Multi-Obfuscation Approach for Preserving Privacy in Smart Transportation

Sami S. Albouq[1], Adnan Ani Sen[1], Nabile Almoshfi[2], Mohammad Bin Sedeq[1], Nour Bahbouth
salbouq1@iu.edu.sa ,adnanmnm@iu.edu.sa, nalmashfi@ju.edu.sa, *m.binsadiq@iu.edu.sa*

[1]Department of Computer Science, Islamic University of Madinah, Madinag, Saudi Arabia
[2]Department of Computer Science, Aljouf University, Skakka, Saudi Arabia

Summary

These days, protecting location privacy has become essential and really challenging, especially protecting it from smart applications and services that rely on Location-Based Services (LBS). As the technology and the services that are based on it are developed, the capability and the experience of the attackers are increased. Therefore, the traditional protection ways cannot be enough and are unable to fully ensure and preserve privacy. Previously, a hybrid approach to privacy has been introduced. It used an obfuscation technique, called Double-Obfuscation Approach (DOA), to improve the privacy level. However, this approach has some weaknesses. The most important ones are the fog nodes that have been overloaded due to the number of communications. It is also unable to prevent the Tracking and Identification attacks in the Mix-Zone technique. For these reasons, this paper introduces a developed and enhanced approach, called Multi-Obfuscation Approach (MOA that mainly depends on the communication between neighboring fog nodes to overcome the drawbacks of the previous approach. As a result, this will increase the resistance to new kinds of attacks and enhance processing. Meanwhile, this approach will increase the level of the users' privacy and their locations protection. To do so, a big enough memory is needed on the users' sides, which already is available these days on their devices. The simulation and the comparison prove that the new approach (MOA) exceeds the DOA in many Standards for privacy protection approaches.

Keywords:
Connected, Vehicles, Privacy, Standard, Communication, Cloud, Computing, Fog, Interworking, Obfuscation

1. Introduction

The world has been witnessing a tremendous amount of development in every field, Transportation, Engineering, Health, and all others. Technology is an essential factor in the quick development fields. For example, Internet of Things (IoT) has emerged and become an important element that integrated with all mentioned fields to either collect or process data. Therefore, IoT can be seen on the road, in business, hospitals, cities, and even our houses. There are billions of intelligent things and devices like smart TV, smartphones, smart cars, smartwatch, etc., either inside or outside our homes. These devices have a unique identity to connect to the internet and share and interact with each other, which is known as Machine-to-Machine relation (M2M) [1][2].

IoT has changed our world and our life to a more sophisticated and adaptive. IoT enables each device to have a unique ID to use it during communications. The device feels, senses their surrounding conditions like heat, pressure, movement, voice, etc., and converts them to data that is provided to users based on their requests. This data is collected or stored to provide users with services such as navigations, shopping, or remote-health-monitoring. However, collecting and analyzing users' data may reveal sensitive information that best to kept secret to maintain a high level of privacy.

The aim of IoT is to equip every object with sensors or RFID Tag, connect to the internet, sending/receive data to/from other devices, users, applications, or services providers to support intelligent decisions. Therefore, IoT objects have two main parts: Wireless Sensor Networks (WSNs) and Radio Frequency Identification (RFID) as shown in Figure 1 [3]. Since IoT objects have limited resources (i.e., storage), they connect to the internet to send processed or collected data to the cloud for temporal or permanent storage. The cloud processes received data and provided IoT object with requested services enable them to complete their tasks. RFID is an old technology used in the Soviet Union in 1945. This technology utilizes radio waves to send data. It consists of Tag and Reader, so the tag is attached to the object that needs to be identified. The Identity can be utilized for tracking objects. The reader is sending data between tags and cloud or fog [2]. WSNs works on the network layer and is responsible for sending data to applications through the internet. The most common protocols for sending and receiving data are ZigBee, Bluetooth, and Wi-Fi. ZigBee and Bluetooth consume low power than Wi-Fi, so the sensors last longer and cover different ranges between 10-70 meters [4][5]. The main job of WSNs is collecting data and sending it to the cloud by the default gateway/sink.

RFID and WSNs do not have enough processing, power, and memory where they can process data. Therefore, they depend on cloud and fog computing. In order to get a faster response than cloud and Fog computing to cooperate together [6]. Fog is a middle layer between the cloud and

Manuscript received April 5, 2023
Manuscript revised April 20, 2023
https://doi.org/10.22937/IJCSNS.2023.23.4.18

the IoT objects, which provides a real-time response in the case of an emergency. Moreover, the fog node can apply some computing and processing on the data before sending it to the cloud, which enhances the performance and reduces the number of connections to the cloud. As a result, the fog node reduces the network traffic and reduces the overload in the cloud [7][9].

Figure 1. Smart Environment with IoT Objects

With all these advancements in technologies and smart services, there is a huge threat related to users' privacy and security data. Users' data is the main factor on those smart services depend. Therefore, smart devices and sensors permanently monitor all aspects and details of the user's life everywhere. Consequently, a real threat lies to the user in revealing many sensitive and private data to malicious parties that may exploit it by putting users in danger. The user may threaten his/her life sometimes [10][22].

For example, most of the services that were provided in the smart transportation field depended on the location feature, which required the user to share his/her location with service providers on an ongoing basis. The user's location data could be collected by a malicious service provider or an external attacker who could have stolen this data. The leaking data could reveal everything about the user, such as personal information. It could contain answers to the following questions [10][22][23]:

- What is the workplace and nature?
- Where does he/she live?
- When is he/she at home and when he/she is outside?
- What is his/her social status and whether he/she is Rich or not?
- Does he/she frequently go to restaurants, stay in hotels, and go shopping for brands?
- Does he/she suffer from chronic diseases that require periodic visits to health centers?
- Does he/she have children at school?

Therefore, many countries have started implementing laws to protect the privacy and have asked service providers to accept these laws, such as the European GDPR, Saudi law, and others [24]. However, these laws were not and will not be sufficient despite their importance. We still need additional privacy protection techniques to ensure users' information security. It is generally difficult to maintain a high level of privacy by relying on a third trusted party to facilitate communications between two parties. This scenario puts users in insecure communications as they are required to send confidential information to the third party for communicating with others. However, any technique involving a third trusted party may expose to insecure operations such as tracing or information leakage.

2. Related Work

In this section, we highlight related work on techniques for location privacy protection. To protect users' privacy, there are many approaches each of which has its own methods to maintain three types of privacy (identity, query, and location) that concluded to two major types of approaches: regular and hybrid.

2.1 Regular Approaches

Anonymity [11][12]: means that users' identities or their data cover up using a nickname or fake name. In this way, there is no connection between users and their data, which is known as "linking Profiling". This approach is simple to implement. It protects users' privacy. Users do not have to trust others in terms of giving access permission to their data.

Mix-Zone [13][14]: It is similar to Anonymity in which both of them rely on nicknames or pseudonyms. Even so, Mix-Zone relates the region where users are located. This region is portioned into small parts called zones. Each user in this zone has his/her nickname. If they leave their zones to enter other ones, their nicknames are changed to the new ones. It is a good approach for protecting privacy.

The disadvantage of the two previous approaches is that the users' privacy can be violated by tracking their IP addresses.

Dummy [15][16]: This type is based on anonymity. It tries to hide the user's information by generating new wrong information (Dummy information) about the user location, or queries. This approach is used in LBS. The dummy approach generates the wrong location for each query to hide the actual one for the user in the whole set of queries. The benefit of this approach is protecting users' privacy. The side effect of this approach is generating a location for a user where it is impossible to reach. It could be on the water or a mountain. So, attackers can easily exclude them.

Trusted Third Party (TTP) [17]: Here data is not sent directly to the cloud. The exact location of the user is still unknown because this approach depends on Anonymizer. The region is controlled by Anonymizer called Cloaking Region. There are two strategies in this approach. The first one is very simple, where users send their queries to TTP instead of SP, and TTP forwards them to SP instead of them. The other one is that all users in the same cloaking area send their data to one Anonymizer (Small TTP) and then it sends data on behalf of them. The advantage of this approach (TTP) is to hide the users' identity which, means protecting their privacy. The disadvantage is that the users change their trust from the service provider to a third party which is the Anonymizer. As a consequence, their data can be endangered. In addition to that, if an anonymizer sends and receives requests of all users, it will lead to traffic congestion in the network. Therefore, a loss of data can happen.

Obfuscation [18]: The exact location of users is changed with a known place (landmarks) or covered in an area based on the mechanism used in graph adjusting. This approach keeps some data of the users unspecific. As with other approaches, it protects the users' privacy, but it provides an imprecise result and requires more power to manage the result when using a large obfuscation area.

Caching [19]: When users request the location of a place, the data is cached to be used again in the future. The main advantage of this approach is minimizing the communication to the server to get the request's answer. In addition, it protects the users' privacy. The drawback of this approach is the users need to connect to the server in case of a new request. Usually, this approach can integrate with others.

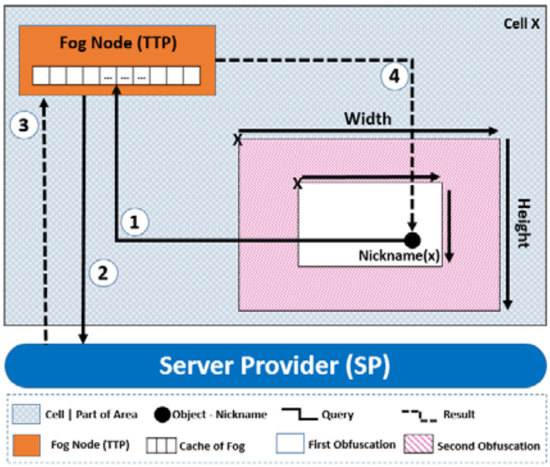


Figure 2. Double Obfuscation approach DOA (Previous approach)

2.2 Hybrid Approaches

Double Cache Approach (DCA) [20]: The main point here depends on two caches to keep nodes' information hidden from the service provider. A user sends a request to cache 1 If he doesn't find the answer for his query in cache 2. Then any other user (Cooperator) can read the request and send it to SP instead of the main user (owner). When SP returns the answer to the query, it is inserted in Cache 2. The main user will check after a few seconds again in cache 2. In this way the identity of the user is unknown, and SP can't trace it back.

The disadvantages of this approach are how to protect the data in the cache and how to manage the cache.

Swap Obfuscation Approach (SOA) [21]: It is a cooperation between two nodes to protect their privacy from the SP and each other, too. The basic concept here is the first node sends a request to the second one with fuzzy information about itself. Then, the second node will send this request to the SP. In this way, the first node protects itself from the second node with unclear information, and from the SP where there is no direct communication. Moreover, the second node will enhance its privacy by sending this query because it is like a dummy in this case. This approach has issues in performance which is resulted by swapping between the nodes.

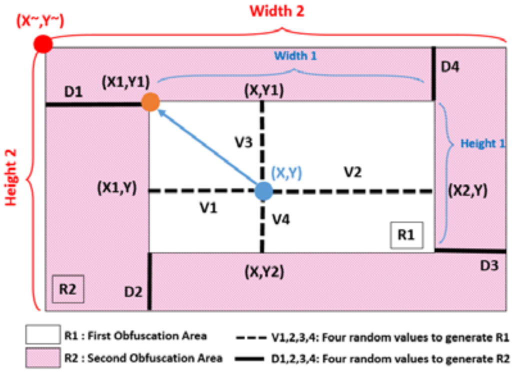

(X~,Y~)
Width 2
Width 1
D4
D1
(X1,Y1)
(X,Y1)
V3
Height 2
Height 1
(X,Y)
V2
(X1,Y)
(X2,Y)
V1
V4
R1
(X,Y2)
D3
R2
D2
R1 : First Obfuscation Area
R2 : Second Obfuscation Area
V1,2,3,4: Four random values to generate R1
D1,2,3,4: Four random values to generate R2

Figure 3. Processing Method of Result in DOA

Double Obfuscation Approach (DOA) [1]: this is a hybrid approach. It mixes four techniques which they are: obfuscation, mix-zone, TTP, and cache. TTP is presented by fog nodes, so users connect them to ask for a Point of Interest (POI). Furthermore, fog nodes have a cache to store historic requests, so the caching technique is used here. If fog nodes have the answers, they returned them to the users, otherwise, they send them to the cloud (SP). The users don't use their real location, instead, they use near a location in an area where one fog node controls. This is the first obfuscation. The second one is when fog nodes send a request to SP, so fog nodes double the obfuscation for more privacy. Fog nodes also divide the area into five zones to enable a user to select the zone of results that is suitable for his current location.

3 Approach

This section addresses the DOA shortages and how MOA eliminates them.

The main shortages of DOA that is shown in Figure 2 can be summarized as follows:

- DOA is powerless to prevent Tracking attacks of the users' point-of-interest, and this risk is increased especially in similar type areas. For example, the area is famed for restaurants, medicals, or others. This will discover extra information about the users such as the query types.

This section describes our enhanced proposed solution that relies on Multi-Obfuscations Approach (MOA). The proposed solution presents an idea that is inherited from the DOA approach to enhance DOA and overcome all shortages that mentioned previously. In this paper, we introduce a different method for selecting several simple obfuscation areas (cells) to help cooperation between adjacent fog nodes and improve the level of privacy of their areas (fog nodes) and reduces the overload of them.

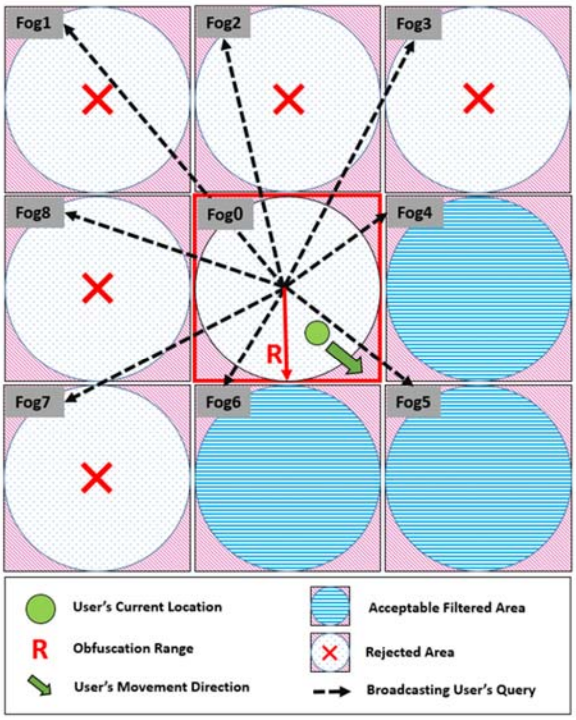


Figure 4. MOA based Fog Computing

3.1 How the Proposed Approach Works

The user does not need to send his/her current location or his/her obfuscation area. It is enough to communicate with the fog node that he/ she resides in by sending the query to it. The fog node (initiator) will then send a Broadcast to a group of cooperative fog nodes surrounding it (nearby). Each fog node will collect all its formed queries and send them all at once to the service provider. This will create K-Anonymity for the queries and therefore a higher level of security is met. Each fog node will return the query result to the fog node (initiator) requesting the query. After that, the fog node will collect the results and send them to the end user based on a fixed sequence. The user will finally select the results for the cells closest to their new location and ignore the other cells. The user can keep some results in their cache for later use to reduce the number of connections and improve performance.

4. Simulation and Results

To evaluate the proposed approach, a simulation has been implemented. It is similar to that in [1] using the Visual Studio.Net platform, and the same conditions and assumptions mentioned in the previous research were applied. The comparison was made based on several criteria:

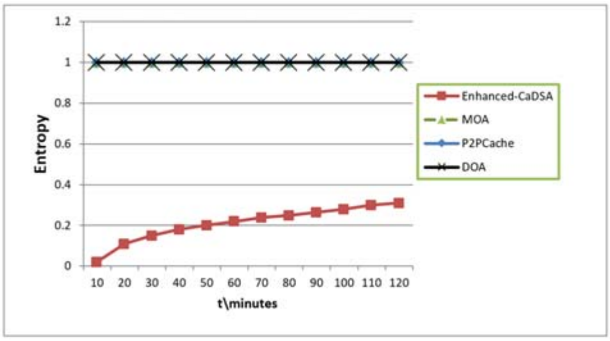


Figure 5. Entropy Comparison (Privacy Level)

Entropy is the correct amount of data that a malicious service provider or an external attacker can collect from the user. The entropy represents the level of protection achieved in terms of the extent to which the attacker is certain that the information he has about the user is correct. The value of entropy is between 0 and 1, and the higher it is, the higher the level of protection.

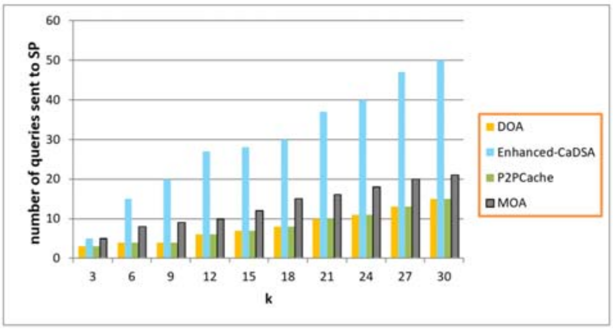
Performance is measured by the speed or response time of queries in addition to the number of queries sent to the service provider. This criterion is positively affected by the adoption of protection technology on Cache to reduce the number of calls, where it is negatively affected by the processing performed. The processing must be implemented on the results or queries before sending.

There are also non-quantitative criteria associated with: Anonymity (hiding the identity): The MOA achieves absolute protection of the user's identity from the service provider as there is no direct connection between the service provider and the user.

Accuracy (maintain accurate results): The MOA has found an effective solution to this open problem in noise-based techniques by obtaining additional results for neighboring nodes. Therefore, the users can select the nodes that are closest to their current location without affecting their privacy. Thus, MOA supports dynamic queries greatly.

Need Trust (the need for trust in a third party): The user doesn't need to trust anyone here, not even the fog node, where he just sends his query without his location or identity.

Attack Resistant: The MOA, compared to the previous approach, protects against two additional types of privacy attacks, namely, the zone tracking attack and the zone profiling attack. As a result, the service provider will not be able to plot a path for the user as he moves between different regions because the user's query originates from different regions. Also, he will not be able to standardize the regions with a specific query pattern (for example, medical, entertainment) due to the participation of each region in the inquiries of its neighboring regions.



Regarding quantitative criteria, first, the entropy criterion, Figure 5 shows the superiority of the proposed approach by achieving the highest level of entropy, similar to the DOA approach and the P2PCache approach. It is because all of this technology does not communicate directly with the service provider, but the fog node does it instead. Therefore, the service provider does not have any confirmed information about the user who requested the query. While in Enhanced-CaDSA technology, the user himself sends his real query with a set of K Dummies to the service provider. The service provider begins to form some real information about the user by linking and comparing his queries, and it can be said that the MOA achieves a very high level of protection from the service provider.

3.1 The Advantages of MOA Compare DOA

- The user will not need to request a new connection or query when moving to a new cell and this solves the problem of the Tracking attack.

- Each fog node can send all the queries at once and most of the queries will be Dummy thus, misinformation about each region is given to the malicious service provider. This will improve the privacy of the region.

- The improved approach provides better results accuracy by merging the results of more than one adjacent cell based on the new user's location and thus greater support for dynamic queries.

- Reducing the overload on the fog node, which will no longer need to process the results for each region.

- Better user privacy, whether from the fog node itself, the server, or the external attacker.

- Reducing the headache of generating an obfuscation zone for the user.

However, MOA requires more user resources related to cache size to be able to store results from more than one cell to take advantage of this approach effectively. Also, the Relating to performance, as we discussed, there are two quantitative criteria. The first one is related to the number of queries sent to the service provider. It is notable that MOA provides an average rate (Figure 6) compared to other methods. MOA depends on a group of fog nodes and not a single node and therefore each node will send a query to the service provider. Thus, five queries, for example, maybe sent, instead of one query. This means better protection and avoidance of zonal tracking attacks by query or zonal profiling attacks to detect the specificity of query type. At the same time, it will not negatively affect performance here because each fog node will be responsible for one query. Also, each fog node has a cache that will significantly reduce the number of queries to be sent to the cloud. In this way, it makes increasing the number of queries virtually ineffective on system performance in return for a significant improvement in the level of protection for the user and for the areas of the fog nodes themselves.

As for the processing time of the query, the improved MOA approach is better than the previous DOA approach on the user side because it does not require the user to process any of the queries or generate a small private noise area. It also does not require from the central fog nodes any processing to divide the results into regions that are divided automatically by each fog node. This explains the results shown in Figure 7

5 CONCLUSIONS

This research presented an enhanced approach for preserving privacy in IoT applications (Transportation domain), which is called MOA. MOA addressed all drawbacks of DOA, which are related to performance, the accuracy of results with dynamic queries, and the need for semi-trust to fog nodes. In this work, we designed the new approach and explained how it works besides all its advantages. In the next work, we will implement the new approach and compare it to other privacy approaches to prove its superiority according to the level of privacy and accuracy of results without significant effect on the performance.

Acknowledgments

This work has been supported by Islamic University of Madina grant Takamul-109. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of supported by the Islamic University in Medina or other research sponsors.

References

[1] S. S. Albouq, A. A. A. Sen, A. Namoun, N. M. Bahbouh, A. B. Alkhodre, and A. Alshanqiti, "A Double Obfuscation Approach for Protecting the Privacy of IoT Location Based Applications," IEEE Access, vol. 8, pp. 129415–129431, 2020, doi: 10.1109/ACCESS.2020.3009200.

[2] X. Jia, Q. Feng, T. Fan, and Q. Lei, "RFID technology and its applications in Internet of Things (IoT)," in 2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet), Apr. 2012, pp. 1282–1285. doi: 10.1109/CECNet.2012.6201508.

[3] Kandris, D., Nakas, C., Vomvas, D., & Koulouras, G. (2020). Applications of wireless sensor networks: an up-to-date survey. Applied System Innovation, 3(1), 14.

[4] Y. Chen and Q. Zhao, "On the lifetime of wireless sensor networks," IEEE Commun. Lett., vol. 9, no. 11, pp. 976–978, Nov. 2005, doi: 10.1109/LCOMM.2005.11010.

[5] [C. M. Ramya, M. Shanmugaraj, and R. Prabakaran, "Study on ZigBee technology," in 2011 3rd International Conference on Electronics Computer Technology, Apr. 2011, vol. 6, pp. 297–301. doi: 10.1109/ICECTECH.2011.5942102.

[6] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications," IEEE Internet Things J., vol. 4, no. 5, pp. 1125–1142, Oct. 2017, doi: 10.1109/JIOT.2017.2683200.

[7] T. Dillon, C. Wu, and E. Chang, "Cloud Computing: Issues and Challenges," in 2010 24th IEEE International Conference on Advanced Information Networking and Applications, Apr. 2010, pp. 27–33. doi: 10.1109/AINA.2010.187.

[8] A. Khan, "What is the relationship between Cloud Computing and Service Orientated Architecture (SOA)?," Welcome to Khan's Blog, Jul. 13, 2015. https://arsalankhan.com/2015/07/13/relationship-between-cloud-computing-and-soa/ (accessed Nov. 27, 2021).

[9] "Fog Computing vs. Cloud Computing: Key Differences | SaM Solutions," Sep. 10, 2019. https://www.sam-solutions.com:443/blog/fog-computing-vs-cloud-computing-for-iot-projects/ (accessed Nov. 27, 2021).

[10] A. A. A. Sen and A. M. Basahel, "A Comparative Study between Security and Privacy," in 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom), Mar. 2019, pp. 1282–1286.

[11] C. Bettini, S. Jajodia, and L. Pareschi, "Anonymity and Diversity in LBS: A Preliminary Investigation," in Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops (PerComW'07), Mar. 2007, pp. 577–580. doi: 10.1109/PERCOMW.2007.23.

[12] E. Elabd, "Enhanced peer-to-peer anonymity approach for privacy preserving in location-based services," J. Locat. Based Serv., vol. 14, no. 4, pp. 252–267, Oct. 2020, doi: 10.1080/17489725.2020.1844326.

[13] B. Palanisamy and L. Liu, "MobiMix: Protecting location privacy with mix-zones over road networks," in 2011 IEEE 27th International Conference on Data Engineering, Apr. 2011, pp. 494–505. doi: 10.1109/ICDE.2011.5767898.

[14] N. Guo, L. Ma, and T. Gao, "Independent Mix Zone for Location Privacy in Vehicular Networks," IEEE Access, vol. 6, pp. 16842–16850, 2018, doi: 10.1109/ACCESS.2018.2800907.

[15] H. Kido, Y. Yanagisawa, and T. Satoh, "An anonymous communication technique using dummies for location-based services," in ICPS '05. Proceedings. International Conference on Pervasive Services, 2005., Jul. 2005, pp. 88–97. doi: 10.1109/PERSER.2005.1506394.

[16] B. Niu, Q. Li, X. Zhu, G. Cao, and H. Li, "Achieving k-anonymity in privacy-aware location-based services," in IEEE INFOCOM 2014 - IEEE Conference on Computer Communications, Apr. 2014, pp. 754–762. doi: 10.1109/INFOCOM.2014.6848002.

[17] P. Jagwani and S. Kaushik, "Secure Cloaking Area Based on User Profile Similarity," Int. J. Eng. Technol., vol. 8, no. 6, pp. 458–461, 2016, doi: 10.7763/IJET.2016.V8.933.

[18] M. L. Damiani, E. Bertino, and C. Silvestri, "Protecting Location Privacy through Semantics-aware Obfuscation Techniques," in Trust Management II, vol. 263, Y. Karabulut, J. Mitchell, P. Herrmann, and C. D. Jensen, Eds. Boston, MA: Springer US, 2008, pp. 231–245. doi: 10.1007/978-0-387-09428-1_15.

[19] M. Yamin and A. Abi Sen, "Improving Privacy and Security of User Data in Location Based Services," Int. J. Ambient Comput. Intell., vol. 9, pp. 19–42, Jan. 2018, doi: 10.4018/IJACI.2018010102.

[20] A. A. A. Sen, F. B. Eassa, M. Yamin, and K. Jambi, "Double Cache Approach with Wireless Technology for Preserving User Privacy," Wirel. Commun. Mob. Comput., vol. 2018, p. e4607464, Aug. 2018, doi: 10.1155/2018/4607464.

[21] Y. Alsaawy, A. Alkhodre, A. Ahmed, A. Abi Sen, and M. S. Siddiqui, "Swap Obfuscation Technique for Preserving Privacy of LBS," Jun. 2019.Nguyen B Truong, Gyu Myoung Lee, and Yacine Ghamri-Doudane. Software defined networking-based vehicular adhoc network with fog computing. In 2015 IFIP/IEEE International Symposium on Integrated Network Management (IM), pages 1202–1207. Ieee, 2015.

[22] Tawalbeh, L. A., Muheidat, F., Tawalbeh, M., & Quwaider, M. (2020). IoT Privacy and security: Challenges and solutions. Applied Sciences, 10(12), 4102.23]]

[23] Sen, A., Ahmed, A., Eassa, F. A., Jambi, K., & Yamin, M. (2018). Preserving privacy in internet of things: a survey. International Journal of Information Technology, 10(2), 189-200.24]]

[24] Zaeem, R. N., & Barber, K. S. (2020). The effect of the GDPR on privacy policies: Recent progress and future promise. ACM Transactions on Management Information Systems (TMIS), 12(1), 1-20.