International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 09 | Sep 2020 www.irjet.net p-ISSN: 2395-0072

Algorithm for Storage to Protect Data Privacy in Location-based Services

S V Charisma[1], S Ramani[2], Mohana Kumar S[3]

[1]S V Charisma Master of Technology, Department of Computer science and engineering, Ramaiah Institute of Technology, Bangalore, India
[2]S Ramani Master of Technology, Department of Computer science and engineering, Ramaiah Institute of Technology, Bangalore, India
[3]Mohana Kumar S Master of Technology, Department of Computer science and engineering, Ramaiah Institute of Technology, Bangalore, India

---

Abstract - Data privacy is an important aspect and major challenge when location-based services are provided to the user as potential hackers are looking for various ways to misuse data for wrong reasons. This paper proposes a combination of two algorithms perfect shuffle algorithm and hash salt algorithm to protect data privacy in location based services. Perfect shuffle algorithm is an encryption algorithm that exchanges and shuffles the data; hash salt algorithm is a cryptographic encryption algorithm that consists of SHA 512 and salt which is a 11-digit random key added to SHA 512 to provide extra layer of security to the user's data. The algorithms are combined to protect user's data privacy. However, an account of the data in a database needs to be kept securely for future legal purposes; therefore, a Super admin had been authorized for handling the data securely in a web application. In this manner, data privacy is protected both at the user end and at the service provider end.

Key Words*:* data privacy, database, data masking, encryption

1. INTRODUCTION

Location based services emerged as a result of internet accessibility through mobile phones, being able to position the user's device and rich user interfaces. Today, location-based services are technology capable of delivering applications personalized to a user's geographic location, based on a specified mobile device for a specific reason, mainly in the fields of emergency and personal health, navigation, and access to on-the-go tourist information, all of which are of interest to the customer [1] [8]. As location based systems primarily depend on the position of the smartphone of the customer, the primary purpose of the service provider is to decide where the customer is [3] [7]. For example, location based services that direct a user to the nearest restaurant. In another example, location-based apps may deliver an SMS message about a local shopping mall advertising a deal. Location-based systems may be used in a number of ways, including health care, indoor object search, sports, job and personal life [4]. The areas where location based services find its maximum use is in locating nearby convenience stores, restaurants, traffic updates through an app, weather forecasts based on location, the social events happening in the city, step by step guide to reach a desired location, availing cab services to reach a particular destination, utilizing food applications at the comfort of your door step.

However, when accessing any location based service, for example, booking a cab, ordering food services, ordering items through e-commerce applications, the data shared by the user is either kept with the service provider for business/accounting purposes or stored in the database for future purposes such as investigation. The data held by the service provider and that stored in the database can be misused or hacked, therefore, the privacy of the user's data is a concern and needs to protected and preserved [9]. This paper addresses a scheme which protects data privacy. Two algorithms have been developed to protect data privacy and successfully implemented in mobile application domain for LBS.

2. LITERATURE SURVEY

Abdur R. Shahid, Niki Pissinou, S.S. Iyengar, Jerry Miller, Ziqian Ding, Teresita Lemus proposed that a concealing region is used to disclose the user's exact position. In this region, the locations become similar so the attacker cannot identify and is resilient against a wide range of interferences [9]. Yuwen Pu, Jin Luo, Ying Wang, Chunqiang Hu, Yan Huo Jiong Zhang proposes an Advanced Encryption Standard (AES) symmetrical encryption with one-time pad keys and IBE (Identity-based Encryption) which preserves location privacy using cryptographic approach [5]. Elena Simona Lohan, Philipp Richter, Vincente Lucas-Sabola, Jose A. Lopez-Salcedo identified the problem that user is not aware of the privacy threats in wireless communication. Therefore, TLS (Transport Layer Security) standard is proposed to ensure security and privacy of the user. In addition, technical solutions are provided to deal with indoor and outdoor localization providing privacy and data security for communications over the internet [11]. Muhammad Usman, Muhammad Rizwan Asghar, Imran Shafique Ansari, emerging technologies such as MIMO (Multiple Input and Multiple Output), NB-IoT (Narrowband IoT) which give

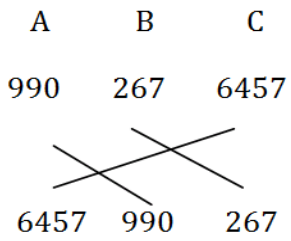
proximity information with high precision are considered and compared with the Localization Proximity (LP) technology for the smart city application. The one that provides higher security among the two is selected [12]. Anahid Basiri, Elena Simona Lohan, Terry Moore identified that random forest method can be used to solve the privacy concerns related to location data. In this method, a forest region is created with a number of tress. The more the number of trees, the higher the accuracy. Three challenges were solved using the random forest method: Quality of positioning service, concerns related to privacy and what all contents were actually made available [13]. Philip Asuquo, Haitham Cruickshank, Jeremy Morley, suggest privacy enhancing schemes and cryptography methods, for example, symmetric key authentication and asymmetric key authentication for privacy in vehicle and cellular networks [14]. Priti Jagwani and Saroj Kaushik brief about the privacy protection schemes such as location obfuscation, k-anonymity to achieve privacy protection and confidentiality while accessing location based services [2]. Huang, H. and Gao S. have taken user's location data into account and uses outdoor positioning technologies such as GNSS and indoor positioning technologies such as WLAN for location estimation and identifies the major privacy threats for which techniques such as obfuscation and k- anonymity are proposed [8]. Savy Gupta, Shagun Seth, Amit Dhawan identified that to achieve security and prevent sensitive data leakage when smartphones used by employees can be used to take pictures of confidential documents at organizations for wrong purposes; an application that makes use of LBS along with GPS is used where as soon as the employee enters the organization the security administrator disables the camera [15]. Priyanka Kumar, Raghul M have developed two applications; parent app and child app on the respective mobile devices. The parent app keeps track of the child when he is outside by himself. Two commands are used: CALL and STOP; the CALL command keeps calling the child's phone when he is lost and STOP command ensures that the calling does not stop until the parent finds the child and clicks on STOP [6]. G. Sriram B., Srikanth Reddy, K.V. Seshadri provide a Location based service (LBS) algorithm for the latitude and longitude coordinates of the location upon which cryptographic encryption algorithm is used providing enhanced security to mobile users [10]. Zhang Lei, Yu Lili, Li Jing, Meng Fanbo have proposed an algorithm to protect the historical data and current data positions of the user, the algorithm blurs the historical data and reduces the probability of matching the locations efficiently [16]. Detian Zhang, An Liu, Gaoming Jin, Qing Li. have proposed an algorithm where efficient route to speed up processing of queries for large paths that will handle large queries without requiring the path data for location-based services which is proven to be very practical in real time [17]. Mohammed Elbes, Eyad Almaita, Thamer Alrawashdeh, Tarek Kanan proposed an algorithm for indoor environments called LSTM (Long Short term Memory) Based indoor localization algorithm which takes fingerprints and produces a mean error equal to one meter which is very good with regard to location based services [18]. The literature survey clearly says that data privacy is still a major concern when location based services are being availed.

3. ALGORITHMS

Two algorithms are proposed to protect user's data privacy: **perfect shuffle algorithm** and **hash salt algorithm.** Perfect shuffle algorithm is an encryption algorithm that exchanges and shuffles the numbers or characters in a string; this algorithm is used to hide the phone numbers. The second algorithm, that is, hash+salt algorithm is a combination of **SHA 512(Secure Hash Algorithm 512)** which is a hashing algorithm that cryptographically encrypts the data by performing hashing function on it and produces an output called **message digest** and **salt** which is 11-digit random string added to SHA 512 to make the data security stronger.

3.1 Perfect Shuffle Algorithm

The **Perfect shuffle algorithm** is used to encrypt phone numbers of the users stored in the database. The phone numbers are shuffled and exchanged. For example, if the original Phone number is 9902676457. It is first split into three parts A, B and C and then exchanged



The Encrypted phone number now becomes 6457990267 and in this format the phone numbers will be stored in the database. This way even if the hacker or attacker tries to get hold of the data, he has to perform a number of permutations and combinations to find the original number. Therefore, the privacy of the user is protected.

3.2 Hash Salt Algorithm

The hash salt algorithm is used to encrypt passwords, names and any other details stored in the database. The hashing algorithm used is **SHA 512(Secure Hash Algorithm 512).** Secure Hash Algorithm belongs to a group of cryptography operations intended to secure information or any data contents. The hash function is an algorithm composed of bitwise operations and modular additions that converts the data [5] [10]. **SHA 512** is a hashing algorithm, which executes a hashing function on the information that you give. Hashing functions take some data as input and produce an output called **message digest.** But hashing algorithm alone cannot protect the data stored in the database. For example, consider an example of a user booking a cab to reach a particular destination. The user must sign in with his user ID

and password first. He then logs in with his user ID and password, and forwards it to the server through a secure connection. The server fetches the ID to check the matching message digest (that is, password). The user's password is verified and authenticated with the help of hashing algorithm which produces an output that must match with the server saved password. The server will not store, or try to display the passwords in this process. So if a hacker or attacker gets hold of the data, he can run through a hashing algorithm a set of widely used passwords, and obtain a list called a **rainbow table**. Comparing a list of hacked passwords with a rainbow table is pretty simple for a computer. There is a password that will match with everything on the table. Therefore the hashing algorithm alone has a disadvantage. The data needs to be protected and the extra step is to add a salt which is 11-digit random string to the hashing algorithm used. A random number needs to be added to all passwords so that it has that added protection. The output which is called a **message digest** would be a product of both the hashing algorithm and the salt. This product does not correspond with any data on the rainbow table. So even though a hacker takes the password and tries to crack the password using the hashing algorithm, it is useless. Therefore, a salt is applied to the hash which is 11-digit random string to make it stronger.

3.3 Data Masking

There has to be an authorized person to keep an account of the data say, for the purpose of investigation, business or accounting. For example, in the scene of crime, the investigation officer may want to look at the original data in order to solve the case. In such a scenario, the original data may have to be revealed. Therefore, a Super admin is authorized to look after the data stored in the database and provide it in the future if required. The super admin is the only authorized person to have full access of the data. But, the admins working at the service provider end need to perform their daily tasks as well; therefore, masking out technique belonging to data masking technology has been applied. **Data masking** is the method of using changed information (characters or other data) to cover original data. The primary justification for adding masking to data is to shield the information which is seen as personal recognizable data, delicate individual information, or financial information. It also needs to look true and appear consistent. For example, data displayed on terminal screens to call centre administrators in certain organisations will have automatically implemented masking technology depending on user access permissions (for example, stopping administrators from accessing credit card information during billing).

4. APPLICATION DEVELOPMENT

In order to test the effectiveness of the algorithms and masking out technique, mobility application has been considered where a user books a cab to reach a particular destination. Two mobile applications and one web application have been developed; the first application called CAB USER is run on android device and the second application called CAB DRIVER is run on android device as well; the database used is MongoDB and MongoDB compass, and a web application has been designed for the Super admin who can access the entire database.

Let's take an example of a user booking a cab and how the information will be transferred to driver app where all the data will be encrypted. The android application is created by using the languages JAVA, XML in Android Studio and the Web application is designed using Visual Studio code with Angular CLI, which is an open source web application framework.

5. RESULTS

The user's data when the service is being provided gets stored in the database at the service provider's end. The proposed algorithms have been applied to the data stored in the database; masking out technique has been applied to the admins working at the service provider's end and the following results have been obtained.

Case 1: Hacker's view

Figure 1: Hacker's view

In this case, if the hacker or attacker tries to get hold of the data stored in the database, he cannot view anything as phone number is shuffled using the perfect shuffle algorithm and password is encrypted using hash salt algorithm. In this way, the privacy of the user is protected in the database as shown in Figure 1.

Case 2: Super Admin's view

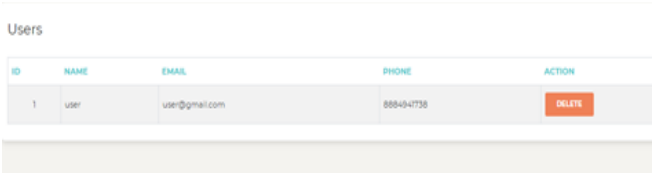


Figure 2: Super Admin's view

Super Admin can view the entire details such as the user's name, email, phone number as shown in Figure 2. He is the only authorized person to have account of the data. By keeping an account of the data, it can be used for future

purposes. The super admin can provide the data for future investigation purposes to the concerned authorities. In this manner, there is a safe account of the data while his privacy being assured.

Case 3: Admin's view

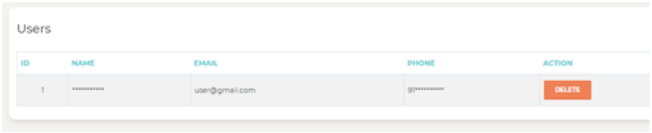


Figure 3: Admin's view

The Super Admin can add other admins who will be assigned business tasks. For e.g. as shown in Figure 3, the added admin can only view the email id and would be required to send advertisement or offer updates to the user using the email. All other user details on the web application are masked using masking out technology. In the way again the privacy of the user is protected and the effectiveness of the masking out technique is proven.

SUMMARY

From the three cases shown above and Table I, we can conclude that the user's personal data i.e. his/her privacy is protected using the combination of perfect shuffle algorithm and hash salt algorithm; and the service provider is also able to carry out his duties without the user worrying about his data since his data is masked using the masking out technique. Therefore, protecting the user's data privacy in location based services is achieved.

TABLE I. DATA BEFORE AND AFTER APPLYING ALGORITHMS

Data	Before Perfect shuffle and Hash Salt Algorithm	After Perfect Shuffle and Hash salt algorithm
Name	Svcharishma	49edbo4583boe26uita96a28
Password	Charishma*123	**Object** Hash:C3eogsb2ar2iobrea26 Salt:bjgdkeuh2rt78
Phone number	8884941739	'1739''888''494'

6. CONCLUSION

In this paper, we have proposed combination of perfect shuffle algorithm and hash salt algorithm to protect user's data privacy. The combination of the two algorithms adds an extra layer of security to the user's data stored in the database. An account of the data needs to be kept, therefore, a super admin has been authorized, yet, daily business and accounting tasks need to be carried out, so, by masking out technique, data is masked allowing the admins working for the service provider to perform their daily assigned business and accounting tasks. In this way, the user's data privacy is protected at the service provider end.

REFERENCES

[1] Haosheng Huang, Georg Gartner, Jukka M. Krisp, Martin Raubal & Nico Van de Weghe (2018) Location based services: ongoing evolution and research agenda, Journal of Location Based Services, 12:2, 63-93, DOI: 10.1080/17489725.2018.1508763.

[2] Priti Jagwani and Saroj Kaushik. "Privacy in Location Based Services: Protection Strategies, Attack Models and Open Challenges," 2017.

[3] Taha S, Shen X, Fake point location privacy scheme for mobile public hotspots in NEMO-based VANETs. In Proc. IEEE ICC, Budapest, Hungary. 2013;630-634.

[4] Haosheng Huang and Georg Gartner. "Current Trends and Challenges in Location-Based Services," 2018.

[5] Yuwen Pu, Jin Luo, Ying Wang, Chunqiang Hu, Yan Huo Jiong Zhang. "POSTER: Privacy Preserving Scheme for Location Based Services Using Cryptographic Approach," 2018.

[6] Priyanka Kumar, Raghul M. "Location Based Parental Control-Child Tracking App using Android Mobile Operating System," 2018.

[7] Gang Sun, Shuai Cai ; Hongfang Yu ; Sabita Maharjan ; Victor Chang ; Xiaojiang Du ; Mohsen Guizani "Location Privacy Preservation for mobile users in Location-based services" 2019.

[8] Huang, H. and Gao, S. (2018). Location-Based Services. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2018 Edition), John P. Wilson (Ed). DOI: 10.22224/gistbok/2018.1.14.

[9] Abdur R. Shahid, Niki Pissinou, S.S. Iyengar, Jerry Miller, Ziqian Ding, Teresita Lemus. "KLAP for Real-World Protection of Location Privacy," in IEEE World Congress on Services 2018.

[10] G. Sriram B., Srikanth Reddy, K.V. Seshadri, K. Hemantha Kumar. "Location Based Encryption-Decryption System for Android," 2018.

[11] Elena Simona Lohan, Philipp Richter, Vincente Lucas-Sabola, Jose A. Lopez-Salcedo,Gonzalo Seco-Granados, Helena Leppakoski, Elena Serna Santiago. "Location Privacy Challenges and Solutions: GNSS Localization," 2019.

[12] Muhammad Usman, Muhammad Rizwan Asghar, Imran Shafique Ansari, Fabrizio Granelli, and Khalid A. Qaraqe. "Technologies and Solutions for Location-Based Services in Smart Cities: Past, Present, and Future," 2018.

[13] Anahid Basiri, Elena Simona Lohan, Terry Moore, A. Winstanley. "Indoor location based services challenges, requirements and usability of current solutions," 2017.

[14] Philip Asuquo , Haitham Cruickshank, Jeremy Morley, Chibueze P. Anyigor Ogah, Ao Lei , Waleed Hathal, Shihan Bao, and Zhili Sun. "Security and Privacy in Location-Based Services for Vehicular and Mobile Communications: An Overview Challenges and Countermeasures," 2018.

[15] Savy Gupta, Shagun Seth, Amit Dhawan, Abhishek Singhal. "Location Based Camera Disable System on Android Platform," 2018.

[16] Zhang Lei, Yu Lili, Li Jing, Meng Fanbo. "Location Privacy Protection Algorithm Based on Correlation Coefficient," 2018.

[17] Detian Zhang, An Liu, Gaoming Jin, Qing Li. "Edge-Based Shortest Path Caching for Location-Based Services," 2019.

[18] Mohammed Elbes, Eyad Almaita, Thamer Alrawashdeh, Tarek Kanan, Shadi AlZu'bi, Bilal Hawashin. "An Indoor Localization Approach Based on Deep Learning for Indoor Location-Based Services," 2019.