IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⸎ Impact Factor 7.39 ⸎ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11758

DATA ANONYMIZATION USING PSEUDONYM SYSTEM TO PRESERVE DATA PRIVACY

Kavya.S[1], Prof.M.N.Chandan[2]

PG Scholar, Dept. of MCA, PES College of Engineering, Mandya, Karnataka, India[1]

Assistant Professor, Dept. of MCA, PES College of Engineering, Mandya, Karnataka, India[2]

Abstract: Every business or organisation regularly collects and stores large amounts of data. Two methods that are frequently utilised to do this are cloud computing and wireless network infrastructure. Customers that use these services can perform their duties more swiftly and easily while obtaining the desired results. In order to store information in a digital database, general services use a special identifier. It might, however, have some restrictions and difficulties. The unique identifier is linked to the data owner's name, address, identity card number, and other facts. Attackers can take the complete data set by altering a unique identifier. Attackers may even use eavesdropping or educated guesses to get the necessary data. The result is a lack of data privacy protection. Therefore, wherever digital data is held, data privacy concerns must be taken into account. When using current services, there is a considerable risk that data or information will be revealed to or leaked to an unidentified user during transfer. Additionally, during exchange of information, attacks against services, such as impersonating and forgery attacks, may occur. This study suggests using a palm vein-based biometric authentication method to assuage these concerns. The database record is made anonymous, and the data is adequately protected, using a pseudonym generation approach. This protects information and data from unauthorised access. The suggested remedy can stop information breaches, and a user's genuine identity is never revealed.

INDEX TERMS: Data preservation, pseudonym, anonymity, unlink ability, and palm vein authentication

OVERVIEW

Using commonly used methods could have the drawback of controllability. Information is linked as a result, and the flow of information is unhindered by any central authority. Their strategy upholds user privacy while promoting data exchange and information control. They concluded that the universal composability (UC) framework has been used to validate the proposed system. They also provided strong evidence in support of a claim regarding the discrete logarithm. Finally, they have asked prospective researchers to carefully consider the numerous methods that have been outlined in the literature for providing safe pseudonyms. A study on the private information data aggregation specifically in wireless sensor networks was done in [3]. The most popular privacy preservation strategies were utilised to classify and compare the various data protection aggregation approaches now in use. The results demonstrated that the majority of possibilities went through a stage known as "initialization." Participants ask significant issuers for substantial authentication during this step using a secure channel. Therefore, it is essential to create effective protocols that are free from reliable authority and have secure two-way communication channels.

LITERATURE SURVEY

Camenisch, J. and A. Lehmann. (Un) linkable Pseudonyms for Governmental Databases. in Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. 2015. ACM

When decentralised data needs to be synchronised or shared between databases, linked data sets are typically given a unique identification. Although this method makes cross-domain data sharing easier, it also has inherent controllability problems. No centralised authority can restrict or regulate the flow of information because data records can be easily connected. Even worse, such linkability poses a serious security and privacy risk when records contain sensitive personal data, as is the case, for example, in national social security systems. Using domain-specific pseudonyms, where only a central authority is aware of the cross-domain relationship between the pseudonyms, is an alternate strategy. The central authority must, however, be completely trusted in order for the existing methods to work because, otherwise, it might supply erroneous conversions and take advantage of the information it gathers from the requests. We suggest a (un)linkable pseudonym system that gets over these restrictions and permits managed but privacy-protective interchange of remote data. We demonstrate the security of our protocol within the UC framework and offer an effective instantiation based on related discrete-logarithm presumptions.

DOI: 10.17148/IJARCCE.2022.11758

Yang, J.-J., J.-Q. Li, and Y. Niu, A hybrid solution for privacy preserving medical data sharing in the cloud environment. Future Generation Computer Systems, 2015. 43: p. 74-86
The development of the healthcare system has been hastened by the explosive expansion of cloud and IoT technologies. The Internet of Things (IoT) devices track the patient's wellbeing and upload the gathered information as Electronic Medical Records (EMRs) to the cloud for archival and distribution. EMR cloud outsourcing creates new security and privacy issues. In this study, we suggested a unique architecture that ensures the security and privacy of the outsourced medical records. The group-based access structure in the proposed model is built using partially ordered set (POSET), and the fine-grained EMR access control is provided by ciphertext-policy attribute-based encryption (CP-ABE). The modified group based CP-ABE (G-CP-ABE) decreases the computational overhead by lowering the number of leaf nodes in the access tree. Additionally, the proposed G-CP-ABE framework combines symmetric encryption and the CP-ABE algorithm to reduce overall encryption time. As a result, G-CP-ABE can be used to monitor health conditions even from an IoT device with limited resources. The proposed model is effective, as demonstrated by the performance study, making it acceptable for usage in real world applications.

Sharma, S. and D. Shukla. Efficient multi-party privacy preserving data mining for vertically partitioned data. in 2016 International Conference on Inventive Computation Technologies (ICICT). 2016
Digitally stored data in the computational domain. This type of data format requires less work and storage. As a result, many institutions and organisations are keeping their information in this format. The primary subject of research in the work that has been provided is data and how to protect their privacy. The proposed work takes into account an organisation where choices are based on information from several departments and their characteristics. Additionally, all departments' characteristics are needed in order to make decisions. However, the departments are unable to reveal the owner of the data's privacy. Therefore, combining the data attributes and mining the combined data require a privacy-preserving technique to address privacy concerns in a central database. In this essay, we show how data from many departments within the same company can be combined without endangering client privacy in order to make judgments that are both accurate and efficient. Thus, the approach of vertical data combining, cryptography, and decision mining is shown. A C4.5 decision tree is employed to extract the decisions from the data. Java technology is used to accomplish the suggested privacy-preserving data mining and decision-making technique. Additionally, the system's performance is calculated in terms of precision, error rate, memory usage, and time usage. Finally, the standard J4.5 tree utilising the WEKA tool is employed with the same data for a comparative performance analysis in order to support the results of the suggested data mining technique. The experimental findings demonstrate the effectiveness and security of the chosen privacy-preserving method.

Liu, Z., et al., Uncoordinated Pseudonym Changes for Privacy Preserving in Distributed Networks. IEEE Transactions on Mobile Computing, 2019.
Participatory sensing applications are emerging as a result of the widespread use of inexpensive embedded sensors, such as cameras, microphones, accelerometers, and others, on mobile devices. While collecting and analysing the position and trajectory data of participants may compromise their privacy, participatory sensing can be extremely beneficial to both individuals and communities. However, the majority of the current ideas concentrate on the location privacy of participants, and very few on the trajectory privacy of participants. Participants' locations and any pertinent personal privacy will be revealed through an effective analysis of trajectory data that contains spatial-temporal historical information. In this research, we suggest a paradigm for participatory sensing called TrPF that protects trajectory privacy. Based on the framework, we enhance the theoretical mix-zones model by taking the time component into account from the standpoint of graph theory. We then compare the effectiveness of our proposal to earlier trajectory privacy measures and undertake threat model analyses using various backgrounds. The research and simulation findings show that our approach may successfully safeguard the privacy of participants' trajectory information while incurring less information loss and expenditure than those offered by the competing ideas.

Balakumar, P. and R. Venkatesan, A survey on biometrics based cryptographic key generation schemes. International Journal of Computer Science and Information Technology & Security, 2012. 2(1): p. 80-85
In all systems, the problem of precise and automatic user detection and authentication is crucial. Currently, only a few scenarios can be adequately protected by shared secrets like PINs or passwords and key objects like smart cards. A method that could confirm the identity of the person is what is needed. Recognition of people is becoming easier because to biometrics. A living individual can be recognised based on their physiological or behavioural characteristics using a biometrics system without the need for a human being to intervene. The use of biometrics to safeguard the system as a result spreads globally. Fingerprints, iris scans, retina scans, and other biometrics are some of the ones used to secure systems. In general, the system is secured by the creation of cryptographic keys from biometrics. Cryptography works well for securing purposes because of its efficiency and adaptability. Occasionally, biometric data can be stolen, allowing attackers to get access to the system whenever they want. By utilising two biometric traits, this difficulty is lessened in

DOI: 10.17148/IJARCCE.2022.11758

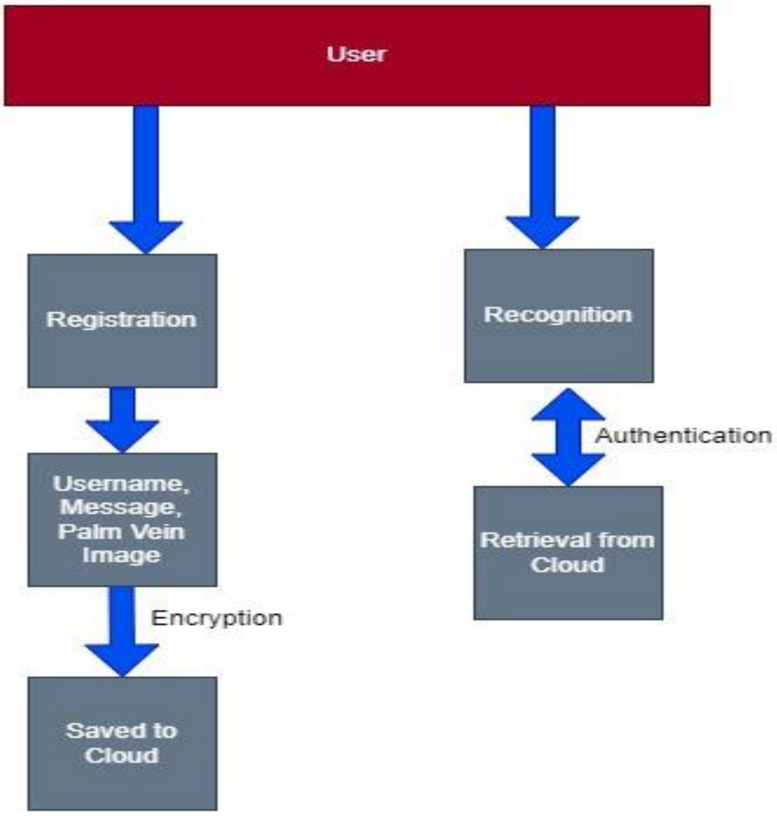
this paper. Fingerprint and iris biometrics are employed in this paper. Using the fusion algorithm, these two features are integrated. A cryptographic key is created using the features combined. The experimental finding indicates that the proposed strategies produce greater security than the ones already in use.

BACKGROUND STUDY

The literature offers numerous strategies for protecting data privacy. On the other hand, as data increases in size, it must be shared with other parties in order to accomplish specific objectives. Furthermore, when data is shared, it may become open to malicious assaults from careless people or organisations. Techniques that are both extremely effective and trustworthy must be used in order to ensure data privacy. The reference covered three techniques that are currently being used to achieve this. The three methods are operations on encrypted data, de-identification, and privacy-preserving aggregation. But the lack of data anonymity rendered the solutions useless. However, they did come to the conclusion that if a very efficient and privacy-protective algorithm could be created for that goal, the de-identification technique was the most private among them. In an effort to protect the privacy of data and information, numerous scholars are said to have produced a number of solutions. A useful classification for privacy-preserving techniques is still lacking, though. As a result, the data preservation methods were divided into two groups: cryptography and non-cryptographic. Attackers do, however, also use cutting-edge technology and consistently show off their own inventive techniques for stealing data and information. This demonstrated how urgently reliable algorithms must be developed in this field.

PROPOSED METHOD

The primary goal of the ongoing research is to use palm vein to ensure data privacy while utilising the pseudonym method. The palm vein has already been proven to be a trustworthy biometric trait that can be utilised to create pseudonyms. Homomorphic encryption is used to create pseudonyms. accentuates the privacy-protection technique that has been proposed. In order to conduct this investigation, the palm vein scanner was used to take the palm vein image. Noise must often be removed from the acquired photos by filtering. The feature border is examined using the edge detection. The appropriate conversion of intravascular signals into the secret key requires consideration of a number of factors in the chosen feature. To demonstrate that the proposed method may successfully defend online privacy against anticipated attacker efforts, two partial keys are produced from the vein in the hand. After that, the last alias is created using the missing keys.



DOI: 10.17148/IJARCCE.2022.11758

CONCLUSION

Today, storing enormous amounts of data and protecting it effectively are requirements for all enterprises and organisations. Cloud computing services are frequently used for the effective handling of such data. Anonymized data are the major focus of the current investigation for data protection considerations. A novel method was introduced for creating pseudonyms that would ensure maintaining confidentiality.

BIBLIOGRAPHY

[1] N. M. Shrestha, A. Alsadoon, P. W. C. Prasad, L. Hourany, and A. Elchouemi, ‘‘Enhanced E-health framework for security and privacy in healthcare system,’’ in Proc. 6th Int. Conf. Digit. Inf. Process. Commun. (ICDIPC), Apr. 2016, pp. 75–79.

[2] J. Camenisch and A. Lehmann, ‘‘(Un)linkable pseudonyms for governmental databases,’’ in Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur. (CCS), 2015, pp. 1467–1479.

[3] J. Xu, G. Yang, Z. Chen, and Q. Wang, ‘‘A survey on the privacypreserving data aggregation in wireless sensor networks,’’ China Commun., vol. 12, no. 5, pp. 162–180, May 2015.

[4] J.-J. Yang, J.-Q. Li, and Y. Niu, ‘‘A hybrid solution for privacy preserving medical data sharing in the cloud environment,’’ Future Gener. Comput. Syst., vols. 43–44, pp. 74–86, Feb. 2015.

[5] S. Sharma and D. Shukla, ‘‘Efficient multi-party privacy preserving data mining for vertically partitioned data,’’ in Proc. Int. Conf. Inventive Comput. Technol. (ICICT), Aug. 2016, pp. 1–7.

[6] M. M. Potey, C. A. Dhote, and D. H. Sharma, ‘‘Homomorphic encryption for security of cloud data,’’ Procedia Comput. Sci., vol. 79, pp. 175–181, Mar. 2016.

[7] V. Akila and T. Sheela, ‘‘Preserving data and key privacy in data aggregation for wireless sensor networks,’’ in Proc. 2nd Int. Conf. Comput. Commun. Technol. (ICCCT), Feb. 2017, pp. 282–287.

[8] K. Gu, N. Wu, B. Yin, and W. J. Jia, ‘‘Secure data query framework for cloud and fog computing,’’ IEEE Trans. Netw. Service Manage., to be published.

[9] Z. Liu, L. Zhang, W. Ni, and I. Collings, ‘‘Uncoordinated pseudonym changes for privacy preserving in distributed networks,’’ IEEE Trans. Mobile Comput., to be published.

[10] C. N. H. Vinh, A. Truong, and T. T. Huu, ‘‘A privacy preserving authentication scheme in the intelligent transportation systems,’’ in Proc. Int. Conf. Future Data Secur. Eng. Cham, Switzerland: Springer, 2018, pp. 103–123.