International Journal of Advanced Information Science and Technology (IJAIST) ISSN: 2319:2682
Vol.3, No.10, October 2014 DOI:10.15693/ijaist/2014.v3i10.67-75

Distributed Processing of Continuous Spatio-temporal Queries Over Road Networks Without Compromising Privacy

Suhelah Sandokji
PhD Student /Department of Computer Science
College of Computing and Information Technology
King Abdul-Aziz University

Jonathan Cazalas
Assistant Professor/ Department of Computer ScienceCollege of Computing and Information TechnologyKing Abdul-Aziz University

***Abstract*—Location-based services are increasingly becoming popular as mobile devices, such as smart phones and tablets, are proliferating the market. While users, are attracted to the usefulness and convenience of LBSs, their privacy is facing serious threats. In order to receive LBSs, users must continuously report their exact location to the LBS server. Entrusting this information to these servers opens the door to various threats. This paper considers a user's privacy concern. We present a distributed algorithm for processing continuous spatiotemporal queries on work networks without compromising privacy. A distributed version of the CASPER algorithm, where the Trusted A nonymization Server-Based Schemes is used. The algorithm also modifies over the Sallam et,al work "Distributed algorithm for processing continuous spatiotemporal queries over road network". As Sallam's et. al. work enhanced PLACE a distributed over IMA and QTP algorithms. The objective of our work is utilizing the conjunction between CASPER and sallam's et.al. work for processing continuous spatiotemporal queries in road network without Compromising Privacy.**

***Index terms*—continuous spatiotemporal queries, protection user Privacy,road network, spatiotemporal data server, Trusted Anonymization Server-Based Scheme.**

I. INTRODUCTION

With the worldwide proliferation of location-detection devices, such as GPS-enabled devices, and the staggering growth of smart phones [2], it is only a matter of time until these applications, often referred to as Location Based Services (LBS), truly become ubiquitous. [2].The uprising mobile devices popularity, accompanied with the dramatic growth in the LBS and wireless technologies, facilitate accessing data anywhere at any time. Given that marketing wireless communication has increased dramatically, along with a flood of location-based applications that have been invaded, it is critical and desirable, to answer users query in the real-time[1].

LBSs are based usually on the processing of spatiotemporal queries such as the range query and the nearest neighbor query. A range query retrieves all data objects located in the specific query region, for instance ''give me the name of all restaurants within 10 miles of my location."On the other hand the result in nearest neighbor query is the closest data object to a specific query point. Such as''find the closest supermarket or taxi from my location."

It becomes essential for researchers to provide efficient query processing techniques for spatiotemporal databases. Initially, the traditional database model that is designed for complex querying on persistent data storage are augmented by adding models and index structures geared to manage the locations of moving objects efficiently [3, 4, 1]. R-trees such asR*-tree X-tree, Lazy Update R-tree are examples for most popular mechanisms for spatial indexing. A variety of research has considered performance issues to support the ever-increasing number of continuous queries and moving objects. Some strategies focused on reducing the continuously computational load of the real-time queries monitoring and evaluating, over these mobile objects. Some examples of these are MQM [1] and MobiEyes [5]. Some researchers achieved better performance and throughput using distributed servers[6]. Scalability is also achieved using distributed computation as either relay on user mobile processors as in [1, 5], or using parallel computing utilized by GPGPU as in [2].

Despite all of these advancements, we still face many challenges due to privacy and accuracy concerns. Consequently, many other researchers have been focused on formulating methods to protect user's privacy.

User privacy protection is one of the significant concerns that must be addressed to gain user confidence regarding system usability and enabling the user to feel service provider trustworthiness.

Registered users with LBSs send data that reveal their position continuously to the server. When a user is requesting a query service, the location based server answers it based on the knowledge of the user's location that is continuously sent by registered users [7].Albeit these location-based query processing and the location-based applications, guarantee convenience and safety, with untrustworthy servers, the customer's privacy and security are threatened. The implicit view that most LBSs are based on, is that the users agree to reveal their private location. Users usually trade security and privacy with the LBSs. If a user is concerned about his privacy, they would turn-off their location device and (temporarily) unsubscribe from the service.[15]For example, by keeping track of the places a user visits, their lifestyle can

be revealed. With the knowledge of a clinic that a person visits, the personal medical records can be inferred. Locations of old friends can be tracked. In fact, in many cases, GPS devices have been used in stalking people [8]. The pseudonymityis not applicable in LBS [9] since the true identity is known directly by revealing the person's location. Nevertheless, an adversary who has compromised the LBS server, may keep tracking the location information surrounded by the LBS queries, in order to deduce sensitive privacy information about the user, such as their lifestyles, home locations, health conditions, and political/religious associations. For example, Hoh *et al.* [10] and Krumm [11] researchers investigated the possibility of inferring a driver's home location from collecting vehicle's GPS data while the location data was anonymized. Matsuo [12] inferred various personal information from knowing a user's indoor location data, such as age, coffee drinker or not, smoker or not, and work role. Furthermore, Gruteser and Hoh [13, 14] concluded that by using multiple hypotheses tracking (MHT),they were able to track individuals from completely anonymized GPS data.

Researchers have long been responsive to the LBS probable privacy risks. Many schemes have been suggested for user's protection [15] that are based on two approaches: *query privacy* (e.g., [16,28]) and *location privacy* (e.g., [12]).In query privacy, the LBS query attributes are the users' private information. On the other hand location privacy is related to any user's private information regarding to or inferred from their location information. Issues that researchers have studied are:

•Whether a malicious LBS provider can identify a user (i.e., deanonymized) or their location specifically.

• Whether a malicious entity can infer, using LBS query content -or location information, -a user's interest and/or habits, among others .Both approaches are related closely. If a malicious entity can accurately track a user then, the user may be relatively easily deanonymized. Also if a user is easily recognized, the location privacy would easily be compromised. This is because of the amount of available revealed information that can be used for achieving location-related inference attacks [15].Consequently, these two approaches of LBS privacy, achieved two types of privacy metrics. K-anonymity metric is the most accepted metric. It specifies that "a release provides k-anonymity protection if the information for each person contained in the release cannot be distinguished from at least k–1 individual whose information also appears in the release." [15]

The second approach is location entropy. This metric is used widely in the query results, for computing the uncertainty of the location information. Some researchers proposed other types of query privacy metrics. For example, congestion, ubiquity, and uniformity used in [17] to refine the anonymity set, that used in LBS queries as the location anonymity measurement. Regarding location metric Location entropy can also be used. Hoh and Gruteser focused on the accuracy of estimating a user's position by a malicious source when using the expected distance error [14].

Privacy protection schemes are classified into three most popular types: policy, location perturbation and obfuscation, and PIR based approaches.[15]. Strong enforcement approaches are ideal to achieve better protection, inspite of the policy-based schemes which have superior state of the art privacy protection technology. Using Trusted Anonymization Server-Based Schemes is one of the stronger enforcement approaches. Usually these servers are based on k-anonymity approach for location and query privacy. An adaptive interval cloaking algorithm is used to generate spatio-temporal cloaking boxes. Boxes that contain at least k_{min} users are used as location information, which is then sent to the LBS server. k_{min},is the global parameter, which indicates how well a user's location is cloaked. It is also the minimum size of the acceptable anonymity set.

Even though users are concerned about protecting their identities and personalization, they may also be concerned about accuracy results for public queries or public data. In contrast to focusing on privacy, other research challenges are the need to model location-dependent queries on real-world scenarios. While many works focus solely on GPS-enabled mobile-devices moving around randomly, most real-world scenarios will force the movement of users (devices) to be restricted to a transportation network (road network).In computing nearest-neighbor queries, an object that is closest to a focal point based on Euclidean distance may actually be farther away when constrained using the underlying road network. This means that the mobile objects are moving over a road network instead of Euclidean distance. Therefore, computing continuous spatiotemporal queries over road networks are still fundamental concerns for accurate query results in the field of spatial databases research .There have been many techniques for continuous kNN(CkNN) monitoring (e.g., [18,19]) aimed at Euclidean spaces. However, when we consider spatiotemporal queries over road networks, where most real-world scenarios are, the length of the shortest path is used to measure the distances between a query and a data object or between two objects. Typically, road segments are represented by the edges of the network. The edges' weights represent the road lengths or the time required to travel them. While several papers (e.g., [21]) discuss the snapshot k-NN queries in road networks from different points of view, ,few of the recently researches consider the continuous Monitoring (e.g., [23]).Salam et.al recent work, introduced a distributed algorithm for processing continuous spatiotemporal queries. Distributed query processing that involves multiple servers is inspired by (1) the need for scalability in terms of supporting a large number of moving objects and a large number of queries, and (2) the need for providing real-time answers to users' queries. To rapidly answer queries, incoming data streams are processed in-memory. The load is distributed among a set of regional servers that collaborate to continuously answer the spatiotemporal queries.

While Sallam's et.al. work was novel and significantly enhanced the research in spatial data bases domain, protecting the user's privacy was not considered in the query computation. We propose to augment this previous work by including privacy-preserving architectures between the querying objects and the location-based database servers. In order to protect user's privacy, in this research we are using CASPER system which is Trusted Anonymization Server-Based Scheme.

The motivation for our work is the need for protecting user privacy and to distribute the processing of the query in order to get the results in real time. This paper presents a distributed version of the CASPER algorithm, where the Trusted Anonymization Server-Based Schemes is used.

A. Work contribution

1. Introduce a distributed CASPER algorithm, the modified CASPER algorithm version over the Sallam et.al.Work "Distributed algorithm for processing continuous spatiotemporal queries over road network". As Sallam's et.al. work enhanced PLACE a distributed spatiotemporal data server over IMA and QTP algorithms.
2. Introduced modified IMA and QTP algorithm versions used in Sallam et,al. work over user cloaked area instead of point position. As distributed CASPER algorithm.

B. Objectives:

Our work is utilizing the conjunction between CASPER and Sallam's et.al. work for processing continuous spatiotemporal queries in road networks without compromising privacy .

Our enhancement would convince LBS users who need to protect their privacy, by securing their location information and giving them control over such information. Besides getting accurate answers regarding public data, as well as in situations when the accurate answers are preferable over the privacy concerns, the system would allow the user to reveal their information in their profile. The service providers, who surely have an incentive to eliminate or all eviate users' privacy concerns, are worried as user get comfortable is the successful marketing element in LBSs.

The remainder of the paper is organized as the following:

Next section is the related work .Followed by the proposed algorithm section. Finally the conclusion is presented.

II. RELATED WORK

The related work is presented in two subfields. First we discuss the spatiotemporal query over road network, followed by protection user privacy.

A. Spatiotemporal query over roadnetwork

Research has been conducted for processing continuous range and k-NN queries on a single-server. For example, CPM [24],YPK-CNN [18], and SEA-CNN [19] monitor exact Continuous k-NN's (CNNs) in the Euclidean space. In addition to monitoring CNNs, CPM supports aggregate nearest neighbor's w.r.t. and a set of query points. SEA-CNN monitors the NN changes assuming that the initial result is available and shares the execution and data structures among multiple concurrent k-NN queries. DISC [20] approximates k-NN queries answer by using "e distance units". The returned kth neighbor is farther from the focal than the actual kth NN by e distance units. SINA [21] uses a three-step spatial, to evaluate continuous range queries between moving objects and moving ranges. All the above algorithms assume that the objects move in Euclidean space; the constrained motion of objects is not taken into consideration. In [25], two algorithms for processing continuous k-NN queries on road networks are presented. In the first Incremental Monitoring Algorithm (IMA) [24], it consists of two steps; first the network is expanded around the issuer to find the k-NN, after computing the query initially, the shortest path between the issuer and the nodes is computed and stored in an expansion tree form to be used later for subsequent updates. Updates from objects and edges' weights falling in the expansion tree, can adjust the NN set of some queries. Then, regions of the affected queries are adjusted. In the second Group Monitoring Algorithm (GMA), the k-NN's of the intersections are monitored using IMA and are used to compute the results of all the queries in the path. IMA can be easily extended to support range queries.However, GMA focuses on storing pre-computed query results in network nodes to be used in solving subsequent k- NN queries and thus cannot be easily extended to solve other types of queries. MOVNet [22] also evaluates on MOVing objects which is the location-based snapshot queries in road Networks .The continuous range query algorithm used on top of MOVNet is C-MNDR (Continuous Mobile Network-Distance-based Range query algorithm) [23].

All the previous algorithms are using single-server, whilst in this paper we presented the distributed algorithm. There have been systems and algorithms, e.g., MQM [1] and Mobieyes [5], where they reduce the load from the server by utilizing the computing capability of moving object clients. By contrast PLACE [6] is a distributed spatiotemporal server. In PLACE network, regional servers cooperate to evaluate queries of moving objects. The mobile objects are hopping from one server to the next according to the coverage region of each server. Sallam'set. al. recent work [43]is an extension to both IMA and PLACE. In contrast to PLACE, Sallam.et. al. support road network constraints on the motion of objects. In contrast to IMA, instead of computing and storing the distance from the issuer to each node in the search tree as in IMA, they store the all-pairs shortest path matrix and the leaf nodes of the search tree. This paper augmented Sallam'set. al. work, by protecting users privacy.

B. Protecting user privacy

In general, LBSs have three approaches that are governed in

providing location privacy. First is spatial and temporal cloaking [26,27].This approach is based on sending an approximate location and time instead of the exact values to the servers. The purpose is to hide the user among k other users (called k-anonymity [26, 27]),and thus improving privacy. The drawback of this approach, however, injures the accuracy and the server responses time liness. Besides the important weakness where user privacy can be broken, there are several simple attacks on these mechanisms [29].However there are other cloaking mechanisms, such as Pseudonyms and silent times [31,32], in which frequently the identifiers of the device are changed along with no data transmissions for long periods at ordinary intervals. Consequently this harms functionality severely and disconnects users. Location transformation is the second approach. This approach protects user location privacy via transformed location coordinates. The challenge in this approach is accurately finding all the real neighbors .In [30] Hilbert Curves can find the approximate neighbors using blind evaluation. In order to find real neighbors, the proximity of transformed locations to actual locations are kept and incrementally processes nearest-neighbor queries.The other choice is using trusted third parties between clients and LBSA servers to achieve location transformation [43].The third approach is based on Private Information Retrieval PIR.This approach stillhas many barriers before it can be used for LBS in real time, even though its performance is improved by using special hard wares [43].

The trusted third party model that is related to the first approach is more suitable for real-time query processing as it requires less computation overhead .Next subsection will discuss metrics types used for trusted anonymization server with more details.

C. Trusted Anonymization Server-Based Schemes

Several metrics used for the architecture of centralized trusted server are: a mix zone, expected distance error and personalized *k*-anonymity model.[15]. Mix zones approach has been introduced by Beresford and Stajano [34]. In a mix zone,user can't update their location information..Each user gets a new pseudonym before leaving to other zone. By using Mix zones,adversaries can't refer new pseudonym to the old one. This makes it useful for protecting LBS query privacy.In order to evaluate the effectiveness of mix zones,the size of an anonymity set and a location entropy-based metric are used. There are some elements that affect the protection effect of mix zones, such as floor plan layout and the movement patterns of users. This show that using the location entropy provide a more accurate estimate of the available uncertainty, and should thus be considered a useful metric for designing privacy-protection systems.[34]

Second metric is the expected distance error that Hoh and Gruteser made use of[43]. The plan is to compute the accuracy with which an adversary can estimate a user's location, then using path confusion for privacy protection. The main initiative consist of letting more paths intersect with each of them in order to have at least two users' paths crossed via a perturbation algorithm in the trusted anonymization server, which confuses the adversary in matching the users paths. Both query and location privacy are protected using this approach. This way increases the distance error between users' accurate locations and an adversary's location estimations which further prevents adversaries from referring LBS users.

Gruteser and Grunwald research introduce the *k*-anonymity concept into the LBS privacy protection research community[38]. In their original work,a tuple consisting of three intervals ([$x1$, $x2$], [$y1$,$y2$], [$t1$, $t2$]) stand for the location information. The first two intervals symbolize a spatial area of user location, while the time for which the user is in the area is represented by the third interval. The spatio-temporal cloaking boxes are generated by the servers that process these data tuple via an adaptive interval cloaking algorithm .These cloaking boxes containing at least $k_{\min}$ users, sent finally to the LBS server as location information. $k_{\min}$is a global parameter, stand for the minimum size of acceptable anonymity set, demonstrating how well a user's location is hidden away. The weakness of single-point failures and the honesty and reliability of the centralized anonymization servers harms this solution .Furthermore, the global setting of $k_{\min}$ makes it hard for tuning the protection level. There is also a lack of guarantee for the resolution of the location information sent to the LBS server. Nevertheless, the lack of protection in sparse areas exists.[35].

Recently, researchers significantly over came the weakness. For example, privacy requirements are tackled to personalize LBS's users requests rather than the global setting in the anonymization server. CliqueCloak [33],as an example, is a personalized *k*-anonymity model where users are able to tolerate their maximum temporal and spatial resolutions and minimum level of anonymity. This work modifies the duty from finding cloaking boxes into finding cliques match certain conditions in the constraint graph as they modeled the anonymization constraints as a constraint graph. This way it collects together a set of users then constructs a clique graph to come to a decision whether certain users can allocate to the cloaked spatial area. These users' minimum bounding rectangle represents the cloaked spatial area. This approach suffers from two points. First, there is the limit of user number due to computation load of Clique graph. Secondly, some information about the possible user locations may be revealed, which threaten the privacy significantly.

CacheCloak scheme is another trusted server-based anonymization [29]. It protects location privacy in real-time and maintenance for the location accuracy from the LBS suppliers' perception. It caches LBS reply and uses them to respond to other users' queries later. LBS supplier are only able to detect interconnecting paths that CacheCloak estimates, but hardly can relate actual user's locations information which is with the query being send. CacheCloak's significant contribution is that it comes over the trade-off

between privacy protection and LBS seffectiveness that many researchers consider as the zero game. In spite of this, with the proliferation of the LBS market, the challenge questionably, is the scalability of the system, as this schema is based on using cached information for respondingto users' queries. The worst case causes the potential scalability bottleneck for the system. This is when each user demanded a special LBS sort,then the trusted server must cache from different LBS servers, an innumerable amount of information .However, this schema aspires to guard both location and query privacy, hence it improves the location anonymity and consequently renders the users' movement untraceable.[29]

Xu and Caialso introduced a feeling-based location privacy model[36].Using the entropy, they compute the popularity of a region, and they use a quadtree-style approach to avoid any referring of the LBS queries to precise users, [37]. Their work is a continuation to the previous work in [38], and it aided the expression of "users' intended protection level".T. Xu and Y. Caispot lighted that the *footprint*(which is the historical locations of different mobile devices)must betaken into account as well as their current locations, by the trusted server, to enhance the protection of users' location privacy. [38]Their modeling tolerates users who designate their preferred protection level. Users' location in LBS queries are the ones that they previously circled and designated as a comfortable region .Specifically, the popularity of a spatial region are identified by the authors as 2*E*, where *E* is the entropy computed using visitors' footprints inside the region. The popularity on a per-user basis is then calculated by the authors and therefore may further identify the *P*-popular trajectory (PPT) for each user according to the popularity configuration .This schema aspire to guard query privacy protection technique since via this approach no adversary would detect specifically the query sender in the selected cloaking set.[38].

Casper [27] is one of the attractive personalizable *k*anonymity based LBS-related privacy protection framework. A location anonymizer is located in a trusted server, and it permits each user privacy profile to be suitable for his desired needs. A user privacy profile consists of both the user's target of protection and privacy (i.e., the k value in *k*-anonymity) combined with the acceptable minimal location resolution (below which the user's privacy is considered compromised, even if the *k*anonymity condition is satisfied).Furthermore, Casper maintain the users' location information adaptively over an area in the anonymization server, since it is offered using incomplete pyramid structure [39],thus lowering both location update and cloaking costs. Nevertheless, another advantage of CASPER is addressing the query processing issue, for instance it uses location-based applications for the anonymous service. CASPER anonymizer differentiates itself from another trusted anonymization server-based scheme as it: (1) Offers a customizable privacy profile for each mobile user that contains the k-anonymity and minimum cloaked area A_{min} requirements, (2)Provided high scalability as it scales well to as huge number of mobile users as needed with random privacy profiles, and (3) unable to reverse engineer inorder to infer any information about the actual user location. *k*-anonymity-based approaches aim to protect LBS query privacy because the very nature of this metric is anonymity that is, how to make query senders in distinguishable. A. Deutsch *et al.* later proposed "policy-aware" *k*anonymity, defending against more realistic adversaries who are aware of the policy for cloaking box generation. [40]

Our work is toward distributing CASPER algorithm in order to decentralize anonymization server. The next section gives more details regarding the new version of the framework.

III. Distributed CASPER Framework

As distributed CASPER system consisted of distribution of two coupled components(data base servers),namely, the privacy-aware query processor and the location anonymizer. See fig1the location anonymizer (TS) is a trusted third party. It is considered the middle layer between location-based database server and mobile users. It is responsible for the following: (1) accept from the mobile users both the actual location information combined with a privacy profile of each user, (2) based on each user privacy profile, the actual location information is blurred into cloaked spatial areas, and (3) finally the cloaked spatial areas sent to the location-based data base server. The other component is the privacy-aware query processor (PAQPS). It is embedded inside the location-based database server to achieve its functionality in processing the anonymous queries and cloaked spatial areas rather than the exact location information. In each type of server there are data structures to present and save the data in the system. Next we define these data representation. [39]See fig1

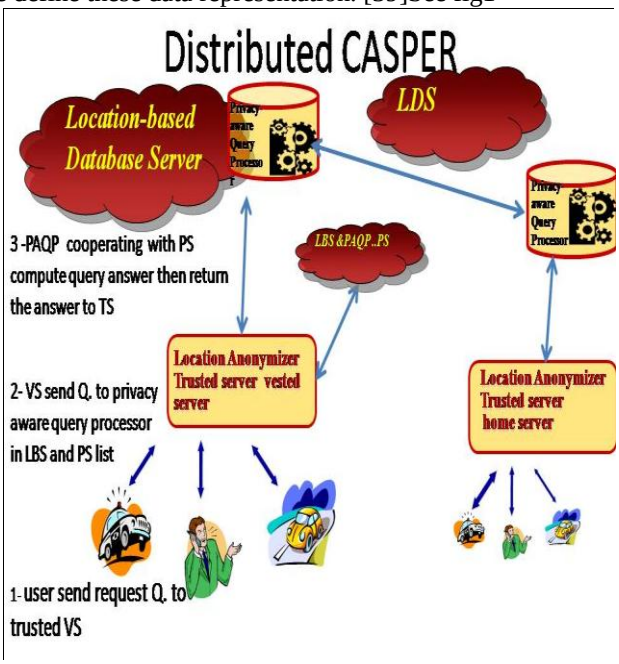

Figure 1Distrebuted CASPER servers connection

A. Data representation

Data in LOCATION ANONYMIZER server database

The road networks in all servers are represented as a bi-directional weighted graph. Each road is divided into cells or cloaked spatial areas that are contained in it. Each cell is represented as (cid, Rid, 2 path cost to both ends of road, N) where cid is the cell identifier. Rid is the road identifier that the cell is contained into. While N is the number of the mobile users within the cell boundaries. (All servers has this data based on their location). Edges in the graph represent roads, nodes represent road junctions (the intersection points between roads), and the weight of an edge is the length of the corresponding road. Also, we keep track of the object state that is saved in the hash table. Each registered mobile user has one record in this hash table. Each record has the form (uid, profile, cid,), where uid stands for the mobile user identifier which is unique.Profile is the user's privacy profile, which is in the form of tuple(uID, x, y, K,A_{min}). Cid represents the cell identifier in which the mobile user is located.Moving objects send their state periodically to the trustedserver. As such, their state is updated in the hash table. Any trusted servers are allowed to transfer object privacy data only to other trusted server, and are forbidden to submit to any other entrusted server.

B. Query representation

Query issuers can submit range and k-NN queries to the trusted server. A range query returns the list of candidate objects, distanced less than the query's anonymized spatial region's range from the query's focal point. To issue a range query, the object sends a request to one of the location anonymizer server database in the form (OID, 'R', R), where OID is the ID of the query issuer, 'R' is used to indicate a range query, and R is the required range.

A k-NN query returns the candidate list k objects that are closest to the query'sanonymize spatial region. To issue a k-NN query, the object sends a request to the trusted servers in the form (UID,'K',k), where UID is the ID of the query issuer, 'K' is used to identify a k-NN query, and "k" is the number of nearest neighbors requested.

For more efficiency, the costs of the shortest paths between all pairs of nodes in the network are pre-computed and stored in-memory.

Queries are continuous, i.e., they are stored in the system. Changes to the query answer as a result of the objects' movements, is done by privacy aware query processor in the visited server, and are sent progressively to the query issuer via visited server location anonymizer. Then this update is sent to the query issuer in the form of update tuples: (±, QID, CID), to add or remove anonymizedspatial region or cell whose ID is CID from the answer of the query whose ID is QID.

A range query returns list of thecandidate objects distanced less than the query range from the query's cell.

To issue a range query, the object sends a request to trusted servers in the form (UID, 'R', R), where UID is the ID of the query issuer, 'R' is used to indicate a range query, and R is the required range.

C. New object connection

1. Object sends connection request to →trusted Server combined with the private information in the tuple form (X,Y,K,Amin).
2. Trusted Server assigns UID and finds, via the user location x,y, its location defined by the road R and store it in → Hash table in memory and the object O state.
3. References in O's road are used to find the queries that O could be part of their answer.
4. The trusted server (TS) assigns the moving object to a road and calculates the cost of the path between the object's location and the start and end junctions of the road.
5. Objects' states (current road ID and path cost to road ends) are found by trusted server and blurs the location to cloak spatial areas that match each user privacy profile (k,Amin), (cell id)and sends the cloaked spatial area to the location-based database and it's embedded privacy aware query processor .
6. The location anonymizer in the visited server (visited trusted server(VTS(O)) receives continuous location updates from mobile users, blurs the location updates to cloaked spatial areas that match each user privacy profile (k,Amin), then sends the cloaked spatial areas to the location-based database server.
7. While cloaking the location information, the anonymizer also removes any user identity to ensure the pseudonymity of the location information and its embedded privacy aware query processor (VPAQP (O)).

D. Range Query evaluation

1. Object sends query request to visited trusted serverTS (O). Server retrieves O's state to get the cell id -his/her cloaked spatial area- (Cell id contains object query server id). As in fig.1
2. Requests transfers from the trusted server to Privacy-aware Query Processor in the object query server PAQPS (visited server)
3. Object's search starts from the end nodes of O's cell road and expands to their neighboring nodes
4. Those nodes are added to a queue. Nodes are removed from the queue one at a time;
5. Objects of public data or cells of private data on the roads connected to these nodes are added to the query answer and a reference to the query is added to investigate roads.
6. Nodes neighboring to visited nodes are added to the queue. Search stops when reaching nodes at a distance greater than the query range from the

issuer (focal) or when a loop in the graph is detected (when reaching a road that was visited before).

7. Nodes that the search stopped at (leaves of the search tree) are stored to help incrementally evaluate the query answer.

E. The distributed CASPER servers

To discuss query evaluation, there is a need to illustrate the types of the distributed severs. For each category of the following servers, there are a few of these two types of servers namely, the location anonymizer (TS) and the privacy-aware query processor server (PAQPS).

The entire road network is divided into Nsub regions that may overlap. N couple of servers exists; each couple of servers are responsible for one sub region. The Default Server (DS) is a specific server that objects initially send connection requests to. Each moving object is connected to one server (the trusted server part TS (O)) to which it sends its position periodically. This server is called O's Visited Server, VS (O).VTS (O) for the trusted part and VQPS (O) for the privacy aware query processor part or VPAQPS (O). The first server that O connects to is called O's Home Server, HS (O)(specifically HTS (O) part. As the object changes its position, it gets out of the range of its visited server; as a result, it disconnects from this server and connects to another one. HTS (O) keeps track of O's Visited Servers.Moving objects can issue range and k-NN queries. Aquery iscontinuously answered by a Querying Server and a set of Participating Servers (the privacy aware query processor part). For a query q, the Querying Server, QS (q), is the regional server that q's issuer belongs to, i.e., QS (q) =VS (spatial region query). A Participating Server for a query q, PS (q), is a regional server whose coverage region overlaps the search region of q.

F. Query evaluation

1. An object O issues a new query q to its visited trusted server. VTS(O)= QTS(q)
2. QTS(q)expands the search from O's position to finds the participated trusted servers, {PTS(q)}
3. QTS(q) → its Privacy Aware Query Processor QPAQPS(q)the specify q's anonymized spatial region, q's parameters and the positions (nodes) and object profilethe anonymized also removes any user identity to ensure the pseudonymity of the location information.
4. QTS (q) sends to a set of Participating Servers→{PTS (q)}: PS(q) is any server in which the spatial regions overlap QS(q) spatial region and participate in computing the query answer .QTS (q) specifies q's anonymized spatial region, q's parameters, the positions (nodes),the object profile , and PS(q)part of the query computation.
5. ∀ PS(q):QTS(q) sends to →PPAQP(q)q's anonymized spatial region, q's parameters, the positions (nodes),object profile, and PPAQP(q) that are part of the query computation. The anonymizer also removes any user identity to ensure the pseudonymity of the location information.
6. QTS (q) stores the query in a table and waitsfor PTS (q) respond.
7. ∀ PS (q):PTS (q) ←PPAQP (q) query result.
8. ∀ PS (q) : QTS(q) ←PTS(q) query result .
9. QTS (q) → O result is a list of candidate's answer after collecting them from servers and gets the final result based on object privacy profile.

G. Object position update

Continuously, the objects send position updates to visited server.

1. VTS check if O's position is in VTS (O)'s region.
2. If the O's new position lies outside VTS (O)'s region, VTS (O) sends the old state of O and queries that O is part of to the new Trusted server and sends a message to HTS(O) containing the ID of the new VS(O).
   That performs the following:
3. Locate O's new cell and the cost from O's cell to the ends of its road.
4. Send +ve updates to the QSs of queries that O entered their search range,
5. Send-ve updates to the QSs of queries that O is not part of in their answers any more.
6. Broadcast O's id and cell's id (O's road and cost from O to road ends) to querying servers of k-NN queries that O is part of.
7. The trusted querying server updates the weight of O in the answers of the queries.
8. Send a message to O to confirm the update, provide O with the new trusted server's connection details.

IV. Conclusion

The perception of a pervasive computing society has spurred a grand research interest in database systems for LBS. However, LBSs also pose a serious threat to users' privacy. An adversary can collect the service recipients' sensitive and privacy location information, which are embedded in the LBS queries if they managed to compromise the LBS server. In this paper, we addressed the challenge of processing of continuous spatiotemporal queries over road networks without compromising privacy. Our work is utilizing the conjunction between CASPER and Sallam's et.al. work for achieving two features :1)convince LBS users by protecting their location information privacy. This protection is flexible in blurring location information based on continuous ability of changing her/his profile, besides not blurring public data. 2) Distributing CASPER algorithm in order to decentralize anonymization server via parallel CASPER computation that would enhance bothof the scalability and performance . In the distributed CASPER algorithm, CASPER system consisted of a peer of two types of servers, namely, the location anonymizer servers, and the privacy-aware query processors that are embedded in

location based servers.

The location anonymizer is a trusted third party that acts as a middle layer between mobile users and one of the location-based database servers. Location anonymizer carry out two duties: the first duty is regarding user data.It receives the exact location information from mobile users along with a privacy profile of each user periodically, it blur it into *cloaked* spatial areas based on each user privacy profile, and sends it to the location-based database server to be saved. The other duty is receiving users query.It finds the other *participated trusted servers*. Then it sends them the user and her query with blurred information in the form of cloaked *spatial areas*. After that, the query answer tasks are parallel computed by privacy *aware query processors* embedded *LBS* of the *participated trusted servers.* Later, results are returned back to the first trusted server, which return it back to the user. The result retuned is *candidate list* objects. The paper illustrated more details regarding Distributed CASPER and different servers' duties algorithms.

In future work, we will implement the incremental processing of continuous spatiotemporal queries over road networks without compromising privacy algorithm, as well as the required measures for scalability, performance and privacy.

V .REFERENCES

[1] Ying Cai; Hua, K.A.; Guohong Cao; Xu, T. Real-time processing of range-monitoring queries in heterogeneous mobile databases. Mobile Computing, IEEE Transactions on, 2006.vol.5, no.7, pp. 931-942.

[2] Cazalas, J.; Guha, R., "GEDS: GPU Execution of Continuous Queries on Spatio-Temporal Data Streams," Embedded and Ubiquitous Computing (EUC), 2010 IEEE/IFIP 8th International Conference on , vol., no., pp.112,119, 11-13 Dec. 2010 doi: 10.1109/EUC.2010.26

[3] Guttman, A. 1984. R-trees: A dynamic index structure for spatial searching. In Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD). ACM Press, New York, NY, 47–57.

[4] Ying Cai; Hua, K.A.; Guohong Cao; Xu, T. 2006. Real-time processing of range-monitoring queries in heterogeneous mobile databases. Mobile Computing, IEEE Transactions on, vol.5, no.7, pp. 931-942.

[5] Gedik, B.; Ling Liu. 2006. MobiEyes: A Distributed Location Monitoring Service Using Moving Location Queries. Mobile Computing, IEEE Transactions on, vol.5, no.10, pp.1384-1402.

[6] Xiong X.,. Elmongui H, Chai X., Aref W., PLACE\: adistributed spatiotemporal data stream management system for moving objects, in: Proceedings of ACM Conference, 2007.

[7] Mokbel M. F. and Aref W. G.. PLACE: A Scalable Location-aware Database Server for Spatio-temporal Data Streams. IEEE Data Engineering Bulletin, 28(3):3–10, 2005.

[8] USAToday. Authorities: GPS system used to stalk woman. http://www.usatoday.com/tech/news/2002-12-30-gpsstalker x.htm.

[9] Pfitzmann A. and Kohntopp M.. Anonymity, Unobservability ,and Pseudonymity - A Proposal for Terminology. In Proceedings of the Workshop on Design Issues in Anonymity and Unobservability, 2000.

[10] Hoh B.et al., "Enhancing Security and Privacy in Traffic-Monitoring Systems," IEEE Pervasive Computing, vol. 5,no. 4, 2006, pp. 38–46.

[11] Krumm J., "Inference Attacks on Location Tracks," PERVASIVE'07, Proc. 5th Int'l. Conf. Pervasive Computing,Springer-Verlag, 2007, pp. 301–09.

[12] Matsuo Y.et al., "Inferring Long-Term User Properties Based on Users Location History," IJCAI '07 Proc. 20th Int'l. Joint Conf. Artificial intelligence, Morgan Kaufmann Publishers Inc., 2007.

[13] Gruteser M. and Hoh B., "On the Anonymity of Periodic Location Samples," Security in Pervasive Computing, vol. 3450/2005, Mar. 2005, pp. 179–92.

[14] Hoh B.and Gruteser M., "Protecting Location Privacy Through Path Confusion," IEEE SecureComm, 2005, pp 194–205.

[15] Shin, K.G.; Xiaoen Ju; Zhigang Chen; Xin Hu, "Privacy protection for users of location-based services," Wireless Communications, IEEE , vol.19, no.1, pp.30,39, Feb.2012 doi:10.1109 MWC.2012.6155874

[16] Gruteser M.and Grunwald D., "Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking," MobiSys, ACM, 2003.

[17] Kido H., Yanagisawa Y., and Satoh T., "An Anonymous Communication Technique Using Dummies for Location-based Services," IEEE Proc. Int'l. Conf. Pervasive Services, ICPS '05, July 2005.

[18] Yu X., Pu K., Koudas N., Monitoring k-nearest neighbor queries over moving objects, ICDE (2005).

[19] Xiong X., Mokbel M., Aref W., SEA-CNN: Scalable processing of continuous k-nearest neighbor queries in spatiotemporal databases, ICDE (2005).

[20] Koudas N., Ooi B., Tan K., Zhang R., Approximate NN queries on streams with guaranteed error/performance bounds, VLDB (2004).

[21] Mokbel M., Xiong X., Aref W., SINA: scalable incremental processing of continuous queries in spatiotemporal databases,SIGMOD (2004).

[22] Wang H., Zimmermann R., Snapshot location-based query processing on moving objects in road networks, in: Proceedings of ACM GIS, 2008.

[23] Wang H., Zimmermann R., Processing of continuous location-based range queries on moving objects in road networks, TKDE (2010).

[24] Mouratidis K., Hadjieleftheriou M., Papadias D., Conceptual partitioning: an efficient method for continuous nearest neighbor monitoring, SIGMOD (2005).

[25] Mouratidis K., Yiu M., Papadias, Mamoulis N., "Continuous nearest neighbor monitoring in D.road networks, VLDB (2006).

[26] Ghinita G., Kalnis P., and Skiadopoulos S., "PRIVE: Anonymous Location-Based Queries in Distributed Mobile Systems," Proc. 16thInt'l Conf. World Wide Web, 2007..

[27] M. F. Mokbel, C.-Y. Chow, and W. G. Aref, "The New Casper: Query Processing for Location Services Without Compromising Privacy," VLDB, ACM, Sept. 2006, pp.763–74.

[28] A. Pingley et al., "Protection of Query Privacy for Continuous Location Based Services," IEEE INFOCOM'11, Apr. 2011.
[29] P. Golle and K. Partridge, "On the Anonymity of Home/Work Location Pairs," Proc. Pervasive Computing, 2009. J. Meyerowitz and R. R. Choudhury, "Hiding Stars with Fireworks: Location Privacy Through Camouflage," MobiCom, ACM, 2009.
[30] Khoshgozaran A.and Shahabi C., "Blind Evaluation of Nearest Neighbor Queries Using Space Transformation to PreserveLocation Privacy," Proc. 10th Int'l Conf. Advances Spatial TemporalDatabases, 2007.
[31] Jiang T., Wang H.J., and Hu Y.-C., "Preserving Location Privacy in Wireless Lans," Proc. Fifth Int'l Conf. Mobile Systems, Applications Services, 2007.
[32] Beresford A.and Stajano F., "Mix Zones: User Privacy in Location- Aware Services," Proc. IEEE Second Ann. Conf. Pervasive ComputingComm. Workshop, 2004.
[33] Gedik B.and Liu L., "Protecting Location Privacy With Personalized k-Anonymity: Architecture and Algorithms," IEEE Trans. Mobile Computing, vol. 7, Jan.2008, pp. 1–18.
[34] Beresford A. and Stajano F., "Location Privacy in Pervasive Computing," IEEE Pervasive Computing, vol. 2, no.1, 2003, pp. 46–55.
[35] M. Gruteser and D. Grunwald, "Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking," MobiSys, ACM, 2003.
[36] T. Xu and Y. Cai, "Feeling-Based Location Privacy Protection for Location-Based Services," CCS, ACM, Nov. 2009, pp. 348–57.
[37] R. Finkel and J. L. Bentley, "Quad Trees: A Data Structure for Retrieval on Composite Keys," ActaInformatica, vol. 4, no. 1, 1974, pp. 1–9.
[38] T. Xu and Y. Cai, "Exploring Historical Location Data for Anonymity Preservation in Location-Based Services," IEEE INFOCOM '08, Apr. 2008, pp. 547–55.
[39] W. G. Aref and H. Samet, "Efficient Processing of Window Queries in the Pyramid Data Structure," PODS '90,Proc. 9th ACM SIGACT-SIGMOD-SIGART Symp. Principles of Database Sys., 1991.
[40] A. Deutsch et al., "Policy-Aware Sender Anonymity in Location Based Services," IEEE ICDE, Mar. 2010, pp.133–44.
[41] B. Hoh and M. Gruteser, "Protecting Location Privacy Through Path Confusion," IEEE SecureComm, 2005, pp.194–205.
[42] A.Sallam et.al., "Distributed processing of continuous spatiotemporal queries over road networks", Alexandria Engineering Journal, Vol. 51, Issue2,June2012,Pages85-93,ISSN1110-0168,http://dx.doi.org/10.1016/j.aej.2012.03.002.(http://www.sciencedirect.com/science/article/pii/S1110016812000440)
[43] Krishna P. N. Puttaswamy, Shiyuan Wang, Troy Steinbauer, Divyakant Agrawal, Amr El Abbadi, Christopher Kruegel, and Ben Y. Zhao. 2014. Preserving Location Privacy in Geosocial Applications. IEEE Transactions on Mobile Computing 13, 1 (January 2014), 159-173.