2014 1st International Congress on Computer, Electronics, Electrical, and Communication Engineering (ICCEECE2014)
IPCSIT vol. 59 (2014) © (2014) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2014.V59.7

Enhanced Privacy Framework for Privacy in Moving Objects

Divya.C [1], Harini Shankar [1], B.A.Sabarish [1+] (Asst. Prof)

[1] Department of Information Technology, Amrita School of Engineering, Ettimadai.
dividivya.cd@gmail.com, harini2905@gmail.com, sabarishpm@gmail.com

Abstract. Advances in wireless communication technology like mobile phones and GPS enabled devices have increased the capability to locate a person and predict their behavioral movements. Retrieval of such personal information is a threat to security which may be used for a variety of intrusive or malicious purposes. A location-based service (LBS), is any information, entertainment, or social media service that is available on a mobile device, and makes use of geographical position. The Location Based Services (LBS) request user's location information to choose the nearby cell to serve the user's request which in turn causes a threat of disclosing location privacy. Under such situations a framework is essential to preserve the location privacy of moving objects. Many frameworks proposed so far focuses on resolving a specific kind of attack, which doesn't provide a complete privacy. So this paper proposes a framework that combines the existing frameworks to ensure privacy accompanied with good accuracy and efficiency. The major attack that poses threat to the privacy of moving objects are user identification attack, sensitive location tracking attack, sequential tracking attack and source identification attack. A framework proposed in this paper for achieving better privacy is Enhanced Privacy Framework (EPF) which is an integration of HERMES framework and Statistical framework that protects against the major attacks assuring location privacy. EPF is designed to provide better security against the major attacks like user identification attack, sensitive location tracking attack, sequential tracking attack and source identification attack.

Keywords. moving object database, privacy, anonymity.

1 Introduction

Spatial information describes the physical location of objects. It is very important in research and computing applications. The integration of digital communication technology based on mobile devices and networks, accompanied with the growth of internet services, offer services that demand the location information about users. Such services are referred to as Location Based Services (LBS). LBS integrates a mobile device's location with other data hence providing value to a user. The benefits of location based services are utilized by a user at the expense of some private information such as the identity of the user and the current location. The rapid growth in mobile devices and such LBS has resulted in Moving Object Databases (MOD), a rich repository of moving object data. This MOD reveals the pattern of user trajectory, current location information and so on by applying the mining operations. Although this information is very useful in areas of research, it poses privacy threats to users. Precise and correct location of a user in real time can be revealed by the moving object data to third party servers which might be untrusted thereby resulting in a privacy concern. Apart from traces about location, study of movement patterns of a specific user causes the sensitive information to be revealed such as the user's home, workplace and can even identify health related issues. To safeguard against privacy and security risks, lots of data anonymization methods, protocols used for securely querying location based information and also cryptographic techniques for sharing location based information in social networks. Apart from all such techniques, frameworks provide a unique solution to the privacy threats on the moving object databases. The paper proposes Enhanced Privacy Framework (EPF) which is very efficient, scalable and ensures perfect security.

2 Analysis of Framework

Though the enhancements in mobile technology and mobility data offers numerous benefits to users in the field of research and mining, there are many security issues faced by mobile users that might leak sensitive information and patterns resulting in various attacks. Privacy Aware Monitoring framework (PAM) [1] protects against the Spatio temporal correlation inference attack by not disclosing the genuine point locations. It uses the concept of “bounding box” and “safe region”. Trajectory Privacy preserving framework (TrPF) [2] focuses on ensuring user’s trajectory privacy. Adversary can analyse the trajectories which might provide rich spatio temporal history of information. Therefore, it is mandatory to unlink the user’s identities from sensitive information collection locations. This is done using TrPF framework thereby ensuring good security. Generic framework [3] ensures to prevent the unauthorized access of data by providing an increased level of authentication using the three factors namely: password, smart card, biometrics. Statistical framework [4] concentrates on providing source anonymity by ensuring that unauthorized users cannot identify the origin of events by analyzing the network traffic. The framework incorporates the concept of “interval indistinguishability”. Spatio Temporal attacks and Participatory Sensing attacks are together addressed and resolved using HERMES framework [5]. HERMES framework is the combination of two other frameworks namely:-PAM [1] and TrPF [2]. Private-HERMES incorporates HERMES [14], a query engine based on a powerful query language for trajectory databases, which enables the support of aggregative queries. Private Information Retrieval is a technique through which data item or record can be extracted from a database while hiding the identity of the item from the database server, hence providing a better privacy.

Table 1.

Models	Attacks Resolved
Statistical framework	Source identification attacks.
Privacy Aware Monitoring framework	Spatio temporal correlation inference attacks.
Unified framework	Disclosure attacks.
Trajectory Privacy Preserving framework	Participatory sensing.
Generic framework	Authentication related attacks.
Vehicle to Grid framework	Cyber-attacks.
Hermes framework	User identification attack, Sequential tracking attack &Sensitive location attack.
Private Information retrieval	Information becomes insecure.

3 Motivation

The motivation of this paper is to achieve good security and thereby provide an expected level of privacy to users by avoiding almost all the common attacks. Most of the frameworks proposed so far focuses on resolving either one or the other issue. Even though each framework is efficient in their own way thereby resolving the attack it is intended for, those frameworks don’t concentrate on providing a maximum degree of privacy to users by overcoming all the common attacks as a whole. This enhanced privacy can be achieved by integrating the possible frameworks efficiently and hence resolve many of the attacks. So EPF is proposed with the intention of overcoming the attacks identified as common.

4 Fundamentals of EPF Framework

Enhanced Privacy Framework aims at achieving maximum security and privacy with the integration of various technologies. It tries to resolve almost many of the most common attacks by providing a very secure architecture which integrates Privacy Aware Monitoring Framework (PAM) [1], HERMES framework [5] and Statistical Framework [4].

EPF has a three tier architecture as shown in Fig.1. It follows a bottom-up approach. As a first level the statistical framework is applied to the incoming Moving Object Datasets being queried. Statistical framework injects fake transmissions, hence unable to distinguish between the real and the fake traffic. The framework incorporates the problem of Statistical Source Anonymity (SSA) that exists in sensor networks. SSA prevents the adversaries from identifying the source location using statistical analysis on node transmissions [6]–[12]. The next level of framework is the HERMES [5]. It consists of HERMES [13] which is a trajectory query engine and HERMES++ [14], a powerful query engine for preserving privacy. The query engine is based on any query language. HERMES++ allows access to trajectory database with restrictions to avoid attacks. This framework allows the user to interact with a Graphic User Interface (GUI) accompanied with 3D rendering capabilities in Java and dependent on the Swing GUI widget toolkit [15]. The results from the program supported operations are visualized in the 3D globe provided by NASA World Wind [16]. HERMES uses two anonymization algorithms namely NWA [17] and W4M [18] to obtain the anonymized trajectory. HERMES also allows the fake and anonymized trajectory to be evaluated using the algorithms of mobility data mining such as TRACLUS [19], T-Optics [20] and CenTR-I-FCM [21], K-medoids [22] and Bisecting K-medoids [23] to take care that adding fake trajectories and disturbing the original trajectory should not alter the patterns hidden in the original MOD. The third level is the Privacy Aware Monitoring Framework (PAM) [1] which concentrates on providing efficiency, privacy, accuracy. The genuine point location of the object is encapsulated in a "bounding box" [1]. PAM uses the concept of "safe region" [1] which reduces the number of updates thereby improving the efficiency by updating the location only when the centroid of the object's bounding box moves out of the safe region. Fig.1. represents the architecture of EPF framework.

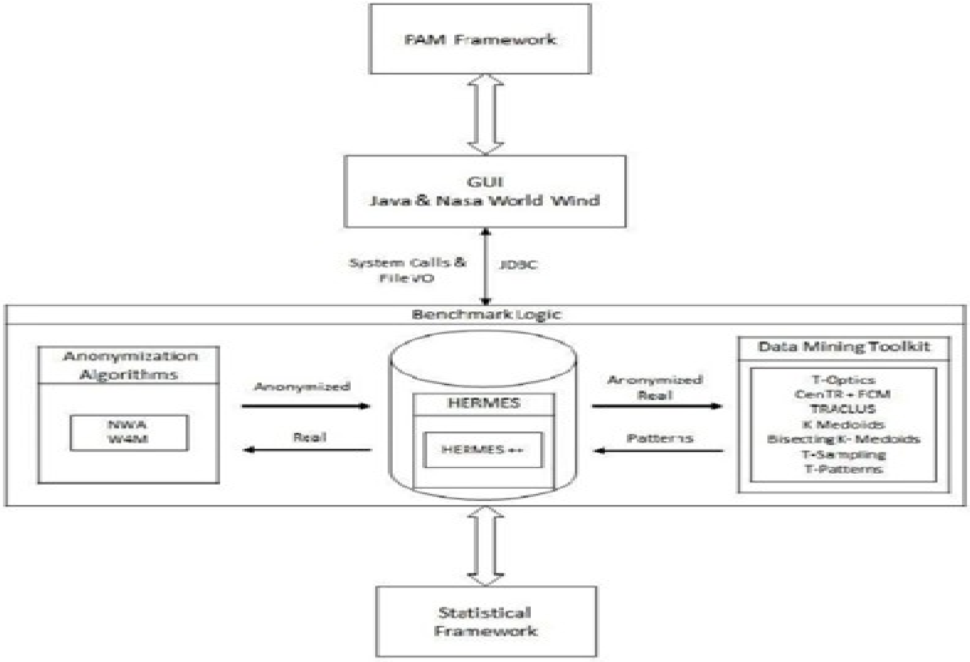

Fig. 1: Architecture Diagram.

Frameworks **→**	**Unified framework**	**Trajectory Privacy Preserving framework**	**Generic framework**	**Vehicle to Grid framework**	**Private Information retrieval**	**EPF framework**
Common Attacks ↓						
Source Identification Attacks				✓		✓
User identification attack, Sequential tracking attack & Sensitive location attack.	✓	✓	✓		✓	✓
Spatio temporal correlation inference attacks.		✓				✓

5 Related Work

Source location privacy in sensor networks is part of a broader area, the design of anonymous communication systems. The foundation for this field was laid by Chaum in [24]. Topics related to location anonymity have been discussed by Reed et al. in [25], who introduced the idea of preserving anonymity through onion routing, and by Gruteser and Grunwald in [26]. Here, privacy is maintained through anonymous routing. In [27], Xi et al. proposed a new random walk routing method. Path confusion has also been proposed as an anonymity-preserving routing scheme by Hoh and Gruteser in [28].

The scheme of [29] was recommended to address this latency/overhead trade off. In [29], Shao et al. introduced the notion of statistically strong source anonymity. More specifically, after the transmission of every fake event, the node draws an exponentially distributed random variable t ~ Exp (λ), where λ is the pre-specified rate of the exponential distribution. The node then waits for t time units and then transmits another fake event. In Privacy Aware Monitoring framework, research work mainly focuses on spatial temporal query processing. Supposing that object movement trajectories are known apriori, Saltenis et al. [30] proposed the Time-Parameterized R-tree (TPR-tree). Using this moving objects are indexed, where the moving object's location is shown as a linear function of time. Benetis et al. [31] developed query evaluation algorithms where search is based on the TPR-tree for both NN as well as reverse NN. Tao et al. [32] extended TPR-tree to the TPR*-tree by optimizing the performance of TPR-tree.

Here, there are two types of work in monitoring continuous spatial queries. The first type makes an assumption that the movement trajectories are well known. Monitoring Continuous kNN queries has been analysed for moving queries over static objects [33] and for objects that move linearly [34], [35]. Iwerks et al. [34] monitored distance semi joins for two data sets moving linearly by extending it [36]. But, as mentioned in [37], the assumption about known-trajectory is not valid for most of the application situations. The second type on object movement pattern does not make any assumption. Xu et al. [38] and Zhang et al. [39] proposed returning to the user both the query result and the scope for validity with the resulting solution being the same. By retaining the same, the query is evaluated again only when the validity scope is exited by the query, where the result remains unchanged. This proposal is useful for stationary object only. PAM focuses on the notion of "safe region" [1].In Hermes framework [5] comparison/evaluation of anonymization algorithms are made use of where the platform integrates two well-known anonymization algorithms, namely NWA [17] and W4M [18]. Both algorithms take as input trajectories which may have been extracted from a query posed to HERMES, and transform them into anonymous equivalents, subsequently stored in the MOD.

6 System Implementation

This paper aims at integrating three other frameworks. But the implementation procedure requires only to consider HERMES [5] and Statistical framework [4] and not about PAM framework as HERMES framework in itself is TrPF [2] and PAM [1]. Statistical framework [4] incorporates statistics based problem to solve the security the issue. To achieve this event triggered transmissions has to be refrained. Here each node transmits some fake messages and the real events are not transmitted as they occur. But they are sent instead of the next scheduled fake message. Therefore the adversary cannot identify and distinguish between the real and the fake events. To this the concept of interval indistinguishability [4] is added, otherwise the adversary can make statistical analysis to distinguish between the fake and real messages. If there are two time intervals with one denoting the real event transmission and other the fake one, these two time intervals are indistinguishable if the inter-transmission distribution of time during these two specified interval could not be differentiated with a significant level of confidence. As mentioned above, to obtain the source anonymity using statistics, the statistical framework proposes the notion of Statistical Source Anonymity (SSA):

Xj: random variable for the time between j and (j+1)th transmission. E[Xj]: desired mean for all j. let E[Xj] = μ.

IF: fake interval.

IR: real interval.

If there are no real event transmissions occurring the node transmits fake messages at a predetermined probability distribution.

During a fake interval, IF, for any Xj-1, Xj € IF

E [Xj| Xj-1 < μ] = μ (1)

since Xj, Xj-1 are independent.

During a real interval, both fake and real event can possibly occur.

Ej is a random variable representing either a real event RE or fake event FE. Ej takes any one of these values with any probability. There are two conditions to be satisfied by design:

The time between the real event transmission and its preceding fake event should be shorter that the mean, μ (to reduce delay).

The time between the real event transmission and its successive real event should be longer than

μ.

During a real interval, IR, for any Xj-1, Xj € IR

$$E\ [Xj \mid Xj\text{-}1 < \mu, Ej = RE] > \mu \qquad (2)$$

$$E\ [Xj \mid Xj\text{-}1 < \mu, Ej = FE] > \mu \qquad (3)$$

By using (2) and (3),

$$E\ [Xj \mid Xj\text{-}1 < \mu] = E\ [Xj \mid Xj\text{-}1 < \mu, Ej = RE].\ Pr\ [Ej{=}RE]$$
$$+\ E\ [Xj \mid Xj\text{-}1 < \mu, Ej = FE].\ Pr\ [Ej{=}FE] > \mu.Pr[Ej{=}RE] + \mu.Pr\ [Ej{=}FE] = \mu$$

Inter-transmission time is either greater or lesser than μ2. With this the real and fake event can be generated without any interval indistinguishability. Now with statistical framework implemented on the mobility data sets, EPF proceeds to implement the HERMES [5] framework which is the next level of the built architecture. HERMES integrates two anonymization algorithms: NWA [17] and W4M [18]. These algorithms take the input trajectories that is extracted from a query posed to HERMES [13] query engine and convert the resulting trajectories into anonymous one and hence stored in the MOD. It also incorporates several data mining methodologies such as TRACLUS [19], T-Optics [20] and CenTR-I-FCM [21], K-medoids [22] and Bisecting K-medoids [23] which is applied and hence the patterns arising from original data is compared with the patterns resulting from anonymized data. With this implementation of HERMES framework all the attack that EPF aimed to resolve is thereby achieved.

7 Conclusion

This paper analyses the possible security and privacy threats arising with the growth in mobile devices and Location Based Services. A solution is proposed to overcome the attacks that have been analyzed and identified as the most common on Moving Object Databases. The presented framework is the integration of already existing frameworks namely Statistical framework [4], HERMES framework [5] and PAM framework [1] thereby resolving the common attacks such as user identification attack, sequential tracking attack, sensitive location tracking attack, source identification attack, spatio temporal correlation inference attack. Although each framework as such resolves at least one of the attack efficiently they don't provide a complete privacy against all the possible attacks and threats. EPF framework thereby aims to solve most of the security and privacy issues in moving objects.

8 Acknowledgement

This acknowledgement is intended to be a thanks giving measure to all those involved directly or indirectly with this research. This work was supported by Mr.B.A.Sabarish by providing necessary information and data regarding the research. The authors extend their sincere thanks to their family and friends for helping and motivating during the course of the project.

9 References

[1] Haibo Hu, Jianliang Xu, Senior Member, IEEE, and Dik Lun Lee. PAM:An Efficient and Privacy Aware Monitoring Framework for Continuously Moving Objects. *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 3, March 2010.

[2] Sheng Gao, Jianfeng Ma, Weisong Shi, Guoxing Zhan and Cong Sun. TrPF: A Trajectory Privacy-Preserving Framework for Participatory Sensing. *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 6, 2013.

[3] Xinyi Huang, Yang Xiang, Ashley Chonka, Jianying Zhou and Robert H. Deng. A Generic Framework for ThreeFactor Authentication: Preserving Security and Privacy in Distributed Systems. *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 8, 2011.

[4] Basel Alomair, Andrew Clark, Jorge Cuellar and Radha Poovendran. Towards a Statistical Framework for Source Anonymity in Sensor Network. *IEEE Transactions on Mobile Computing*, vol.12, no.2, 2013.

[5] Nikos Pelekis, Anargyros Plemenos, Aris Gkoulalas-Divanis, Despina Kopanaki, Marios Vodas and Yannis Theodoridis. A Benchmark Framework for Privacy Preserving Mobility Data Querying and Mining Methods. Extending Database Technology, 2012.

[6] M. Shao, Y. Yang, S. Zhu and G. Cao. Towards Statistically Strong Source Anonymity for Sensor Networks. In *Proceedings of the 27th Conference on Computer Communications–INFOCOM'08. IEEE Communications Society*, 2008, pp. 466–474.

[7] Y. Yang, M. Shao, S. Zhu, B. Urgaonkar and G. Cao. Towards event source unobservability with minimum network traffic in sensor networks. In *Proceedings of the first ACM conference on Wireless network security–WiSec'08. ACM*, 2008, pp. 77–88.

[8] Y. Zeng, J. Cao, S. Zhang, S. Guo, and L. Xie. Random-walk based approach to detect clone attacks in wireless sensor networks. *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 5, pp. 677–691, 2010.

[9] N. Li, N. Zhang, S. Das and B. Thuraisingham. Privacy preservation in wireless sensor networks: A state-of-theart survey. Elsevier Journal on Ad Hoc Networks, vol. 7, no. 8, pp. 1501–1514, 2009.

[10] H. Wang, B. Sheng, and Q. Li. Privacy-aware routing in sensor networks. *Elsevier Journal on Computer Networks*, vol. 53, no. 9, pp.1512–1529, 2009.

[11] M. Shao, W. Hu, S. Zhu, G. Cao, S. Krishnamurthy and T. La Porta. Cross-layer Enhanced Source Location Privacy in Sensor Networks. *In Proceedings of the 6th Annual IEEE communications society conference on Sensor, Mesh and Ad Hoc Communications and Networks–SECON'09. IEEE Communications Society,* 2009, pp. 324–332.

[12] B. Carbunar, Y. Yu, W. Shi, M. Pearce, and V. Vasudevan. Query privacy in wireless sensor networks. *ACM Transactions on Sensor Networks*, vol. 6, no. 2, pp. 1–34, 2010.

[13] N. Pelekis, E. Frentzos, N. Giatrakos, and Y.Theodoridis. HERMES: Aggregative LBS via a trajectory DBengine. *In Proceedings of SIGMOD.*

[14] N. Pelekis, A. Divanis Gkoulalas, M.Vodas, D.Kopanaki and Y.Theodoridis. Privacy-Aware Querying over Sensitive Trajectory Data. *In Proceedings of CIKM*.

[15] Oracle, the Swing Tutorial. URL: http: //download.oracle.com/javase/tutorial/uiswing.(Accessed 19 Jan.2012).

[16] NASA, World Wind Java SDK. URL:http://worldwind.arc.nasa.gov/java. (Accessed: 19 Jan. 2012).

[17] O. Abul, F. Bonchi and M. Nanni. Anonymization of moving objects databases by clustering and perturbation. Information Systems, 35(8):884-910.

[18] O. Abul, F.Bonchi and M.Nanni. Never walk alone: Uncertainty for anonymity in moving objects databases. In *Proceedings of ICDE*.

[19] J.G. Lee, J.Han and K.Y. Whang. Trajectory clustering: a partition-and-group framework. *In Proceedings of SIGMOD.*

[20] M. Nanni and D.Pedreschi. Time-focused clustering of trajectories of moving objects. *Journal of Intelligent Information Systems*, 27(3):267-289.

[21] N. Pelekis, I. Kopanakis, E. Kotsifakos, E. Frentzos and Y. Theodoridis. Clustering uncertain trajectories. Knowledge and Information Systems, 28(1):117-147.

[22] L. Kaufman, P.J. Rousseeuw. Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, NY.

[23] M. Steinbach, G. Karypis, V. Kumar. A comparison of document clustering techniques. *In Proceedings of KDD Workshop on Text Mining.*

[24] D. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 1981.

[25] M. Reed, P. Syverson, and D. Goldschlag. Anonymous connections and onion routing. *IEEE Journal on Selected Areas in Communications*, 1998.

[26] M. Gruteser and D. Grunwald. Anonymous Usage of Location-Based Services through Spatial and Temporal Cloaking. *In Proceedings of the 1st international conference on Mobile systems, applications and services*, 2003.

[27] Y. Xi, L. Schwiebert and W. Shi. Preserving source location privacy in monitoring-based wireless sensor networks. In IPDPS 2006. *The 20th International Parallel and Distributed Processing Symposium*, 2006.

[28] B. Hoh and M. Gruteser. Protecting Location Privacy through Path Confusion. In Secure Comm 2005. *First*

International Conference on Security and Privacy for Emerging Areas in Communications Networks, 2005.

[29] M. Shao, Y. Yang, S. Zhu, and G. Cao. Towards Statistically Strong Source Anonymity for Sensor Networks. INFOCOM 2008. *The 27th IEEE Conference on Computer Communications*, 2008.

[30] S. Saltenis, C.S. Jensen, S.T. Leutenegger and M.A. Lopez. Indexing the Positions of Continuously Moving Objects. *Proc.ACM SIGMOD*, 2000.

[31] R. Benetis, C.S. Jensen, G. Karciauskas and S. Saltenis. Nearest Neighbor and Reverse nearest Neighbor Queries for Moving Objects. *Proc. Int'l Database Eng. and Applications Symp. (IDEAS),* 2002.

[32] Y. Tao, D. Papadias and J. Sun. The TPR*-Tree: An Optimized Spatio-Temporal Access Method for Predictive Queries. *Proc. Int'l Conf. Very Large Data Bases (VLDB),* 2003.

[33] Y. Tao, D. Papadias and Q. Shen. Continuous Nearest Neighbor Search*. Proc. Int'l Conf. Very Large Data Bases (VLDB),* 2002.

[34] G. Iwerks, H. Samet and K. Smith. Continuous k-Nearest Neighbor Queries for Continuously Moving Points with Updates*. Proc. Int'l Conf. Very Large Data Bases (VLDB)*, 2003.

[35] K. Raptopoulou, A. Papadopoulos and Y. Manolopoulos. Fast Nearest-Neighbor Query Processing in Moving Object Databases. GeoInfomatica, vol. 7, no. 2, pp. 113-137, 2003.

[36] G.S. Iwerks, H. Samet and K. Smith. Maintenance of Spatial Semi join Queries on Moving Points. *Proc. Int'l Conf. Very Large Data Bases (VLDB),* 2004.

[37] Y. Tao, C. Faloutsos, D. Papadias and B. Liu. Prediction and Indexing of Moving Objects with Unknown Motion Patterns*. Proc. ACM SIGMOD*, 2004.

[38] J. Xu, X. Tang and D.L. Lee. Performance Analysis of Location Dependent Cache Invalidation Schemes forMobile Environments*. IEEE Trans. Knowledge and Data Eng*., vol. 15, no. 2, pp. 474-488, Mar./Apr. 2003.

[39] J. Zhang, M. Zhu, D. Papadias, Y. Tao and D.L. Lee. Location Based Spatial Queries. *Proc. ACM SIGMOD*, 2003.

Divya.C, a final year Information Technology student of Amrita Vishwa Vidyapeetham was born in the year 1992 in the southern part of India -Tiruppur, Tamilnadu. She's pursuing her Bachelor's Degree in Computer Science Engineering and set to receive her graduation in the year 2014. She's interested in the research and developments in the fields of Data Mining and Information Security.

Harini Shankar, a final year Information Technology student of Amrita Vishwa Vidyapeetham was born in the year 1992 in Kolkatta, West Bengal. She's pursuing her Bachelor's Degree in Computer Science Engineering and set to receive her graduation in the year 2014. She's interested in the research and developments in the fields of Data Mining and Information Security.

B.A.Sabarish, was born in Tamil Nadu. He received the M.Tech degree in computer science and currently is an assistant professor in the Department of Information Technology, Amrita Vishwa Vidyapeetham, Coimbatore. His research interests include Wireless Sensors, Data Mining.