An Efficient Mix-Zone based Secure Location Proof Activities in Urban Areas

Kiruthiga N[1], Sandhiya M[2], Saranya P[3] ,Tharani Priyangha M[4]

[1] Assistant Professor,Department of Computer Science and Engineering,Sri Krishna College of Technology, Coimbatore

[2,3,4] UG Scholar,Department of Computer Science and Engineering,Sri Krishna College of Technology, Coimbatore

Abstract— Activity-based social networks, where people used to upload and share information about their location-based activities (e.g., the routes of their activities), are increasingly popular. Such systems,raise their privacy and security issues: the service providers know the exact locations of their users; the users can report fake location information to, for example, unduly brag about their performance. In this paper, propose a secure privacy-preserving system for reporting location-based activity summaries (e.g., the total distance covered and the elevation gain). Our solution is based on a combination of homomorphic encryption cryptographic techniques and geometric algorithms, and it relies on existing Wi-Fi access point networks deployed in urban areas. The mix-zone network is divided into different zones brought under the control of different certification authorities (CAs), forcing users to change its certificate when moving from a zone to another.The characteristics of the proposed solution is that it prevents attackers from tracking the mobility of the vehicles. Finally, it shows by simulations that the proposed mix system is effective in various scenarios.

Index Terms— **Mix-zone, Secure location proof**, **Homomorphic encryption cryptographic, Certification authorities.**

I. INTRODUCTION

The presence and usage of embedded sensors in mobile devices has significantly increased. Location based services (LBSs) are nowadays able to keep users informed about traffic conditions, significant events happening in proximity and the presence of other people with similar interests. More recently, LBSs are increasingly used by people to track, monitor and share physical activities and performance over time; in particular, health- and wellness related applications.

A popular feature of such applications is the ability to share the summaries of users' activities and their performance statistics with other users or service providers on social networks. For instance, users can share the total distance covered during their activities, the cumulative elevation gain and the actual path. Inorder to exchange their data, users can be rewarded with coupons and discounts or even with cash, with awards in competitions, or simply with a better "social reputation" within their social circles.

The activity tracking and sharing services are gaining popularity, there are two important issues that can hinder their wide-scale adoption and viability. First, users' location data,known to service providers,which is used to infer private information about them, such as their home/work social networks. Second, users might be tempted to cheat when reporting their performance, in order to obtain a better reward, which can endanger the viability of the system for the service provider and its affiliates, as well as its attractiveness to other users. Location cheating is achieved by making mobile devices report erroneous location information to the activity tracker app, or by spoofing the GPS or Wi-Fi signals used to geo-locate the users'.

In exchange for their data, users are offered various incentives. For example, users can receive discounts, coupons or even cash, awards at competitions or simply points to improve their social reputation. In addition, many companies, including big names such as British Petroleum (BP), Bank of America and Autodesk, are giving activity-tracking devices to their employees to encourage healthier lifestyles and, as a result, improve productivity and lower corporate insurance costs[10]. Similarly, health insurance companies such as United Health, Kaiser Foundation Group, Humana Group and Aetna have created programs to include activity-tracking devices into their policies, i.e., consumers are rewarded by the insurers with lower rates based on their activity summaries.

To assess the awareness and concerns of users of activity tracking applications regarding opportunities to cheat and privacy issues, we conducted a user survey of 50 participants. Our survey participants are active Run Keeper users who we recruited on the Amazon Mechanical Turk platform. In the survey questionnaire, we first informed the participants about existing opportunities to cheat and privacy issues of fitness-tracking applications such as Run-Keeper, and we then polled them about their awareness and their concerns (see Section 6.1 and Appendix F of the supplemental material, includes the full transcript of the

questionnaire). Regarding opportunities to cheat, we found that all the participants were unaware of them and 48% were extremely concerned about it. Regarding privacy issues, we found that 90% of the participants were unaware of them and 82% were very or extremely concerned about them[9]. These results raises the awareness and the need for technical solutions to build cheat-proof and private activity-tracking apps.

In this paper propose to gain a proper understanding of the privacy properties of mix zones it is important to and out how hard it is to break the anonymity the system provides. The mix zone approach for calculating anonymity gives the degree of success in playing the attacker role attempting to recover the long-term user identities hidden by the constantly changing pseudonyms is an inverse measure of the anonymity offered by the system. Our approach consists of two phases: First, users obtain secure and privacy-preserving proofs of performance during their activities,that relys on a lightweight message exchange protocol on a user's mobile device and the Wi-Fi access points encountered while pursuing the activity; second, the service provider calculates an accurate summary of a user's activity, such as the total distance covered between two time instants or the elevation gain, without learning any additional information about the user's actual location.

II. Related Work

Cheating on activity-based social networks is becoming a serious problem the users can easily override four square's GPS verification mechanisms by modifying the values returned by the calls to the geo-location API of smartphones. Similarly, another proposed system based on black-box approach to uncover the mechanisms used by Foursquare and Facebook Places to detect location attacks and propose several ways to circumvent them. The analyse data from Foursquare and Gowalia and find that incentives to cheat exist because people actively check-in and collect rewards[1]. Thus, it is important to balance incentives with a more effective verification of users' location claims. In this regard, the fake check-ins lead not only to monetary losses for the venues offering deals for anlocation-based check-ins but also to the degradation of the quality of service provided by recommendation systems that rely on users' location information[5]. The privacy and correctness of location-based applications,in which users are unable to prove that they have satisfied badge conditions without revealing the time and location of their check-ins.

The cheating, researchers have also proposed several mechanisms that offer secure verification of location information. From a broad perspective, such mechanisms can be grouped in three categories:independent-infrastructure,infrastructure-dependent and hybrid mechanisms. In the infrastructure-independent approach, a user obtains location evidence from her neighbours by using short-range communication technologies, such as Bluetooth. Specifically, the propose a location authentication protocol where a set of users help verify each other's' location claims and this protocol operates by keeping a centralized authority that, based on users spatio-temporal correlation, decides whether such claims are authentic or not. Similarly, propose a system in which mutually co-located users rely on Bluetooth communications to generate their location claims that are then sent to a centralized location verifier. In addition to the security and privacy guarantees enable individual users to evaluate their own location privacy and decide whether to accept location proof requests by other users[3]. Provide a formal analysis of the conditions needed in an ad-hoc network which enables for distance-based localization protocols in wireless networks. Similar approaches have been explored in mobile sensor networks.

Inline with our work, the infrastructure-dependent studies assume the presence of a centrally-operated set of access points (AP) to produce and verify location claims. For instance,ensure that the availability of a user in a given region, the AP can require her to be execute together a nonce-based, challenge-response protocol, with constraints on the maximum round-trip delay of the messages exchanged between the user and the AP, or any distance bounding protocol, which enables the AP to check the minimum distance between itself and the user. In particular, propose a verifiable multiliterate protocol that can be used to securely position nodes in a wireless network[4]. Once the secure localization phase is done, the location proof can be obtained by the user , which is a document signed by the witnesses to certify that at a specific time, the user is at a specific geographical location; for example, an AP can embed its coverage range, its centre coordinate and a timestamp in the location proof, in order to certify that at the specified timestamp, the user is in the coverage area of the AP. Alternatively, a user can choose to obtain location proofs for the levels of granularity for the precision of location, and choose the one to disclose to the service provider depending on her preferences and privacy sensitivity[2].

Location-Based Services (LBS) personalize the service which provides the grant access to resources according to the current location of users. They are used in a variety of contexts, such as geosocial network, real-time traffic monitoring, discount tied to the visit of a particular shop or local electronic election. In most of current schemes, the location of a user/device is determined by the device itself (e.g., through GPS) and forwarded to the LBS provider. Secure distance-based localization in the presence of cheating beacon (or anchor) nodes is an important problem in mobile wireless adhoc sensor networks. Despite significant research efforts in this direction, some fundamental questions still remain unaddressed: In the presence of cheating beacon nodes in which the necessary and sufficient conditions to guarantee a bounded error during a two-dimensional distance-based location estimation? Under these necessary and sufficient conditions, what class of localization algorithms can provide this error bound.

The addressed the problem of secure distance-based localization in the presence of cheating beacon nodes. By means of a sound mathematical analysis, we have derived the conditions for securing the distance-based localization in the presence of cheating beacons. Specifically, we have outlined the necessary and sufficient conditions for

achieving a bounded localization error, and defined a nonempty class of algorithms that can achieve such a bounded error[8]. We have also proposed three novel distance-based localization algorithms, specifically a polynomial-time algorithm and two heuristic-based algorithms which belongs to this class of error bound distance-based localization algorithms. We have verified the localization accuracy and execution efficiency of these algorithms using measurements from simulation experiments.

To counter this threat, a LBS needs the requesting device to formally prove that it really is at the claimed location[7]. This notion has been formalized through the concept of location proof, which is a digital certificate attesting that someone was at a particular location at a specific moment in time. A location proof architecture is a system by which users can obtain location proofs from neighbouring witnesses (e.g., trusted access points or other users) that can later be verified by other entities. In recent years, several location proof architecture is proposed in the literature. Most of these approaches require the users to disclose their identities or their positions to a centralized server, thus raising privacy issues such as the possibility of tracing the movements of users of the location proof architecture.

III. Proposed Approach

The proposed research carries potential broader impacts and significant intellectual merits in the following aspects. (1) To investigate the impact of inferential attacks on LBS users in wireless networks and mobile, and prove the vulnerability of using long-term pseudonyms for camouflaging users' real identities. (2) The propose a novel privacy metric to quantify system's resilience to such attacks. (3) An effective and extensible privacy architecture based on the mix zone model is designed. (4) To conduct rigorous analytical study, and design a privacy protection mechanism under urban mobility model constraints, e.g., traffic density and heterogeneity on different roads. (5) This proposal addresses the privacy preservation problem from a novel angle and lays a solid foundation for future research in protecting the users' location privacy.

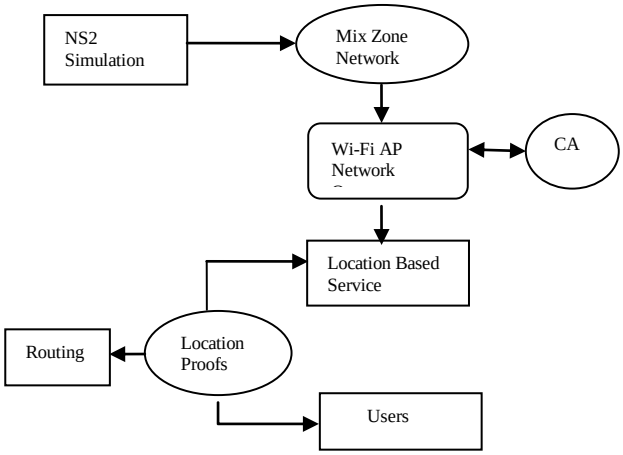

Fig 1.Block diagram for proposed system

A. Location Service Infrastructure

The target LBS system involves mainly three parties: Location Service Infrastructure (LSI),Users, and third party LBS Service Providers, shown in fig.1. Each user in the system is represented by his device id or real identity when communicating with LSI, and represented by a pseudonym at LBS server side. The location information of the users is actively or passively updated to LSI. Since a specific BS application is typically interested in a set of Point-of-Interests (POIs) within certain area is shown in fig 1, these physical locations are registered to LSI. When a user approaches one of the POIs, LSI notifies LBS server about the location of the user using the user's pseudonym. Service information is then delivered back to the user through LSI. As a result, the users' real identities are shielded by LSI, and the third-party LBS service providers can only obtain users' pseudonyms.

These POIs in an area can be modelled as an undirected graph G (V,E), where V is the set of vertices representing the POIs, and E is the set of road segments that connect consecutive POIs. In the proposed system architecture, all vertices are potential mix zone deployment locations. While providing services to the users, LBS applications can obtain a trajectory file recording a user's footprints. Each entry in the trajectory file is a 3-tuple: <pseudonym, timestamp, location>. A trajectory record belonging to a user defines a path consisting of one or a sequence of possible repeated vertices.

B. Adversary Model

To consider LSI to be trustworthy, whereas third-party LBS application servers are not trustworthy. They may directly attack a mobile user's privacy, or secretly sell information to organizations or other individuals. An adversary A refers to any entity formed by one or more malicious parties (by colluding) who aim at learning the locations associated with mobile users' true identities. We do not consider the case that A actively stalks the particular user. Since an adversary has the complete trajectory profiles camouflaged by pseudonyms, it is often characterized as a global passive eavesdropper that becomes the major threat. Accidental leakage of a user's real identity may become side information to an adversary. The goal of the adversary A is to identify the target mobile user in the trajectory file is based on side information matching, and learn the complete footprints left by the tracking targe information matching based attack, i.e., inferential attack. It must be noted that while the trajectory files contain accurate location records for service purposes, the side information may be noisy or even incorrect[6]. This is because of the source that the side information is unreliable, e.g., personal encounter or context inference.

C. Mix Zone Model

The concept of mix zone service restricted area where mobile users can change their pseudonyms so that the mapping between their old pseudonyms and new pseudonyms are not revealed. The users with pseudonyms A-E enter the mix zone from different entrances and exit the mix zone with a different set of pseudonyms F-J at

approximately the same time. The links between old and new pseudonyms are not observed by any outsider. This change effectively "mixes" the identities of all the users to achieve privacy protection. In the aforementioned system model, a mix zone is established by LSI at the software level. And LSI selects mix zone from the set of registered POIs. Once a POI is chosen as a mix zone, LSI will assign a set of new pseudonyms to the users leaving this POI. Such a software level mix zone establishment approach has considerable flexibility over physical deployment of mix zones, because the location and the size of the mix zones are not constrained by the terrestrial borders and can be easily adjusted.

A commonly used metric to quantify a mix zone's protection effectiveness is information entropy given by.

$$H_m = -\sum_u p_u \log p_u \qquad (1)$$

where p_u stands for the probability of mapping from an old pseudonym to a new pseudonym. If multiple mix zones are deployed alongside a user's routes, the user's continuous trajectory is broken into a set of discrete segments, and each discrete segment is associated with a unique pseudonym. This causes an adversary to lose the tracking target.

D. Users in Mix Zone

Users pursue their location-based activities, where they move in a given geographical region, and that they want to obtain statistics or summaries of their activities. These users are equipped with GPS- and WiFi-enabled devices and they have sporadic Internet connectivity (at least at some point in time before and after the activity). Therefore, they can locate themselves and communicate with nearby Wi-Fi access-points. It assumes a unit-disc model for Wi-Fi communications, in which a user and an AP can communicate only if the distance between them is lower than a given radius R, which is constant across all users and all APs. In particular, we assume that users cannot violate this model by, for example, increasing the transmission power of their devices. We assume that users can obtain random identifiers (or pseudonyms) from the online service provider, and that they can use such pseudonyms to protect their privacy while pursuing their activities.

E. Wi-Fi AP Network Operator

Multiple Wi-Fi network operators, and that each operator controls a set of fixed Wi-Fi APs deployed in the regions where the users pursue their activities. Each AP is aware of its geographic position and of its communication radius. We assume that all the APs have synchronized clocks, and that they are able to compute public-key cryptographic operations. In particular, we assume that all the APs from a same network operator share a public/private group key pair (GKpub;GKpriv), where GKpub is known by the users and the service provider, whereas GKpriv is only known to the network operator and to its APs.

F. Social network provider

The network assumes that there is a social network provider that offers activity summaries and sharing services for its registered users. The provider is able to generate sets of pseudonyms for its users, through using a suitable public-key encryption scheme. Moreover, it is able to verify the authenticity of messages signed with the network operators' group keys (by using their public group keys).

G. Location proofs

The sampling time t_i, a user begins to collect location proofs from the access points in her communication range. To do so, she periodically broadcasts (during a short time interval starting at time t_i) location-proof requests that contain one of her pseudonyms P. Note that a different pseudonym is used for each sampling time. All the access points in her communication range send back messages that contain the pseudonym P,and a timestamp t (i.e., the time at which the request is processed by the access point) and their coordinates (x; y), digitally signed with the private group key GKpriv, namely a location proof $LP = Sig_{GK\ priv}\{P, t, (x, y).$ denotes by $LP_{i,j} = \{P_i, t_{i,j}(x_{i,j}, y_{i,j})\}$ the j-th location proof collected at sampling time t_i (note that we omit the signature for the sake of readability). As the communication and processing delays differ from one access point to another, the location proofs collected from different access points at a same sampling time have different timestamps. Under the unit-disc communication model (with radius R), such a location proof certifies that, at time t, the user is at a distance of at most R to the access point that issues the location proof. In other words, it certifies that the user is in a disc of radius R, cantered at the point of coordinate (x; y). We denote such a disc by C ((x, y), R).

H. Activity proofs

To obtain an activity proof (i.e., a distance proof or an elevation proof), a user sends to any access point (whenever she needs it) the location proofs she collected at two consecutive sampling times t_i and t_{i+1}. The contacted access point first combines the different location proofs, collected at each of the two sampling times, into more precise location proofs, by aligning them by time and intersecting them. As these location proofs have different timestamps, the first step of the combination consists in aligning the different location proofs.

KeyGen: Given the domain parameters (a,b,p,G,n,E) of an elliptic curve E over a finite field Fp where p is a large prime that satisfy . Where G is the base point of order n, note that n*G = ∞, the private key x is randomly selected from [1, n-1], where the public key is Y=xG, another point on the curve.

Encryption: Given the plaintext m and Y, output C
1. k ∈ [1, n – 1]
2. M = map (m)= mG
3. C= (R, S) = (kG, kY+mG)

Homomorphic operation: Given that C1, C2... Cn, output C'C'= (k1G, k1Y+m1G)+(k2G, k2Y+m2G)+…+(knG, knY+mnG)
C'= ((k1+k2+..kn)G, (m1+m2+mn)G+(k1+k2+..kn)Y)

Decryption: Given C' and the private key x, output m
1. M = S – xR
2. m =rmap(M)

In fig.2 the map function satisfies the desired additive homomorphic property. However, the reverse mapping function is the shortcoming of this scheme, and the reverse function maps a given point M into a plaintext m, thus, on M must be resolved.

A public-key cryptosystem is as usual defined by algorithms E, D for encryption and decryption and a key generation algorithm KG. The proposed looking at systems that are homomorphic, in the following sense: the set of cipher texts is an Abelian group, where the group operation is easy to compute given the public key.

Non-interactive Zero-Knowledge with Key Setup. CA engage in the following steps:

Step 0: V sends her public key (n, c) to CA;

Step 1: V proves to CA that n is well-formed;

Step 2: V proves knowledge of the plaintext e hidden within c; and that this value e lies in the specified interval.

For an efficient homomorphic encryption scheme, it is crucial to make sure that the size of the cipher texts remains polynomials bounded in the security parameter σ during repeated computations.

IV. Experimental Results

During the simulation, each node starts its journey from a random spot to a random chosen destination. This process repeats throughout the simulation, causing continuous changes in the topology of the underlying network. PDR is the ratio of the number of data packets received by the destination node to the number of data packets sent by the source mobile node. It can be evaluated in terms of percentage (%). This parameter is called as "success rate of the protocols", and is described as follows:

$$\text{PDR} = \left(\frac{\text{Send Packet no}}{\text{Receive packet no}}\right) \times 100 \qquad (2)$$

Throughput is the average rate of successful message delivery over a communication channel as shown in the fig.3. This data may be delivered over a physical or logical link, or pass through a certain network node.

$$X = \frac{C}{T} \qquad (3)$$

Where X is the throughput, C denotes the number of requests that are accomplished by the system, where T denotes the total time of system observation.

Average end-to-end delay Average end-to-end delay signifies how long it will take a packet to travel from source to destination node. In fig.4 it includes delays due to route discovery, queuing, propagation delay and transfer time.

$$D_{end-end} = N(d_{trans} + d_{prop} + d_{proc} \qquad (4)$$

Where dend-end= end-to-end delay, dprop= propagation delay, dtrans= transmission delay,dproc= processing delay,dqueue= Queuing delay and N= number of links. This metric is useful in understanding the delay caused while discovering path from source to destination.

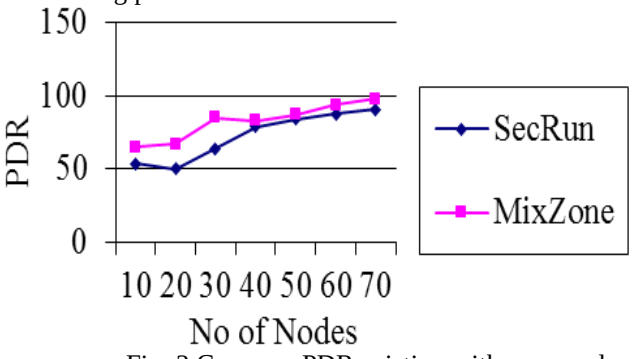

Fig. 2 Compare PDR existing with proposed

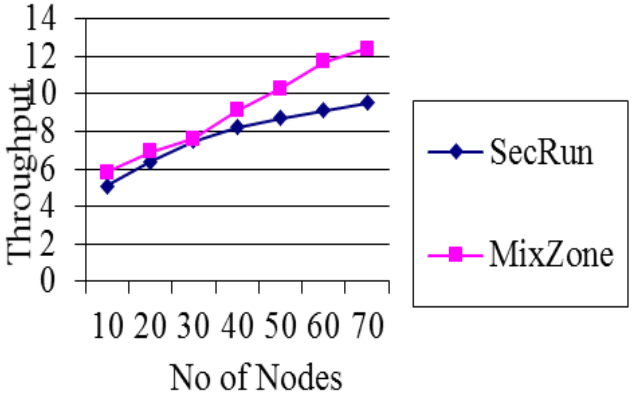

Fig. 3 Compare Throughput existing with proposed

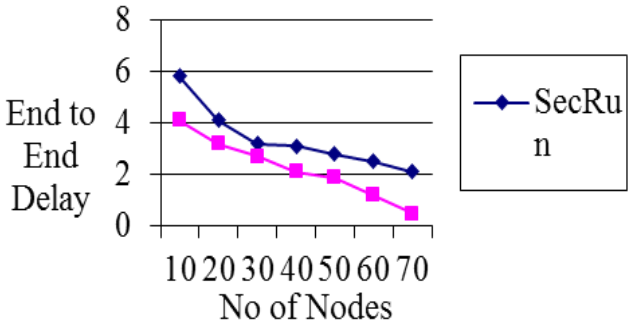

Fig. 4 Compare delay time existing with proposed

V. Conclusion

This propose a secure privacy-preserving system for reporting location-based activity summaries (e.g., the total distance covered and the elevation gain). Our solution is based on a combination of homomorphic encryption cryptographic techniques and geometric algorithms, and it relies on existing Wi-Fi access point networks deployed in urban areas. The mix-zone network is divided into different zones that are brought under the control of different certification authorities (CAs), forcing users to change its

certificate when moving from a zone to another. First, users obtain their secure and privacy-preserving proofs of performance during their activities, by relying on a lightweight message exchange protocol between a user's mobile device and the Wi-Fi access points encountered while pursuing the activities; second, the service provider computes an accurate summary of a user's activity, such as the total distance covered between two time instants or the elevation gain, without learning any other additional information about the user's actual location. Finally, to quantify the users' location privacy, we contemplate modeling the system (in the presence of many users pursuing location-based activities in the same region) as a mix-zone problem, define formal privacy metrics and evaluate them on real data-sets or through experiments. Finally, our solution is able to take advantage of the co-existence of multiple access point operators to improve the accuracy/privacy trade-off.

REFERENCES

[1] Guyon S. Abbas, M. Merabti, D. Llewellyn-Jones, and K. Kifayat, "Lightweight sybil attack detection in manets," IEEE Syst. J., vol. 7, no. 2, pp. 236–248, Jun. 2013.
[2] S. Eidenbenz, G. Resta, and P. Santi, "The COMMIT protocol for truthful and cost-efficient routing in ad hoc networks with selfish nodes," IEEE Trans. Mobile Comput., vol. 7, no. 1, pp. 19–33, Jan. 2008.
[3] E. Hernandez-Orallo, M. D. Serrat, J.-C. Cano, C. M. T. Calafate, and P. Manzoni, "Improving selfish node detection in MANETs using a collaborative watchdog," IEEE Comm. Lett., vol. 16, no. 5, pp. 642–645, May 2012.
[4] E. Hernandez-Orallo, M. D. Serrat Olmos, J.-C. Cano, C. T. Calafate, and P. Manzoni, "Evaluation of collaborative selfish node detection in MANETS and DTNs," in Proc. 15th ACM Int. Conf. Modeling, Anal. Simul. Wireless Mobile Syst., New York, NY, USA, 2012, pp. 159–166.
[5] J. Hortelano, J.-C. Cano, C. T. Calafate, M. de Leoni, P. Manzoni, and M. Mecella, "Black hole attacks in p2p mobile networks discovered through Bayesian filters," in Proc. Int. Conf. Move Meaningful Internet Syst., 2010, pp. 543–552.
[6] Y. Li, G. Su, D. Wu, D. Jin, L. Su, and L. Zeng, "The impact of node selfishness on multicasting in delay tolerant networks," IEEE Trans. Veh. Technol., vol. 60, no. 5, pp. 2224–2238, Jun. 2011.
[7] M. D. Serrat-Olmos, E. Hernandez-Orallo, J.-C. Cano, C. T. Calafate, and P. Manzoni, "A collaborative Bayesian watchdog for detecting black holes in MANETs," in Proc. 6th Int. Symp. Intell. Distrib. Comput. VI, 2012, vol. 446, pp. 221–230
[8] C. K. N. Shailender Gupta and C. Singla, "Impact of selfish node concentration in MANETs," Int. J. Wireless Mobile Netw., vol. 3, no. 2, pp. 29–37, Apr. 2011.
[9] Y. Zhang, L. Lazos, and W. Kozma, "AMD: Audit-based misbehavior detection in wireless ad hoc networks," IEEE Trans. Mobile Comput., vol. PP, no. 99, 2012
[10] Y. Zhang, W. Lee, and Y.-A. Huang, "Intrusion detection techniques for mobile wireless networks," Wireless Netw., vol. 9, no. 5, pp. 545–556, Sep. 2003.

Ms.N.Kiruthiga is currently working as Assistant Professor in Department of Computer Science and Engineering at Sri Krishna College of Technology. Her research interest includes Wireless Sensor Networks, Network and Information security.

Ms. M.Sandhiya is currently pursuing Bachelor's degree in Computer Science and Engineering at Sri Krishna College of Technology, Coimbatore. Her area of interest is Network security .

Ms. P.Saranya is currently pursuing Bachelor's degree in Computer Science and Engineering at Sri Krishna College of Technology, Coimbatore. Her area of interest is Network security .

Ms. M.Tharani Priyangha is currently pursuing Bachelor's degree in Computer Science and Engineering at Sri Krishna College of Technology, Coimbatore. Her area of interest is Network security.