a quicker search time. Unfortunately, with this scheme the client that intends to conduct queries has to perform several interactions with the server. Also, the necessity to keep a local index will entail (large) local storage expenses. In 2006, Zhu et al. [18] designed a more secure scheme that is resistant to cloud server assaults and does not need the use of shared keys. Zhu et al. [19] in 2017 proposed an efficient scheme for k-nearest neighbour search, which increases the security of the decryption key and reduces the load on data owners.

There are a number of efficient ways to determine the estimated nearest neighbour, designed to improve efficiency in space and time at the

sacrifice of accuracy. The scheme based on Locality Sensitive Hashing (LSH) is a well-known and successful method that addresses the nearest neighbour query issue in high-dimensional space. The LSH scheme [20] embeds data in low-dimensional subspaces and leverages hash tables to improve performance. In 2004, Datar et al. [21] proposed a basic LSH based on the original LSH scheme, which leveraged the property of p-stable distribution to expand the LSH method from the Hamming space to the Euclidean space. There is a large amount of extra space required for this scheme. In 2007, Andoni et al. [22] introduced the Leech lattice into the LSH scheme of [20], which reduces the query time and memory usage. Hashing methods are beneficial to cope with high-dimensional and large-scale data, but they demand extra cost when employed in the precise nearest neighbour search. Due to the demand for accuracy in the healthcare business, these LSH-based methods are not suitable for handling the medical images difficulties.

Tree-based data structures have also been proposed by researchers as a method for quickly locating the plaintext domain's nearest neighbour. As early as 1975, Bentley [23] proposed KD-tree and utilised this particular data structure to hold information that may be accessed via associative searches. This single data structure is capable of handling a wide range of query types. The K-nearest neighbour (KNN) search issue involving high-dimensional data was addressed by Jagadish et al. [24] three decades later and an efficient B+ tree structure was presented to them. Other common examples are R-tree variants [25], [26], [27] and Cover tree [28]. As a search of the data being based on a tree-like data, these algorithms may reduce the cost of searching. The preprocessing time and memory space required by these data structures means that they are not suitable for large-scale or high-dimensional data.

It was proposed by Ahn et al. [27] in 2012 that a plaintext nearest neighbour search algorithm may be developed. This algorithm reduces the size of data points by enclosing them in a small space. Non-nearest neighbours are eliminated by a comparison of the distances in this space. Due to this property, this algorithm is suitable for processing high-dimensional and large-scale data.In particular, unlike most current algorithms, this method may obtain the actual nearest neighbour rather than an approximation one. The beauty of this method is simplicity, in the sense of straightforward preprocessing without needing complicated data structures.

In summary, most of the current approaches have limitations and are not appropriate to the healthcare industry's medical imaging challenge. It's worth noting, however, that the algorithm proposed by Ahn et al. [29] is capable of ensuring both efficiency and accuracy at the same time. Based on this algorithm, we provide an efficient scheme to search for the precise nearest neighbour in an outsourced medical picture database. Using the data's mean and standard deviation, we compute at a lower limit on Euclidean distance in our scheme. It is not essential to import all data to compute the Euclidean distances in the original high-dimensional space. Checking the lower bound may remove a huge number of the closest neighbour possibilities.

III. EXISTING SYSTEM

The nearest neighbor search is a vital operation in data mining, machine learning, and information retrieval, and more recently the healthcare industry as well. Recently, due to the emergence of high-dimensional medical images, the importance of having an efficient and effective nearest neighbor search algorithm (in terms of speed and space) has become more pronounced. Generally, it is a difficult task to process encrypted data without first executing the decryption operation. The challenge is how a cloud server can process the queries over encrypted medical images.

Disadvantages

In summary, most of the existing techniques have limitations and are not applicable to the healthcare industry's medical imaging problem.

IV. PROPOSED SYSTEM

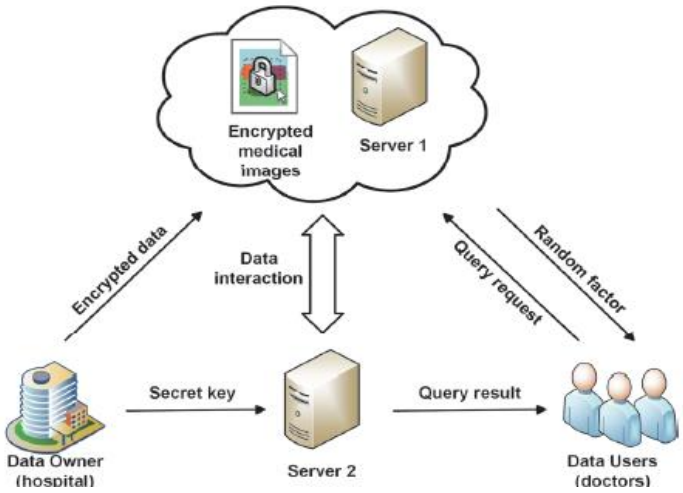

Fig1 Architecture for searching over encrypted data

We propose an efficient and effective scheme to search for the exact nearest neighbor over outsourced encrypted medical images. Specifically, in the paper, we discuss the problem of exact nearest neighbor search over encrypted medical images and propose a secure and efficient solution. Our scheme supports dynamic updates. It allows data users to easily add or delete medical images whenever necessary. Additionally, to protect query privacy, permitted users should send their requests to the cloud for evaluation after encryption. However, by analyzing the data access patterns, the cloud (or a malicious insider) can derive private information about the real data items even though the data and queries are encrypted. Encrypting data by the data owner is a naive method to ensure privacy [7], while it ensures the secrecy of the outsourced data from the cloud and unauthorized users. Additionally, to protect query privacy, permitted users should send their requests to the cloud for evaluation after encryption. In the paper, we propose a secure and efficient scheme to find the exact nearest neighbor over encrypted medical images. Instead of calculating the Euclidean distance, we reject candidates by computing the lower bound of Euclidean distance that is related to the mean and standard deviation of data.

Advantages

However, we also observe that the algorithm proposed by Ahn et al. [9] can simultaneously ensure both efficiency and accuracy. Based on this algorithm, we present an efficient scheme to search for the exact nearest neighbor in an outsourced medical image database. In our scheme, we compute the lower bound of Euclidean distance that is related to the mean and standard deviation of data. It is not necessary to load all data to compute the Euclidean distances in the original high-dimensional space. A large number of the nearest neighbor candidates can be eliminated by checking the lower bound.

V. IMPLEMENTATION

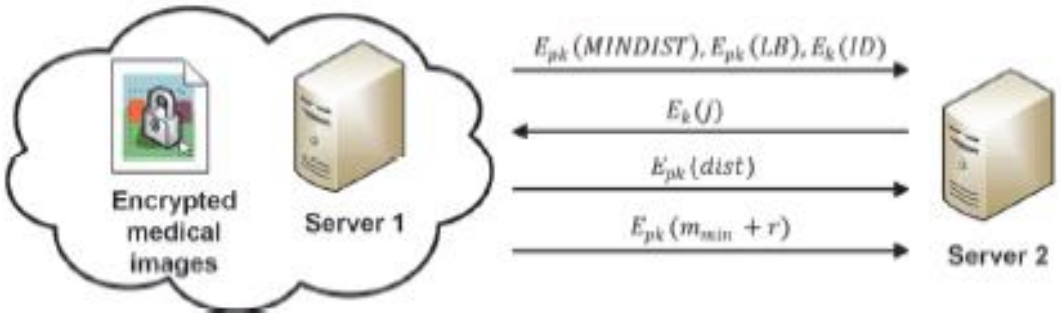

Fig 2 Implementation of two servers

- Data Owner

  In this module, the data owner Collect Patient data and Upload to Cloudlet like pid,pname,paddress,pcno,pemail,ppulse,pecg,pSymptoms,brwose and attach about symptoms with Digital sign,addpimage(Encrypt all parametes except pname) and View all patient collecte data in enc format with digital sign.

- Server A

  The server-A manages which is to provide data storage service for the wearable devices and also View all patients and authorize and View all doctors and authorize ,Vliew all patient Cloudlet data with enc format ,View Patient data access request and authorize ,View all Cloudlet Intruders details and View patient details recovered details ,View No.Of same symptoms in Chart(Symptom name vs No. Of Patients),ViewNo.Of Patients refered same doctor in Chart(Doctor name vsNo.Of Patients).

- Data User

  In this module, the patient Register and Login, View profile ,Request Data Access permission from cloudlet and view Response, Access Your data and select doctor from combo box and send to corresponding doctor and View doctor response with Medical prescription, Verify your data and recover and View and delete your details.

- Server-B

  The Server-B is the one who will perform the following operations such as Register and Login,View Profile, View patient details and give solution like Medicine details,Medical prescription details View all patient Medical prescription Details.

VI. CONCLUSION

Cloud-based electronic healthcare systems will be increasing popular, particularly due to the capability to share and access data in real-time across organizations (e.g., between medical practitioners and healthcare providers) and countries. One process becomes challenging, if not impractical. In the paper, we presented a secure and efficient scheme to locate the exact nearest neighbor over encrypted medical images stored in the remote cloud server. For the purpose of rejecting candidate data points, our scheme securely computes the lower bound of the squared Euclidean distance between a data point in the database and the query submitted by a legitimate user. The performance of our scheme is evaluated using real-world medical images.

REFERENCES

[1] J. Li, L. Huang, Y. Zhou, S. He, Z. Ming, "Computation partitioning for mobile cloud computing in big data environment," IEEE Trans. Ind. Informat., vol. 13, no. 4, pp. 2009-2018, Feb. 2017.

[2] K.-K. R. Choo, "Cloud computing: Challenges and future directions,"

Trends & Issues in Crime and Criminal Justice, vol. 400, no. 400, pp. 1– 6, Oct. 2010.

[3] M. Pajic, R. Mangharam, O. Sokolsky, D. Arney, J. M. Goldman, and I. Lee, "Model-driven safety analysis of closed-loop medical systems," IEEE Trans. Ind. Informat., vol. 10, no. 1, pp. 3–16, Feb. 2014.

[4] B. Xu, L. D. Xu, H. Cai, C. Xie, J. Hu, and F. Bu, "Ubiquitous data accessing method in IoT-based information system for emergency medical services," IEEE Trans. Ind. Informat., vol. 10, no. 2, pp. 1578– 1586, May. 2014.

[5] G. Yang et al., "A health-IoT platform based on the integration of intelligent packaging, unobtrusive bio-sensor, and intelligent medicine box," IEEE Trans. Ind. Informat., vol. 10, no. 4, pp. 2180–2191, Nov. 2014.

[6] H. Huang, T. Gong, N. Ye, R. Wang, and Y. Dou, "Private and Secured Medical Data Transmission and Analysis for Wireless Sensing Healthcare System," IEEE Trans. Ind. Informat., vol. 13, no.3 pp. 1227-1237, June. 2017.

[7] M. Li, S. Yu, W. Lou, and Y. T. Hou, "Toward privacy-assured cloud data services with flexible search functionalities," in Proc. ICDCSW. IEEE, Macau, CHN, 2012, pp. 466–470.

[8] P. Williams, R. Sion, and B. Carbunar, "Building castles out of mud: practical access pattern privacy and correctness on untrusted storage," in Proc. CCS. ACM, Alexandria, VA, USA, 2008, pp. 139–148.

[9] M. S. Islam, M. Kuzu, and M. Kantarcioglu, "Access pattern disclosure on searchable encryption: Ramification, attack and mitigation," in NDSS, San Diego, CA, USA, 2012.

[10] D. E. Knuth, "Sorting and searching," in The art of computer programming, vol. 3, Boston, USA: Addison-Wesley, 1973.

[11] D. Song, D. Wagner, and A. Perrig, "Practical Techniques for Searches on Encrypted Data," in Proc. of IEEE S&P, DC, USA, 2000, pp. 44-55.

[12] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," J. Comput.Secur., vol. 19, no. 5, pp. 895-934, 2011.

[13] S. Kamara, C. Papamanthou, and T. Roeder, "Dynamic Searchable Symmetric Encryption," in Proc. of ACM CCS, Raleigh, NC, USA, 2012, pp. 965–976.

[14] S. Kamara, C. Papamanthou, "Parallel and dynamic searchable symmetric encryption," Financial Cryptography and Data Security, Springer Berlin Heidelberg, 2013, pp. 258-274.

[15] G. S. Poh, J.–J. Chin, W.-C.Yau, K.-K. R. Choo, and M. S. Mohamad, "Searchable Symmetric Encryption: Designs and Challenges," ACM Comput. Surv.vol. 50, no. 3, pp. 40:1-40:37, 2017.

[16] W. Wong, D. W. Cheung, B. Kao, and N. Mamoulis, "Secure KNN Computation on Encrypted Databases," in Proc. ACM SIGMOD, Providence, RI, USA, 2009, pp. 139–152.

[17] H. Hu, J. Xu, C. Ren, and B. Choi, "Processing Private Queries over Untrusted Data Cloud through Privacy Homomorphism," in Proc. IEEE ICDE, Hannover, NI, GER, 2011, pp. 601–612.

[18] Y. Zhu, Z. Huang, and T. Takagi, "Secure and controllable k-NN query over encrypted cloud data with key confidentiality," Journal of Parallel & Distributed Computing, vol. 89, pp. 1-12, Mar. 2016.

[19] L. Zhou, Y. Zhu, and A. Castiglione, "Efficient k -NN query over encrypted data in cloud with limited key-disclosure and offline data owner," Computers & Security, vol. 69, pp. 84-96, Aug. 2017.

[20] P. Indyk and R. Motwani, "Approximate nearest neighbors: Towards removing the curse of dimensionality," In Proc. STOC, Dallas, TX, USA, 1998, pp. 604– 613.

[21] M. Datar, N. Immorlica, P. Indyk, and V. S. Mirrokni, "Locality sensitive hashing scheme based on p-stable distributions," in Proc. SCG. ACM, Brooklyn, NY, USA, 2004, pp. 253–262.

[22] A. Andoni, P. Indyk, "Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions," IEEE Symposium on Foundations of Computer Science, Berkeley, CA, USA, 2006, pp. 459-468.

[23] J. L. Bentley, "Multidimensional binary search trees used for associative searching," Communications of ACM, vol. 18, no. 9, pp. 509–517, Sep. 1975.

[24] H. V. Jagadish, B. C. Ooi, K.-L. Tan, C. Yu, and R. Zhang, "iDistance: An adaptive B+-tree based indexing method for nearest neighbor search," ACM Trans. Database Syst., vol. 30, no. 2, pp. 364–397, June. 2005.

[25] S. Arya, D. Mount, N. Netanyahu, R. Silverman, and A. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," Journal of the ACM, vol. 45, no. 6, 1998, pp. 891–923.

[26] S. Berchtold, D. A. Keim, and H. P. Kriegel, "The x-tree: An index structure for high-dimensional data," in Proc. VLDB Conf., Mumbai (Bombay), India, 1996, pp. 28– 39.

[27] A. Guttman, "R-trees: A dynamic index structure for spatial searching," in Proc. SIGMOD. ACM, Boston, MA, USA, 1984, pp. 47–57.

[28] A. Beygelzimer, S. Kakade, and J. Langford, "Cover trees for nearest neighbor," In Proc. ICML. ACM, Pittsburgh, PA, USA, 2006, pp. 97– 104.

[29] Ahn H K, Han B, and Hwang Y, "A fast nearest neighbor search algorithm by nonlinear embedding," in Proc. CVPR. IEEE, Providence, RI, USA, 2012, pp. 3053-3060.