International Journal of Computer Networks and Applications (IJCNA)
Volume 2, Issue 5, September – October (2015)

RESEARCH ARTICLE

Efficient Collaborative Technique using Intrusion Detection System for Preserving Privacy in Location-based Services

Muhammad Jawad Ikram
Department of Computer Science, Faculty of Computing and Information Technology
King Abdul-Aziz University, Jeddah, KSA.
mshahid@stu.kau.edu.sa

Jonathan Cazalas
Department of Computer Science, Faculty of Computing and Information Technology
King Abdul-Aziz University, Jeddah, KSA.
jcazalas@kau.edu.sa

Abstract – The ubiquitous nature of smartphones and GPS-enabled devices, coupled with the increasingly popular usage of location-based services, has effectively created an environment where data access truly is anywhere at any time. While said environment is indeed convenient and quite useful, the unfortunate reality is that users are exposed to a variety of privacy and security threats. User location information can be tracked and then used in malicious ways by non-trusted applications and adversaries. We address this problem by proposing an efficient, collaborative technique that is integrated with an intrusion detection system and rekeying techniques. The algorithm is parameterized by defining performance and security metrics, which can then be used to find optimal settings, both in terms of privacy and quality of service. Based on the proposed performance-security metrics, the tradeoff between privacy and quality of service can be quantified.

Index Terms— Intrusion Detection System (IDS), LBRA, Location-Based Services (LBS), Privacy, Rekeying.

1. INTRODUCTION

The last decade has witnessed an explosive growth of GPS-enabled smart phones and tablets, resulting in an ever-expanding use of spatial-aware technologies. With these advancements in technology, mobile users can now access a wide range of services from Location-based Service providers [1]. Examples include traffic reports ("Determine the route with the least congestion"), location-based points of interest ("Where is the nearest gas station"), and even location-based advertisements ("Send e-coupons to all users within one mile of my restaurant"). These services are provided based on both the spatial and temporal, or spatio-temporal, information of mobile users [1]. Accordingly, registered users must continuously share their location with a dedicated location-based server, and when a service is requested, users issue location-based queries, which are then executed based on their current location.

Location-based services have been traditionally subdivided into three categories based on the mobility of the client and the object(s) being queried [2]. The first category includes mobile clients that query static, or stationary, objects, such as a mobile client searching for the nearest pizza restaurant. The second category includes static clients that query mobile objects, such as traffic management systems. And finally, the third category includes mobile clients that query other mobile objects, such as a pedestrian searching for a nearby taxi.

Within the domain of location-dependent query processing, two of the more common spatial queries are continuous, range-monitoring queries and k nearest neighbor queries (kNN) [1]. Range queries are such that users request to retrieve all data objects within a given query range. An example may be a user requesting to find all shopping malls within 15 kilometers from their location. Nearest neighbor queries are such that users request to retrieve the k nearest objects relative to their location. An example would be a user requesting to find the 4 nearest hospitals, again, relative to their location.

While the advancements of said technologies have spawned a seemingly endless number of location-based mobile applications, these same applications pose a tremendous privacy and security threat to their customers. In order to achieve optimal results, the location-based query processor must have the most accurate and updated locations of the users performing queries. Users, therefore, are forced to reveal their locations, sacrificing their privacy, all in order to access and benefit from the service. The consequence is exposing themselves to both network and service providers, who can use their trajectories and historical movements to track them. There are chances that these operators may misuse this rich data, such as by selling this information to some

RESEARCH ARTICLE

advertisers or private investigators [3]. Thus, efficient privacy preserving techniques are needed to be developed that provide secured location-based services to mobile users.

As described in Shokri et. al., there is always a tradeoff between privacy and quality of service [1]. By improving one component, we pay in terms of the other. For instance, as the user demands higher privacy levels and chooses to hide, or blur, their location information from the location-based server, the result will be a lower quality of service; not having the exact location information, the server simply cannot return accurate and optimal results. On the flipside, if a user chooses to release his or her location information, effectively sacrificing privacy, the location-aware query processor should be able to provide an optimal answer. Therefore, optimal settings for a system should be such that maximize the privacy of users, while still satisfying the quality of service requirements.

Toby [4] explored the issue of preventing an adversary to locate nodes on the basis of their location information, which is exposed explicitly during communication. They achieved a desirable level of protection simply by reducing location resolution. To accomplish this, they characterize the safety level of a region as the ratio of the area and the number of nodes in that area, with the higher safety levels resulting in less adversarial exposure. The consequence is that every time a node has to release its location, it must re-compute its cloaking box in which the required safety level is obtained. While the idea is straightforward, there are several challenges in implementation. Cloaking boxes should be as small as possible to minimize the effect of reduced location information on overall performance. The capability of a node to compute its cloaking box, without revealing its location, is inherently challenging. Finally, there should be no correlation in a given sequence of cloaking boxes, in order to refine an area which has less than the required level of safety. In the context of ad hoc networks, they overcame these challenges and provided cost-effective solutions.

Mokbel et. al. [5] proposed a novel privacy-aware query processing framework, Casper. Under this framework, both mobile and stationary nodes are capable of obtaining snapshots, as well as performing continuous location-based queries, without exposing their precise coordinates. They present a location anonymizer, in which cloaked spatial regions are used to obfuscate the exact position information. The cloaked spatial regions are determined based on the privacy requirements of users. Mokbel also proposes a privacy-aware query processor [5]. The privacy-aware processor is simply a location-based database server, whose function is to cloak spatial areas in order to protect the exact location information. The privacy-aware server is independent of the computation of user cloaked regions. More specifically, any other anonymization techniques that blur the private location of users into cloaked rectilinear areas can be used for preserving location privacy of users. They propose a shared execution prototype for improving system scalability of processing continuous queries. The experimental results obtained reveal that high quality snapshots and continuous location-based services can be obtained by query processor even as "supporting queries and/or data with cloaked locations". [5] Thus, using this technique, user location can be protected and tracking can be overcome.

Jha [6] deal with the location based resource allocation (LBRRA) in WiMAX and WiMAX - WLAN interface technology. This technique have three phases, firstly to discover solution for trouncing radio link connections that are rejected by the Base Station (BS), which can be helpful for determining users with better quality of service for various applications in real time analysis for any location based networks. Secondly, WiMAX and WLAN technology for LBRA is defined. Finally, the request bandwidth and admitted bandwidth [7] are analyzed with WLAN and WiMAX interface with adaptive modulation and coding.

Limkar [8] proposes a new technique that is based on hidden markov model (HMM). In this technique, HTTP flooding attacks and legitimate HTTP traffic are differentiated. Anomaly is described with the help of an extended hidden Markov model. The technique proposed in Limkar [8] gives predictive pattern of detecting distributed denial of service attacks (DDoS).

Recently, new methods have focused on attempting to overcome tracking in location-based services. Buchanan et al. proposes a new technique, which is based on private equality (PE), a primitive to overcome tracking [9]. The main contribution was in forming a single encrypted table of identities, which allows users to privately compare the identities of their locations by using PE primitive. The advantage of this private match is that the user is able to identify encrypted identities of interesting records simply by performing comparisons. The algorithm proposed in [9] has substantial improvements in computation speeds but has little to no protection with respect to the current location of mobile users. The authors themselves state that their proposed solution "does not strictly protect the current location." [9]

To address these limitations, the focus must be on both improving computation speeds as well as the privacy of locations of mobile users. We propose a collaborative technique that tries to address the problems related to privacy and quality of service that exists in both centralized and user-centric approaches. Users, which demand queries from an LBS, form a mobile, wireless ad hoc network, integrated with efficient intrusion detection and rekeying techniques. The ad hoc network formed by the users is a secured group communication system (GCS). Before a user performs a query, they first broadcast the query within the GCS. If the

RESEARCH ARTICLE

answer exists in the GCS, the service is provided; otherwise, the LBS is contacted. Users in the GCS store the gained information in buffers that can later be provided, on demand, to other nodes in the GCS.

Additionally, to achieve better privacy, efficient IDS and rekeying techniques are used. Rekeying provides safety against outsider attacks, while IDS provides security against insider attacks. Due to collaboration, dependence on the LBS server can be minimized. Finally, performance-security metrics are defined, which can be used to determine optimal settings both in terms of privacy and quality of service. The design objectives are as follows: reduce dependency on the LBS server, make no changes to the architecture of the LBS, of course including the main entities of the LBS infrastructure, maximize mean time to security failure to improve privacy of mobile users, and, finally, minimize service response time to improve quality of service.

In general, the contributions of this paper can be summarized as follows:

1. To minimize reliance on LBS server, a secured and distributed group communication system of mobile nodes is proposed.
2. The group communication system is integrated with efficient intrusion detection and rekeying techniques, which provide better privacy and quality of service.
3. The algorithm is parameterized by defining performance and security metrics, which are then used to quantify the tradeoff between privacy and quality of service.
4. Based on the proposed algorithm, optimal settings, in terms of privacy and quality of service, can be achieved in location-based services.

The rest of this paper is organized as follows. Section 2 reviews related work, both on privacy preserving techniques, in general, and methods to overcome tracking, in specific. In Section 3, the novel collaborative technique is proposed, with a thorough presentation of the integrated IDS and rekeying operations, as well as a how the proposed algorithm efficiently merges with current LBS infrastructure. In Section 4, parameterization of the proposed algorithm is performed, with each state of the system described based on the mathematical equations. In Section 5, performance-security metrics are defined on the basis of the aforementioned parameterization, and optimal system settings are chosen, both in terms of privacy and quality of service. Finally, Section 6 includes concluding remarks and highlights future research directions.

2. RELATED WORK

A number of techniques have been proposed with the objective of providing better privacy and quality of service. These techniques can be categorized into two broad classes: centralized and user-centric [1]. In centralized techniques, a third party (an intermediate proxy server) in the system is introduced, which then encrypts/protects the information by operating between mobile users and the LBS [1]. In this technique, the function of the intermediate proxy server is to anonymize the queries generated by users, such that any information that may identify the user is removed. Alternatively, the intermediate proxy server can merge the query generated by one user with those of others [10]. Unfortunately, with this technique, there are chances that the intermediate proxy server may become untrustworthy, as it is attractive for attackers as a centralized LBS server [1].

In other centralized approaches, users' queries are submitted to an LBS in a different form than actual user queries, usually encrypted using private information retrieval protocols [11]. Other techniques in centralized approaches focus on various techniques of storing the data, such as in encrypted or encoded form [12]. The user-centric approaches are operated on mobile devices of the users [1]. In these techniques, the primary objective is to blur the user location information, for instance by submitting inaccurate GPS coordinates to LBS [1]. However, this affects the quality of service [12].

In addition to providing privacy and higher quality of service, much research has been on methods to overcome tracking in location-based services. The identity of users, position, and path of the users are three important things that need protection [13]. One technique to protect these three things is grid-based cloaking, in which the actual location of a user is blurred without limiting access to location-based services [9].

A common method called obfuscation, or cloaking, is used to protect the user location. In cloaking, a coarse user location is forwarded to LBS rather than the actual information [9]. In Truong [14], it is defined that numerous systems obfuscate the location of users within number of cells. A flexible grid is used in their system, in which users are able to change their cell size. Casper [5] presents a location anonymizer, in which cloaked spatial regions are used to obfuscate the exact position information. The cloaked spatial regions are set by privacy requirements of users. Additionally, a privacy-aware query processor is used, which is simply a location-based database server, whose function is to cloak spatial areas in order to protect the exact location information. Dewri [15] utilizes cloaking regions, and the actual user is hid in a number of other users by providing k-anonymity [9]. To achieve this, it is significant that mutually diverse queries are issued by users in each of the cloaked regions l-diversity.

Another technique to overcome tracking is identifying the context. In Damiani [16], the geometric method used in grid-based cloaking is criticized as actual information can be revealed if the geographical context is known to the untrustworthy party, with this problem occurring particularly

RESEARCH ARTICLE

in semantic locations. Personalized cloaking of semantic locations resolves this problem [9]. In order to preserve privacy, a cloaking region enclosing the user location is generated in Ghinita [17]. In spite of this, untrustworthy entities can correlate the cloaking regions from a number of timestamps and can accurately pinpoint the location of user within a cloaking region. In Pingley [18], the need for a trusted third party anonymizer is overcome, and it is identified that location services are dependent of the degree of privacy protection and the context. Context aware privacy (CAP) is proposed in [18] to minimize the problems of getting information from the queries. They integrate the proposed technique with Google maps. In Gkoulalas [19], the underlying movement of users is determined by requests of users on trajectory data in location-based services. Reconstruction of movement is performed on the basis of location updates. As a result, routes that are sensitive to risks are identified, and then with the use of a spatial database engine, this is converted into anonymous form.

Ad hoc clustering techniques are also widely used to overcome tracking in location-based services [9], and they are an alternative to grid-based techniques. It creates ad hoc groups in which requests are provided by trusted nodes on behalf of the nodes that want to hide their locations. Chow [20] presents a peer-to-peer spatial cloaking algorithm in which a group is formed by mobile users within a single hop. They avoid multi-hop system because of security issues in ad hoc networks. The technique proposed in [20] shows that on-demand protocols results in lower communication costs with better quality of service than the proactive protocols. But response time in on-demand protocols is higher than that of proactive protocols [20]. In Magkos [21], wireless ad hoc network is used to hide the users requesting for queries. Their technique is applicable to both sporadic and continuous LBS queries [9]. In their technique [21], nodes that are within the ad hoc network are trusted and all nodes outside the ad hoc network are considered as threats.

Tracking is also overcome by using private information retrieval (PIR) protocols. Ghinita proposes an alternative to cloaking regions, where a hybrid two step approach is defined, in which the privacy of location queries is ensured. [4] In this technique, the location is initially generalized to coarse-grained cloaking regions, and, finally, the query is submitted to the LBS using a PIR protocol. In Yan [22], location services are classified, and a hierarchical distribution technique is used to support them. The location information with the keys is secured by using the hierarchical encryption. Finally, it is distributed to only trusted members.

Anonymity-based mechanisms are also used to overcome tracking in location-based services. Due to less integration of location-based services in IEEE 802.1x wireless networks, they are less exposed to risks than mobile networks. In spite of this, a number of details, which include MAC layer and IP layer, can be used for tracking a device. For that reason, the anonymity-based techniques are focusing on reducing the chances of mapping unique device identifier, for instance physical or logical addresses, in IEEE 802.1x networks. In this context, Gruteser [23] proposes a technique, which is further enhanced in Gruteser [24]. They [23, 24] present a technique, which identifies, assesses, and compares privacy of locations and risks associated with location tracking.

Finally, several works have explored group-based and in-network collaboration schemes. Cheng et al. [25] propose a protocol for group based location services for mobile ad hoc networks (MANETs), named as GrLS. In this protocol, group mobility is improved by using different location management strategies for singe nodes and for group of nodes. In this protocol, a single node recruits its own location services and performs location updates. In contrast, the group leader in a group of nodes is responsible for recruiting location servers and updating its location to a particular home region, referred to as group home region. This protocol significantly reduces overhead of the location service protocol. Yan et al. [26] explores issues related to privacy of customers and considers a number of security vulnerabilities of positioning Advanced Metering Infrastructure (AMI) in smart grid. They also study customer behavior and authentication of message for meter reading and control messages.

3. Novel Collaborative Technique using IDS and Rekeying for Preserving Privacy in LBS

In this section, we introduce the novel method for preserving the privacy of users who want access to secured, location-based services. The proposed algorithm extends the ad hoc clustering technique with network collaboration and by integrating it with efficient IDS and rekeying techniques. Based on these techniques, the tradeoff between privacy and quality of service is quantified in terms of performance and security metrics. Performance metrics are focused on service response time, while security metrics are defined here as mean time to security failure. These performance-security metrics can be used to identify optimal settings for the system under which mean time to security failure is maximized while meeting performance requirements.

The proposed algorithm is explained as follows;

- A collaborative platform for communication is proposed in order to minimize reliance on location-based server. Minimum reliance on LBS improves the security as nodes are less exposed to adversaries.
- All communicating nodes have the capability of communicating in an ad hoc fashion.

RESEARCH ARTICLE

- The communication in the GCS has no centralized control. This minimizes the chances of single point failure.
- Nodes are capable of forming the ad hoc network on the fly.
- An encrypted group key is shared amongst all the nodes in the group communication.
- Nodes can be part of group communication only if they have knowledge of the group key.
- Rekeying is performed to provide security from attacks that can occur from outside of the GCS.
- A new key (rekeying) is generated every time a node joins or leaves the GCS or if a node is detected and excluded from the GCS by the IDS.
- Once a secured group communication is established, the nodes can then demand their required queries.
- If a node has a query, it first broadcast its request to all the group members.
- Nodes contact the LBS only if the required information is not available amongst the neighboring nodes.
- In the event that users/nodes do not find the required information, then it can contact the LBS server. However, as the number of nodes in the GCS increase, they become less dependent on the LBS server and can often find query answers within the GCS.
- Nodes, who obtain location-based services either within the GCS or from the LBS, store the gained information in a buffer.
- Anytime a node within the GCS requests information found within the buffer, the buffered information is sent to the node which generated the query request.
- The LBS is not part of the group communication in ad hoc fashion; rather, access to the LBS server is achieved based on infrastructure of the available network.
- The group communication can be attacked from outside or inside the GCS. Outsider attacks are controlled through rekeying operations, while insider attacks are controlled by IDS operations.
- IDS is pre-installed in all communicating nodes to provide protection against attacks, which occur from inside the GCS. Once a node is detected as compromised or suspicious, it is evicted from the group communication system and rekeying operations are performed.
- Security failure occurs if a compromised node is not detected by the IDS and subsequently gains unauthorized access of data. Security failure also occurs if more than one-third of the nodes in the GCS become compromised (Byzantine failure model [27]).
- Two performance-security metrics, both based on the stochastic Petri net (SPN) model used in Cho [28], are used to identify optimal settings for the GCS that can maximize lifetime of system, whilst minimizing service response time, thereby providing better privacy and quality of service.

The flow chart shown in Figure 1 summarizes the proposed algorithm. A connection between various states of the system is presented.

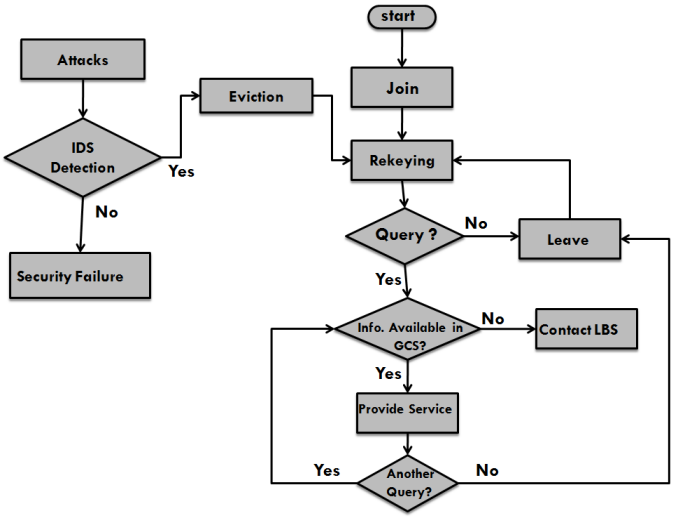

Figure 1System Flow Chart

3.1. IDS Operations

Intrusion detection system (IDS) operations provide security against insider attacks. An IDS is either a device or a software application that is responsible for monitoring system or network activities. It keeps the system or network safe from malicious activities. In case of anomaly detection or policy violation, the IDS generates reports to a management station. Normally, the IDS stores information that is related to observed events. In some cases, a detected threat can be responded by the IDS and thus preventing it from succeeding [29, 30]. We consider two common types of IDS, which are host-based IDS and voting based IDS. The former performs local detection in order to detect suspicious nodes while the latter provides robustness against collusion.

RESEARCH ARTICLE

In host-based IDS, each mobile node performs a local detection in order to detect suspicious nodes. Signature-based detection or anomaly detection [24] are standard techniques, which are commonly used to implement host-based IDS techniques. Route related and traffic related information is collected from the neighboring nodes, which are used to evaluate the nodes [24]. Two probabilities characterize host-based IDS, which are false positive and false negative, denoted by p_1 and p_2, respectively. False negative is the probability with which the IDS may not detect a compromised node in the GCS, while false positive is the probability with which the IDS may erroneously detect a correct node as compromised [28].

In Voting-based IDS, robustness is provided against collusion [24]. If a node is found suspicious or observed as abnormal, it is selected as a target node. All nodes in the GCS share information with each other and is based on routing, location, and identity number (ID). The node with the smallest ID is selected as a coordinator, who then selects m voting participants (including himself) to cast votes for or against the target node. The coordinator also broadcasts the list of m voting participants to all the nodes in the GCS. The eviction process is performed after regular intervals. The target node is either evicted or allowed to remain in the GCS based on the voting result. Like host-based IDS, voting-based IDS is also characterized by two probabilities, false positive (P_{fp}) and false negative (P_{fn}). These two probabilities are calculated based on three parameters: host-based false positive and false negative probability, the number of voting participants (m), and the approximation of current compromised nodes in the GCS. The equation for voting-based false negative and false positive probability is given in section 4, in which the whole algorithm is parameterized.

The working of the IDS is depicted in Figure 2. Figure (2a) shows a secure group communication among the nodes, as maintained via an encrypted group key. A node may become compromised, and the IDS may not detect it (false negative). This node may still exist in the GCS and can request for unauthorized access of data, as shown in Figure 2b. Figure 2c depicts the scenario in which the IDS erroneously detects a correct node as compromised (false positive). Figure 2d shows the case where the IDS correctly detects a compromised node. The compromised node is then removed from the GCS, as shown in Figure 2e. Finally, an encrypted key is generated and shared amongst all nodes in the GCS to maintain secure group communication. This is shown in Figure 2f.

3.2. Rekeying Operations

The group communication among mobile nodes seeking location-based services can be compromised from outsider attacks, as ad hoc networks have no strong line of defense [31]. Accordingly, rekeying operations are performed to

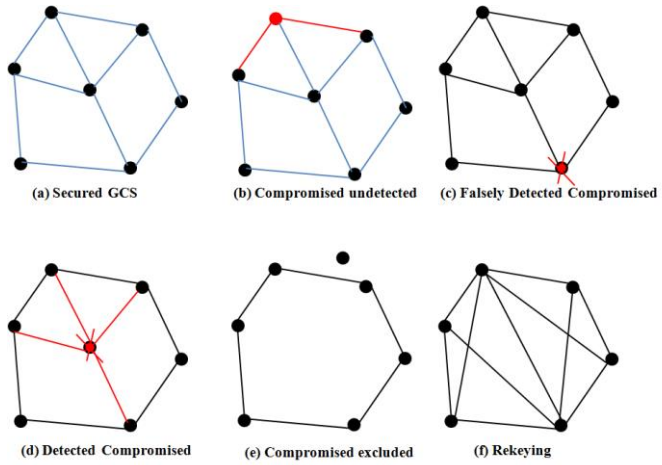

Figure 2: IDS operations

provide security against attacks that come from outside the group communication system. When external intruders attempt to gain unauthorized access to data, prevention methods, such as authentication and encryption, can control such attacks. Additionally and to assume the worst case scenario, a trusted node within the GCS can get compromised and can share the group key with outsider attackers. Thus, attacks that come from inside the GCS are also controlled by the intrusion detection system discussed in the previous section [31].

3.2.1. Rekeying Protocols

A number of rekeying protocols are discussed in [32-34], a few of which are taken under consideration in this paper, namely, individual rekeying, batch rekeying and interval-based rekeying. Among these three, individual rekeying is the considered the simplest protocol [34]. Individual rekeying protocol performs CKA [31] rekeying on every join, on every leave, or if a compromised node is excluded from the GCS. While great in theory, this protocol becomes prohibitively expensive due to the frequent rekeying. To avoid frequent rekeying and to minimize the cost of rekeying operations, threshold-based periodic batch rekeying protocols are proposed in [32]. These protocols are very useful to find optimal settings for the GCS in terms of both performance and security [28]. Two threshold-based rekeying protocols are Trusted and Untrusted Double Threshold-based Rekeying with CKA (TAUDT-C) and Join and Leave Doubled Threshold-based Rekeying with CKA (JALDT-C). In TAUDT-C protocol, CKA [31, 34] rekeying is performed every time two thresholds, k_1 and k_2, are reached, where k_1 represents the number of joins and leaves and k_2 represents the number of nodes that are detected as compromised by the IDS and excluded from the GCS [28]. JALDT-C protocol performs CKA [31, 34] rekeying every time thresholds k_1 and k_2 are reached. For JALDT-C, k_1 represents the number of nodes joining the system, and k_2 represents the number of trusted nodes that leave the GCS plus the number of nodes

RESEARCH ARTICLE

that are detected by IDS as compromised and evicted from the GCS [28]. Both in TAUDT-C and JAUDT-C, GDH.3 protocol [34] is considered for generation of secret key.

3.3. The Proposed System Architecture

Figure 3 gives a visual representation of the proposed algorithm. The architecture of the underlying LBS is kept unchanged, as any such changes may be not practical and difficult to adopt. Three components of the LBS are shown in Figure 3, which are mobile users, positioning technology, and location based service providers. Mobile users in a unit area form an ad hoc network in which security is maintained by using efficient IDS and rekeying techniques. In case a node does not find its required information in the GCS, its spatial location is determined by the positioning technology. Then, the query is sent to the location based service providers to provide services.

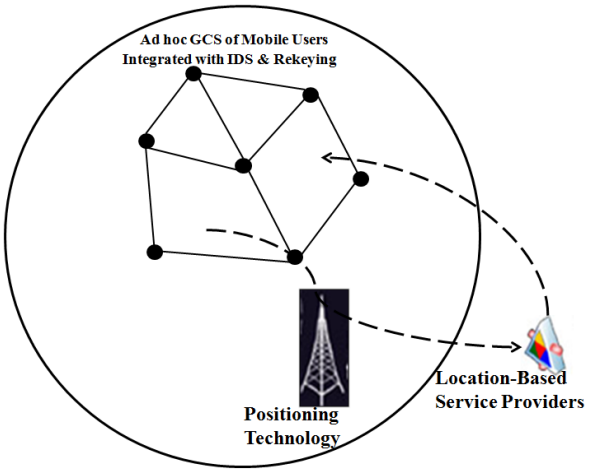

Figure 3: The Proposed System Architecture

4. PARAMETERIZATION

In order to parameterize the proposed algorithm, we consider several states of the system. The system can be in a number of states, which are as follows:

- The joining state: join requests come from mobile nodes to join the GCS to get location-based services. This state is denoted by 'J'.
- The leaving state: leave request come, and nodes leave the system intentionally after getting location-based services. Such leaves are stated as trusted leaves. This state is denoted by 'TL'.
- The trusted nodes state: represents the correct or trusted nodes in the system. This state is represented by 'Tn'
- The undetected compromised state: represents the situation in which a mobile node in the GCS can become compromised, and the IDS does not detect it (false negative). This state is denoted by 'Cn'
- The falsely detected compromised state: represents the situation in which the IDS detects a correct node as compromised (false positive). This state is denoted by 'FDC'.
- The detected compromised state: depicts the situation when IDS correctly detects a compromised node. This state is denoted by 'DC'.
- The security failure state: represents the situation when compromised nodes get unauthorized access of data or if more than one-third of the nodes in the GCS get compromised (Byzantine failure model [27]). This state is denoted by 'SF'.
- With the various system states defined, the goal is to identify appropriate rekeying intervals, which achieve optimal results, both in terms of performance and security. One solution can be found by considering the triggering conditions of the rekeying protocols discussed above.
- The triggering conditions for the rekeying protocols for the three protocols are as follows:

  1) For individual rekeying: rekeying is performed after the following conditions get satisfied [28]:

     $N(J) >= 1$ or $N(TL) >= 1$ or $N(FDC) >= 1$ or $N(TJ) >= 1$

  2) For TAUDT-C rekeying: rekeying is performed if the following condition holds [28]:

     $(N(J) + N(TL)) = k_1$ or $N(FDC) + N(DC)) = k_2$

  3) For JALDT-C rekeying: rekeying is performed if the following condition holds:

     $N(J) = k_1$ or $N(TL) + N(FDC) + N(DC)) = k_2$

  - k_1 and k_2 are two predefined thresholds,
  - N(J) denotes the number of nodes joining the GCS to get services.
  - N(TL) denotes the number of nodes leaving the GCS after getting services.
  - N(Tn) denotes the number of trusted nodes in the GCS.
  - N(DC) denotes the number of detected compromised nodes in the GCS.
  - N(FDC) denotes the number of falsely detected compromised nodes in the GCS.
  - N(UCn) denotes the number of undetected compromised nodes in the GCS.

- The communication time to broadcast a rekeying message is denoted by T_{cm}, which is calculated based GDH.3 protocol given by the following equation [34]:

RESEARCH ARTICLE

$$T_{cm} = \begin{cases} \frac{Nb_{GDH}(2H+1)-b_{GDH}(H+2)}{BW} & for\ N > 1 \\ \frac{b_{GDH}}{BW} & otherwise \end{cases}$$

- N = $N(T_n)$ + $N(UC_n)$, i.e. the number of current active nodes in the system.
- b_{GDH} = the length of an intermediate value.
- BW= bandwidth in Mbps.
- H = the number of hops between any two nodes.

- The behavior of insider attacker are modeled by linear time attacker function with rate $A(m_c)$ [27], which is given by [28]:

$$A(m_c) = \lambda_c \, x \, m_c \,,$$

where m_c denotes the ratio of current active nodes (N) to the number of good nodes in the system, and it can be calculated as follows [28]:

$$m_c = \frac{N(UC_n)+N(T_n)}{N(T_n)}.$$

- The undetected compromised nodes can get unauthorized access of data with rate as follows [28]:

$$p_1 \, x \, \lambda_q \, x \, N(UC_n) \,,$$

where λ_q denotes data packets issued by a node for group communication.

- The detected compromised nodes can get unauthorized access of data with the following rate [28]:

$$p_1 \, x \, \lambda_q \, x \, N(DC)$$

where λ_c denotes base compromising rate.

- The join and leave rates are represented by λ and μ.
- The false positive and false negative probabilities are defined by the following equation [28]:
- $P_{fn} or\ P_{fp} = \sum_{i=0}^{m-\lceil \frac{m}{2} \rceil} \left[\frac{\binom{N(UC_n)}{\lceil \frac{m}{2} \rceil + i} \times \binom{N(T_n)}{m-(\lceil \frac{m}{2} \rceil + i)}}{\binom{N(T_n)+N(UC_n)}{m}} \right] + \sum_{i=o}^{m-\lceil \frac{m}{2} \rceil} \left[\frac{\binom{N(UC_n)}{i} \times \sum_{j=\lceil \frac{m}{2} \rceil - i}^{m-i} \left[\binom{N(T_n)}{j} \times (p)^j \times \binom{N(T_n)-j}{m-i-j} \times (1-p)^{m-i-j} \right]}{\left(\binom{N(T_n)+N(UC_n)}{m} \right)} \right]$$
  - p_1 denotes host-based false negative probability,
  - p_2 denotes host-based false positive probability,
  - P_{fn} denotes voting-based false negative probability.
  - P_{fp} denotes voting-based false positive probability.
  - m denotes the number of voting participants,
  - $p = p_1$ for P_{fn},
  - $p = p_2$ for P_{fp}
- The IDS detects compromised nodes with linear time detection function with detection rate $D(m_d)$ [27].
  - $D(m_d)$ denotes detection rate (the rate at which the IDS is invoked), and is given by [28]:

$$D(m_d) = m_d / T_{IDS}$$

  - T_{IDS} denotes base intrusion detection interval.
  - m_d denotes the degree of nodes detected by the IDS, and it is given by [28];

$$m_d = \frac{N_{init}}{N} = \frac{N_{init}}{N(UC_n) + N(T_n)}$$

  - N_{init} denotes the initial number of nodes in the GCS.
- For voting-based IDS [28],

$$D(m_d) = N(UC_n) \, x \, D(m_d) \, x \, (1-P_{fn})$$

- The rate at which IDS selects correct nodes as compromised is given by [28]:

$$N(T_n) \, x \, D(m_d) \, x \, P_{fp}$$

5. OPTIMAL PERFORMANCE AND SECURITY

Thus far, we have discussed in detail the proposed algorithm for preserving privacy in location-based services, and we have parameterized each state of the system. We now consider how to measure performance and security of the group communication system of mobile, ad-hoc nodes who are generating queries in a collaborative fashion. We start by identifying the metrics of interest to measure performance and security of the system.

As discussed in section 3, the system reaches a security failure state if a compromised node gains access to unauthorized data or if more than one-third of the nodes in the system get compromised, as detailed in the Byzantine failure model [27]. In theory, a secure system should never arrive at the security failure state. Thus, the objective is to maximize lifetime of the system. We measure security in terms of the famous reliability metric, mean time to security failure (MTTSF). Maximum MTTSF equates to maximum lifetime of the system and, therefore, maximum security. Regarding performance, the interest of the nodes seeking LBS queries should be in how fast the system responds to their queries. The pertinent performance metric here is service response time (SRT). Following the parameterization presented in Section 4, we can know summarize the MTTSF and SRT performance metrics.

RESEARCH ARTICLE

MTTSF can be calculated as the total expected reward until the system reaches security failure state. Security failure occurs when either a compromised node gain unauthorized access of data or if more than one-third (Byzantine failure model [27]) of the nodes in the GCS get compromised [28]. SRT is the response time per group communication packet over the lifetime of the system. It can be calculated as total wireless contention delay and transmission delay [35] over MTTSF divided by MTTSF [28].

Using the above definitions, the performance/security metrics can be evaluated based on the parameterization discussed in Section 4. The optimal system settings can then be identified based on said performance and security metrics, with the focus on maximizing mean time to security failure and minimizing service response time. Maximizing MTTSF translates to the system being secured for maximum possible time, allowing mobile nodes to request queries without hesitation. Minimizing SRT simply means that if the user asks for a query and if the information exists in the group communication system, then the service will be provided at the fastest possible time. All in all, optimal settings for the system can be chosen that can provide both location privacy and quick service at the same time. Based on the optimal double thresholds (k_1 and k_2) for rekeying operations, the tradeoff between privacy and quality of service can be quantified, such that it is possible to identify such double thresholds (k_1 and k_2), under which MTTSF is maximized and SRT is minimized. Thus, better privacy can be achieved at better quality of service.

6. CONCLUSIONS

The ubiquitous nature of smartphones and GPS-enabled devices, coupled with the increasingly popular usage of location-based services, has effectively created an environment where data access truly is anywhere at any time. While said environment is indeed convenient and quite useful, the unfortunate reality is that users are exposed to a variety of privacy and security threats. A number of techniques that are currently used for preserving privacy in LBSs are reviewed. It is suggested that reliance on the LBS server should be minimized in order to provide better privacy as well as better quality of service. Towards this end, a new collaboration technique for preserving privacy in LBSs is proposed, and it is integrated with an intrusion detection system and rekeying techniques. It is shown that the collaborative network formed by ad hoc nodes can be merged with the existing LBS infrastructure. The algorithm is parameterized by defining performance and security metrics, which can then be used to find optimal settings, both in terms of privacy and quality of service. Based on the proposed performance-security metrics, the tradeoff between privacy and quality of service can be quantified.

REFERENCES

[1] Reza Shokri, George Theodorakopoulos, Panos Papadimitratos, Ehsan Kazemi, Hiding in the Mobile Crowd: Location Privacy through Collaboration. IEEE transactions on dependable and secure computing, special issue on “security and privacy in mobile platforms”, 2014.

[2] Mohamed F. Mokbel, Xiaopeng Xiong, Walid G. Aref, SINA: Scalable Incremental Processing of Continuous Queries in Spatio-temporal Databases. SIGMOD 2004, Paris, France, June 13-18, 2004.

[3] “Pleaserobme: http://www.pleaserobme.com.”, date accessed: 12-05-2014.

[4] Ghinita G, Kalnis P, Kantarcioglu M, Bertino E. A hybrid technique for private location-based queries with database protection. In: Mamoulis N, Seidl T, Pedersen T, Torp K, Assent I (eds) Advances in spatial and temporal databases. Lecture notes in computer science, vol 5644. Springer, Berlin/Heidelberg, 2009, pp 98–116.

[5] Mokbel MF, Chow C-Y, Aref WG. The new casper: query processing for location services without compromising privacy. In: Proceedings of the 32nd international conference on very large data bases, VLDB ’06, VLDB Endowment, 2006, pp 763–774.

[6] Jha, R.K.; Dalal, U.D., "Location based radio resource allocation (LBRRA) in WiMAX and WiMAX-WLAN interface network," in *Communication Systems and Networks (COMSNETS), 2012 Fourth International Conference on* , vol., no., pp.1-2, 3-7 Jan. 2012.doi: 10.1109/COMSNETS.2012.6151362.

[7] G.Rohini , A.Srinivasan. “Dynamic Transition of Bandwidth and Power Saving Mechanism to Support Multimedia Streaming Using H.264/SVC over the Wireless Networks”, International Journal of Computer Networks and Applications Volume 2, Issue 2, March – April 2015.

[8] Suresh Limkar, Rakesh Kumar Jha, “Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2012 (INDIA 2012) held in Visakhapatnam, India, January 2012, Volume 132 of the series Advances in Intelligent and Soft Computing pp 943-950.

[9] William J. Buchanan · Zbigniew Kwecka · Elias Ekonomou, A Privacy Preserving Method Using Privacy Enhancing Techniques for Location Based Services. Mobile Netw Appl, , 2013, 18:728–737.

[10] J. Meyerowitz and R. Roy Choudhury, “Hiding stars with fireworks: location privacy through camouflage,” in MobiCom’09: Proceedings of the 15th annual international conference on Mobile computing and networking, 2009.

[11] F. Olumofin, P. K. Tysowski, I. Goldberg, and U. Hengartner, “Achieving efficient query privacy for location based services,” in Privacy Enhancement Technologies (PETS), 2010.

[12] G. Ghinita, P. Kalnis, A. Khoshgozaran, C. Shahabi, and K.-L.Tan, “Private queries in location based services: anonymizers are not necessary,” in Proceedings of the ACM SIGMOD inter-national conference on Management of data, 2008.

[13] Chow C-Y, Mokbel M. Privacy in location-based services: a system architecture perspective. SIGSPATIAL Special 1, 2009, 23–27.

[14] Truong A, Truong Q, Dang T. An adaptive gridbased approach to location privacy preservation. In: Nguyen N, Katarzyniak R, Chen S-M (eds) Advances in intelligent information and database systems. Studies in computational intelligence, vol 283. Springer, Berlin/Heidelberg, , 2010, pp 133–144.

[15] Dewri R, Ray I, Whitley D. Query m-invariance: Preventing query disclosures in continuous location-based services. In: Eleventh international conference on mobile data management (MDM), 2010, pp 95–104.

[16] Damiani ML, Bertino E, Silvestri C. The probe framework for the personalized cloaking of private locations. Trans Data Privacy, 2010, 3:123–148.

[17] Ghinita G, Damiani ML, Silvestri C, Bertino E. Preventing velocity-based linkage attacks in location-aware applications. In: Proceedings of

RESEARCH ARTICLE

the 17th ACM SIGSPATIAL international conference on advances in geographic information systems, GIS '09. ACM, New York, NY, USA, 2009, pp 246–255.
[18] Pingley A, Yu W, Zhang N, Fu X, Zhao W. Cap: A context-aware privacy protection system for location-based services. In: ICDCS '09. 29th IEEE international conference on distributed computing systems, 2009. pp 49–57.
[19] Gkoulalas-Divanis A, Verykios VS, Bozanis P. A network aware privacy model for online requests in trajectory data. Data Knowl Eng, 2009, 68(4):431–452.
[20] Chow C-Y, Mokbel MF, Liu X. A peer-to-peer spatial cloaking algorithm for anonymous location-based service. In: Proceedings of the 14th annual ACM international symposium on advances in geographic information systems,GIS '06. ACM, New York, NY, USA, 2006, pp 171–178.
[21] Magkos E, Kotzanikolaou P, Sioutas S, Oikonomou K. A distributed privacy-preserving scheme for location-based queries. In: IEEE international symposium on a world of wireless mobile and multimedia networks (WoWMoM), 2010, pp 1–6.
[22] Sun Y, La Porta TF, Kermani P. A flexible privacy enhanced location-based services system framework and practice. IEEE Trans Mob Comput 8(3): 2009, pp304 –321.
[23] Gruteser M, Grunwald D. A methodological assessment of location privacy risks in wireless hotspot networks. In: Security in pervasive computing, volume 2802 of lecture notes in computer science. Springer, Berlin/Heidelberg, 2004, pp 113–142.
[24] Huang, Y. and Lee, W., "A Cooperative Intrusion Detection System for Ad Hoc Networks," Proc. 1st ACM Workshop on Security of Ad-hoc and Sensor Networks, Fairfax, Virginia, 2003, pp. 135-147.
[25] Hui Cheng, Jiannong Cao, Hsiao-Hwa Chen, Hongke Zhang, GrLS: Group-based Location Service in Mobile Ad Hoc Networks, IEEE Transactions on Vehicular Technology, Volume: 57, Issue: 6, Page(s): 3693-3707, 2008.
[26] Y. Yan, Qian, Y., and Sharif, H., "A secure and reliable In-Network Collaborative Communication Scheme for Advanced Metering Infrastructure in Smart Grid", in Wireless Communications and Networking Conference (WCNC), 2011 IEEE, 2011.
[27] Gärtner, F. C., "Byzantine Failures and Security: Arbitrary is not (always) Random," Technical Report IC/2003/20, EPFL, April, 2003.
[28] Cho, J.-H., Chen, I.-R. and Feng, P.-G. "Performance analysis of dynamic group communication systems with intrusion detection integrated with batch rekeying in mobilead hoc networks." AINAW '08: Proceedings of the 22nd International Conference on Advanced Information Networking and Applications Workshops, Washington, DC, USA, 2008, pp. 644-649.
[29] Scarfone, Karen; Mell, Peter (February 2007). "Guide to Intrusion Detection and Prevention Systems (IDPS)" (PDF). *Computer Security Resource Center* (National Institute of Standards and Technology) (800–94). Retrieved 1 January 2010.
[30] Ravi Shankar P, Santosh Naidu P. "A Dynamic Approach of Malicious Node Detection for Internet Traffic Analysis", International Journal of Computer Networks and Applications Volume 1, Issue 1, November – December, 2014.
[31] Brutch, P., and Ko , C., "Challenges in Intrusion Detection for Wireless Ad-hoc Networks," Proc. Symposium on Applications and the Internet Workshops, 27-31 Jan. 2003, pp.178 – 373.
[32] Cho, J.H., Chen, I. R., and Eltoweissy, M., "On Optimal Batch Rekeying for Secure Group Communications in Wireless Networks," ACM/Springer Wireless Networks, 2007.
[33] Patrick P.C. Lee, John C.S. Lui, and David K.Y. Yau, "Distributed Collaborative Key Agreement and Authentication Protocols for Dynamic Peer Groups," IEEE/ACM Transactions on Networking, vol. 14, no. 2, April 2006, pp.263-276.
[34] Li, X., Yang, Y.R., Gouda, M. G. and Lam, S.S. "Batch Rekeying for Secure Group Communications," Proc. of the Tenth Int'l Conf. on World Wide Web, Hong Kong, July 2001, pp. 525-534.
[35] Bianchi, G., "Performance Analysis of the IEEE 802.11Distributed Coordination Function," IEEE Journal on Selected Areas in Communications, vol. 18, no. 3, Mar. 2000, pp. 535-547.

Authors

Muhammad Jawad Ikram got his B.E degree in Computer Systems Engineering from University of Engineering and Technology, Peshawar, Pakistan in 2010. He received his M.Sc. Degree with Distinction in Networks and Performance Engineering from the University of Bradford, UK in 2012. He is currently PhD research scholar in Department of Computer Science, at King Abdul Aziz University, Jeddah, KSA. His current research interests include network security, performance modeling, GPU computing, and high-performance computing**.**

Jonathan M Cazalas received the MS and PhD degrees in computer science, both from the University of Central Florida, in 2009 and 2012, respectively. He joined King Abdul Aziz University in 2012 and is currently an Assistant Professor in the College of Computing and Information Technology. His research interests include location-based services, mobile computing, privacy and security concerns therein, and high-performance computation. He is a member of the IEEE and the IEEE Computer Society.