Enhanced Query Processing Technique for Location based Services

Srikanth. R
Department of Computer Science and Engineering,
National Institute of Technology Hamirpur,
Hamirpur (H.P.) India

L. K. Awasthi
Department of Computer Science and Engineering,
National Institute of Technology Hamirpur,
Hamirpur (H.P.) India

ABSTRACT
The emerging location-detection devices together with ubiquitous connectivity have enabled a large variety of location-based services (LBS). Location-based services are becoming popular for mobile users. The mobile users' location plays a key role to provide the service from one side, but it can be considered as a dimension of their privacy and so necessary to keep it anonymous to the other parties. Since one important issue is to achieve an accurate service, it is important to use the mobile's accurate location. Using the location accurately raises some concerns on behalf of the user's privacy. One solution for meeting this requirement is using by the means of a anonymizer. Anonymizer uses K-anonymity cloaking the user location to K-anonymizing spatial region (K-ASR). Traditional K-anonymity method needs complex query processing algorithms at the server side and have drawback of tracking user and path for mobile users. In this paper we are going to propose a new model for mobile users to retrieve the result quickly and increases users privacy.

Keywords
Location based services (LBS), privacy, K- Anonymizing spatial region (K-ASR), Anonymizer

1. INTRODUCTION
Nowadays, location-detection devices —such as cellular phones, GPS-like devices and RFID, etc—are more and more widely used. These location-detection devices together with ubiquitous connectivity have enabled a large variety of location-based services (LBS) which are able to tailor services according to the location of the user requiring the services. LBS can be used in a variety of contexts[6], such as health, work, personal life, etc. LBS include services to identify a location of a person or object, such as discovering the nearest banking cash machine or the whereabouts of a friend or employee. LBS include parcel tracking and vehicle tracking services. LBS can include mobile commerce when taking the form of coupons or advertising directed at customers based on their current location. They include personalized weather services and even location-based games.

There are several reasons that explain the great popularity of LBS in the research community, of these one of the main research area in LBS is privacy of the user. Unfortunately, LBS may threaten our privacy. Malicious attacker may collude with LBS provider to steal users' location information and query logs[1].

Such services mainly rely on *k*-nearest-neighbor queries (*k*NN) , which retrieve *k* points-of-interest (POIs) closest to the user's location. *K*-anonymity has been widely studied to protect privacy in LBS. Its main idea is to make the user issuing the query indistinguishable from at least *K*-1 other users. Most existing works [1][2][3][5], adopt the framework shown in Fig.1. The user sends its location, query and *K* to the anonymizer, which is a trusted third party. The anonymizer cloaks the exact user location to *K*-anonymizing spatial region (K-ASR) including at least *K*-1other users. Then the anonymizer sends the *K*-ASR and query to the LBS sever, which calculates the candidate results respect to the cloaked region and sends them back to the anonymizer. At last, the anonymizer calculates the actual results and sends them back to the user. Two serious drawbacks of this framework are: 1) high processing cost since the LBS server has to process range *k*-nearest-neighbor queries [1], and 2) high communication cost since the number of candidate results can be large.

Different from *K*-anonymity is SpaceTwist query processing[4] technique which sends a false location to the server instead of a cloaked region. SpaceTwist requires only simple query processing algorithm on the server—namely, incremental nearest neighbor (INN) retrieval. However, SpaceTwist may fail if the attacker already knows the locations of all the users. According to [4], the location of the user issuing the query can be bounded in a region Ω. If only one user lies in the region Ω, then attacker can easily infer that the query is issued by the user, which may threaten the user's privacy. The reason why SpaceTwist may fail is that it does not guarantee *K*-anonymity.

So far in the area of the LBS lot of research has been done in the privacy issues of the stationary LBS, there is only small research area in the field of the mobile client (eg Client moving from one city to other city) who dynamically moves from one cell network to other cell network randomly while moving to the destination. This is the main motivation for the work on this paper.

In this paper, our contributions are as follows:

- We propose a new framework to protect privacy in LBS. In this model we protect privacy and user traceability attack in LBS.
- We compare our model with traditional *K*-anonymity and SpaceTwist in terms of communication costs.

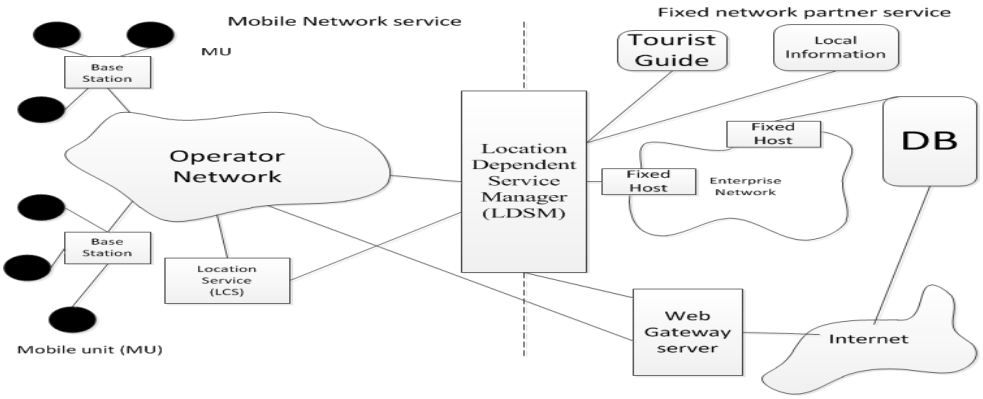


Fig-1: Architecture of the Location based service

The rest of this paper is organized as follows: Section II presents the related works. Next, in Section III, we introduce our approach. Our extensive surveys of various systems are illustrated in Section IV. Finally, in Section V, we conclude this paper and figure out our future works.

2. RELATED WORK

In our methodology, we draw upon work done in the areas of both cryptography and location-based services. In this section, we briefly go over related work in these areas.

2.1. Location-based services

In the peer-to-peer model of LBS, there would be a group of clients or peers who want to mutually compute some location-related function in the absence of a centralized trusted third party server. One of the common queries in this model are aggregate nearest neighbor or group nearest neighbor queries where the "nearest neighbors" are points of interest in the vicinity such as restaurants, hospitals, gas stations, etc. So a typical group nearest neighbor query for a group of peers would be "Which is the restaurant that is closest to all of us" or "Find a meeting spot that is within 2 miles of all of us?"

Figure 1, represents the general architecture of LBS. The left part of the figure is Mobile Network service which is connected by wireless network this part is responsible for collecting query from mobile users (MU). Location dependent service manager (LDSM) acts as the middle ware to the Mobile network service and service provider. The right part of the figure is the service provider which is connected by wired network.

Some of the general work in this direction (non-privacy preserving) includes, among others. One of the widely referenced works in this area is SpaceTwist [4] which is a protocol where a querying client asks the LBS server to return a set of *k* points closest to its own location or *k-N*earest neighbors. The authors do not apply their protocol to the group nearest neighbor problem which is quite different from *k*-nearest neighbor and has a different set of privacy requirements. There has been quite some work in the area of aggregate nearest neighbor queries which show how to do query processing of aggregate nearest neighbor queries in road networks.

2.2. Traditional K-anonymity

Most existing works on LBSs adopt *K*-anonymity by using the framework illustrated in Figure 2. This framework works as follows: A user sends its location, query, ID of the user and cloaks the exact user location to *K*-ASR including at least *K*-1 other users. Then anonymizer sends the *K*-ASR and query to the LBS sever, which calculates the candidate results respect to the cloaked region and sends them back to the anonymizer. At last, the anonymizer which knows the locations of all the users calculates the actual results and

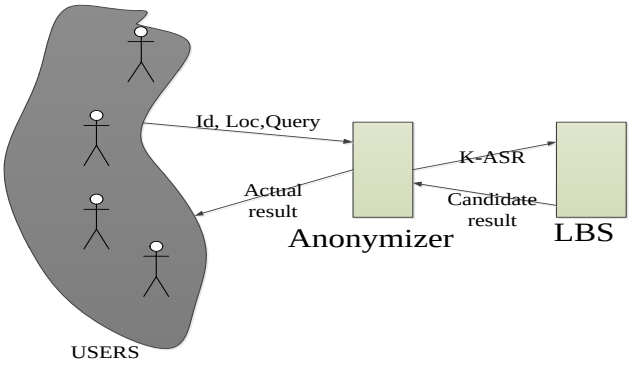


Fig-2: Architecture of traditional K- Anonymity

sends them back to the user. There are two drawbacks of this framework: 1) high processing cost at the server side since the LBS server has to process range *k*-nearest-neighbor queries, and 2) high communication cost since the number of candidate results can be large.

2.3 Public-Key Cryptography

One drawback of a private-key (traditional cryptography) system is that it requires the prior communication of the key K between Alice and Bob, using a secure channel, before any cipher text is transmitted. In practice, this may be very difficult to achieve.

The Figure 3, describes the public key cryptography system. The idea behind a public-key system is that it might be possible to find a cryptosystem where it is computationally infeasible to determine. If so, then the encryption rule could be made public by publishing it in a directory (hence the term public-key system). The advantage of a public-key system is that Alice (or anyone else) can send an encrypted message to Bob (without the prior communication of a secret key) by using the public encryption rule. Bob will be the only person that can decrypt the cipher text, using his secret decryption rule.

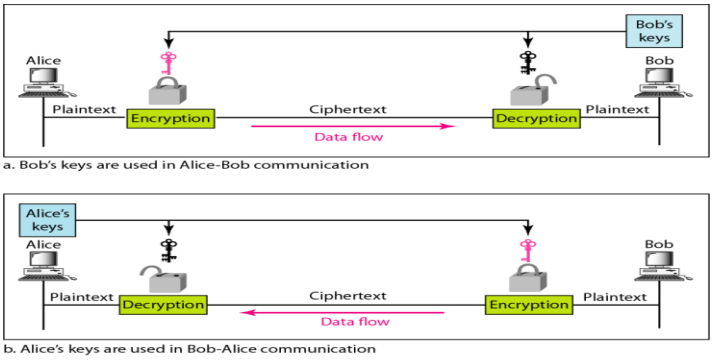


Fig-3 Description of the public-key system

3. SYSTEM DESCRIPTION

In this paper we have proposed a new technique to process LBS query for mobile users, who frequently change their position. Earlier system's as described in section 2 Traditional k- Anonymity uses a single server and multiple anonymizers, hence the workload of the server increases and traceability of the users is increased as there is only server preserve track of all the mobility of the user. To overcome these issues we have developed a new system as described in following paragraph.

3.1 System environment:

In the proposed system environment, we have considered the base station (BS) as LBS server. The BS which basically serves the Voice and data communication can also simultaneously serve as the LBS application server, by using BS station as the LBS server there is no need of setting up another environment for handling the LBS application. Distribution of the LBS server in various cells is shown in the figure- 5(a).

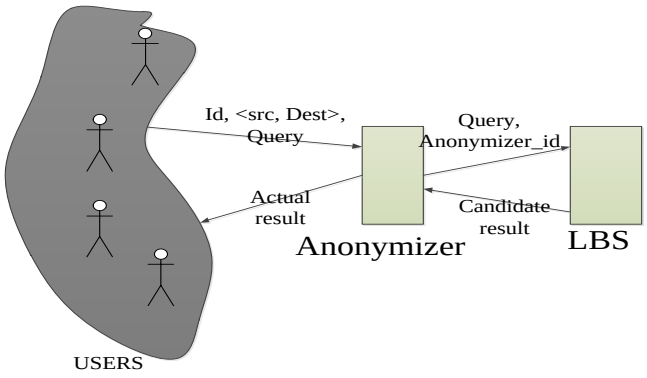


Fig-4: Architecture of proposed model

Each LBS server will be storing the Point of interest data such as shopping malls, petrol pump, and all information that normal user queries in location based service this information is restricted to particular area of the server which it is covering.

The proposed architecture of the system is as shown in the Figure 4, proposed system is divided into three categories here we have a user, anonymizer and server. In this system we have also distributed anonymizers according to the areas as shown in the Figure 5(b). We have distributed anonymizer into area and theses anonymizers can receive the data and sense its local environment of user mobility and update the required information to the server/user depending on the situation.

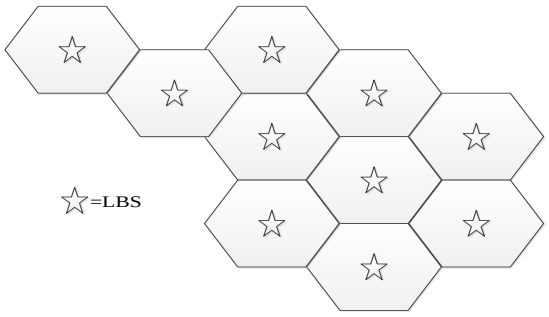


Fig-5(a) Deployment of LBS server over a geo-graphical Area

3.2 Proposed model:

The proposed system is as shown in Figure 4, we have a group of mobile users who issues the LBS query, these user frequently change their position as per their requirements. Mobile user system has the CA(coordinate Agent) , QA(Query Agent). CA is used to get the current time, Location of user Loc(Long, Lati.). QA is used to collect the information of the query and other relevant information of the query.

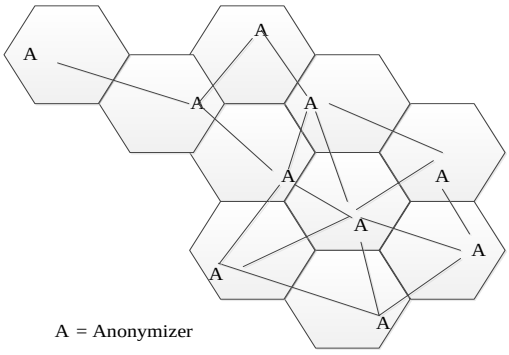


Fig-5(b) Deployment of Anonymizers over geo-graphical area

Query issued by the mobile users include the following information:

1. id – Mobile user id who issues the query
2. <Source, Destination>-- specifies the source point and destination point
3. Query – Query depends on the location
4. Public key of the Mobile user

Above information is delivered to the anonymizer by encrypting with the public key of the Anonymizer, so that information delivered to the Anonymizer is confidential and secure. Anonymizer stores the above information in the database.

Anonymizer forward's the above query information to the server as the following query:

1. Anonymizer_id – Id of the Anonymizer
2. Query – Query forwarded by the user.

Above information is delivered to the sever by encrypting with the public key of the server. Server opens the information using its private key. At the server side the query is processed by the server, and returns the result back to the anonymizer. Server knows the information of the anonymizer so it encrypts with the Anonymizer public key and forward the result Anonymizer. It forwards the result to the desired user based on the id by encrypting the result by user's public key.

In Figure 6, we illustrate a working environment for the proposed system, number in the cell reference the cell-id and the path of the user is calculated by the anonymizer as the user specifies the source station and the destination station. By this the anonymizer knows the cell through which the user travels, so the anonymizers

After getting the result from the server, anonymizer sends the result to the mobile user, based on his position in the cell, if the user moves from one cell to another then source node passes the query, <src, dest> and user id to next sequence anonymizer which comes in the path. When the source anonymizer passes the information to other anonymizer then it changes the source station of the query to the next anonymizer of the present source station. By this help the tracking of the user probability reduces.

Even if the server is the attacked by the intruder, he can't trace the path, and identify the user who had issued the query, he only knows which anonymizer issued query and can't identify the user who issued the query.

Table-1: Database at the anonymizer

Public_key_of _user	Location__user_last_found		Time_ of_issue _Query	Source	Destination	Query	Query_result
	Longitude	Latitude					
1A122-3343f-….			10:30:45	ABC	ASD		
2de67-16281-….			10:35:00	CDE	QWE		
:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:
45452-45df2-….			11:45:42	DFG	ZXC		

of the source station shares the mutual information between the all the anonymizers that fall in the path. Anonymizer is the trusted third party. Through which user can share information, LBS queries.

After getting the result from the server, anonymizer sends the result to the mobile user, based on his position in the cell, if the user moves from one cell to another then source node passes the query, <src,dest> and user id to next sequence anonymizer which comes in the path. When the source anonymizer passes the information to other anonymizer then it changes the source station of the query to the next anonymizer of the present source station. By this help the tracking of the user probability reduces.

Even if the server is the attacked by the intruder, he can't trace the path, and identify the user who had issued the query, he only knows which anonymizer issued query and can't identify the user who issued the query.

In Figure 4, we illustrate a working environment for the proposed system, number in the cell reference the cell-id and the path of the user is calculated by the anonymizer as the user specifies the source station and the destination station. By this the anonymizer knows the cell through which the user travels, so the anonymizers of the source station shares the mutual information between the all the anonymizers that fall in the path. Anonymizer is the trusted third party. Through which user can share information, LBS queries.

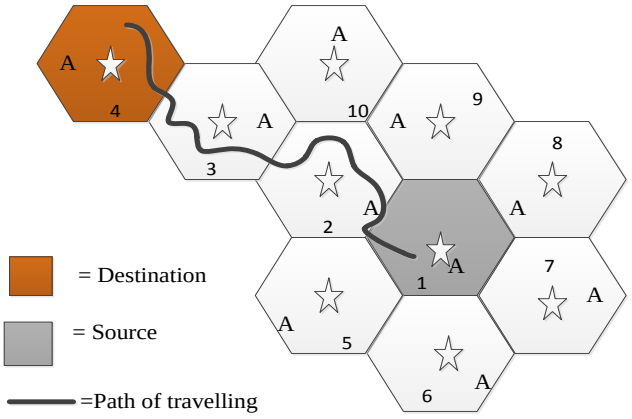


Fig- 6: Map of a mobile user for proposed system

4. SIMULATION RESULTS

Next part of simulation is deployment of the mobile nodes. Mobile nodes deployed in the network may either be stationary or mobile. Mobile nodes issue the query about LBS to the anonymizer; anonymizer sends the query to the server. Server receives the encrypted vehicle *id* to all the anonymizer falling under its coverage area. Anonymizer then adds the *id* of vehicle for sensing the locality, if the anonymizer finds the *id* it updates the location of the vehicle to the server so that server computes accurate result invalidating the

previous location. Server sends better accurate results to the anonymizer.

Parameters that we have used during our simulation are given in Table 2 below.

Table 2 Parameters used for simulation

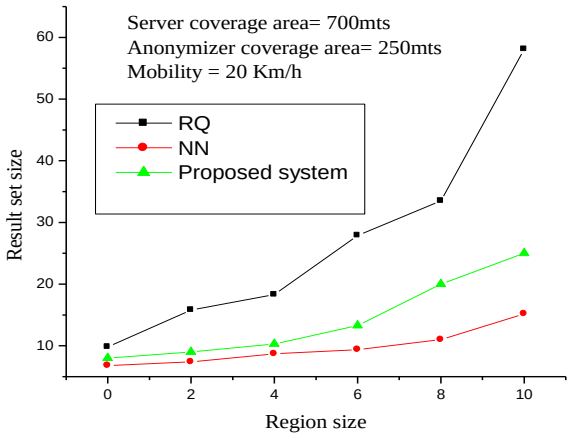
Parameter	Value
Network size (km)	3 Kms
Diameter of server coverage (m)	500-1000 mts
Diameter of anonymizer coverage (m)	200-400 mts
Users mobility (km/h)	0 to 60 km/h
Number of mobile users	21
Number of Anonymizer	14
Number of Servers	4

smaller area then the result delivered by the server is limited to the small location if the area size covered by the anonymizer is increased then the output of the query set size increases. The comparative study of our study on result data set size of proposed scheme along with the RQ and NN scheme is shown in Figure 7(a, b).

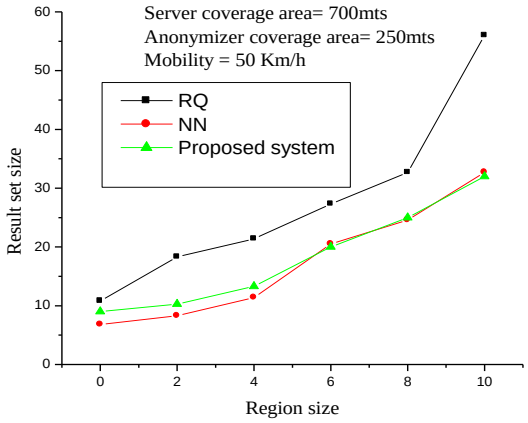
In this condition we have set the range of server to 700mts, whereas the range of anonymizer is set to 250mts. In Figure 7(a), mobility of user assumed to be 20 Km/h, then the result set size of the RQ scheme varies from 9.8 to 58.1 whereas NN scheme result size varied from 6.8 to 15.2 and in proposed scheme the result size varies from 8 to 25 varying on the size of the cloaked region from 1 to 10. In Figure 7(b), mobility of user is increased to 50 km/h, then the maximum result set size of RQ, NN and proposed scheme is 56, 32.7 and 32, when the cloaked region size is 10.

4.2 Impact on Query Processing Time

Query processing time is the time taken to process the query issued by the user. The RQ system processes the resultant POI which falls in the given range from the user present location, if the POI is not present in the range specified then network range is increased and queried from the starting. The NN system processes the query until the given set number of POI's matches with the user defined range.



(a)



(b)

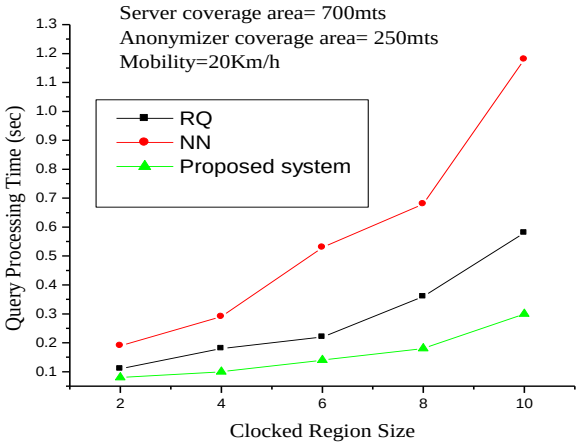
Fig-7: Influence of region size on result set size

After simulating the proposed system we simulated the proposed scheme and derived various set of results. The performance of the proposed system is done under various parameters some of the parameters include: Result set size, query processing time, bandwidth consumption, CPU time etc.

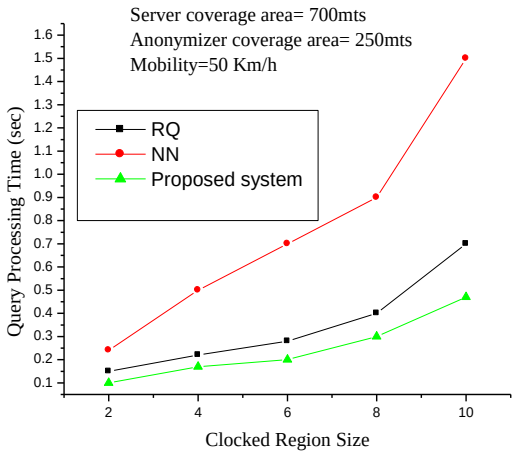
4.1 Impact of Result set size

Result set size delivered by the server to the anonymizer is dependent on the region size of the POI located in the particular area. In the Range Query (RQ), the query is processed based on the region size, as the region range increases the output of the query by the server increases, where as in the Nearest Neighbor (NN) scheme result is based on the number of the neighbors around the surrounding area, the result set size of the NN is not much affected by the size of the area, but affected by number of POI's distributed over a particular area. As the proposed scheme is based on the division of the area into smaller group, if the anonymizer covers the

Time taken to process the given limit of POI by the user takes huge time. The proposed system just checks the possible POI's which fall under the anonymizer area under which user is present and returns back's the result to the user. Figure-8 represents the query processing time taken by the RQ, NN and proposed system.



(a)



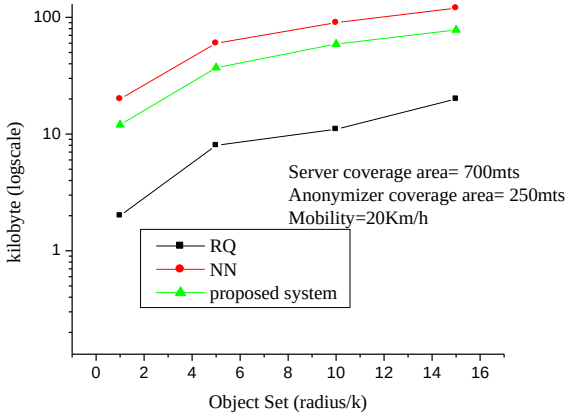
(b)

Fig-8: Influence of region size on query processing time

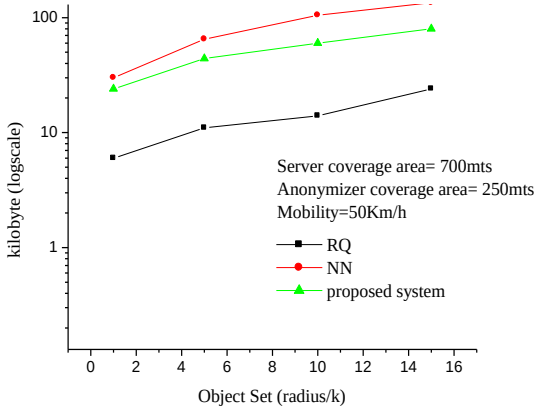
In this condition we have set the range of server to 700mts, whereas the range of anonymizer is set to 250mts. In Figure 8(a), mobility of user assumed to be 20 Km/h, then the query processing time of RQ scheme varies from 0.11 to 0.58(sec) whereas NN scheme result size varied from 0.19 to 1.18(sec) and in proposed scheme the result size varies from 0.08 to 0.3 varying on the size of the cloaked region from 2 to 10. In Figure 8(b), mobility of user is increased to 50 km/h, then the maximum query processing time taken by RQ, NN and proposed scheme is 0.7,1.5 and 0.47(sec), when the cloaked region size is 10

4.3 Bandwidth utilization

Bandwidth utilized by various systems is as shown in Figure 9(a, b). Bandwidth utilized by the system's are computed in Kilobytes based on the object set distribution over the given space. RQ system utilizes lesser bandwidth as compared to NN and proposed system because in this system the server computes POIs, based on the region of fixed size from the user location, after that it stops processing. Whereas NN system continues the search until it reaches the specified limit given by the user. Proposed system search is limited but there is frequent communication between the anonymizer and the server hence the bandwidth consumption is slightly higher than RQ. In this condition we have set the range of server to 700mts, whereas the range of anonymizer is set to 250mts. Other parameter which we used here are range or *k* ie user specified range for RQ scheme and k for NN scheme. In Figure 9(a), mobility of user assumed to be 20 Km/h, then the query processing time of RQ scheme varies from 2 to 20(KB) whereas NN scheme result size varied from 20 to 120(KB) and in proposed scheme the result size varies from 12 to 78(KB) varying on the object set distribution from 1 to 15(range or *K*). In Figure 9(b), mobility of user is increased to 50 km/h, then the maximum bandwidth utilization by RQ, NN and proposed scheme is 24,135 and 80(KB), when the cloaked region or *K* is 10.



(a)



(b)

Fig-9 Bandwidth consumption of various query processing techniques.

This paper has mainly focused on user privacy for mobile users in LBs environment. Proposed system performs better than other system since our system uses new model for query processing and is not based on K-ASR so drawbacks of K-ASR have been eliminated and moreover the user path tracking attack has been overcome since the server doesn't store the information regarding user and his activities. Server load in the proposed system is very low since here we had deployed many servers which store information only regarding that particular area and can only process that small area than compared to existing system which store information regarding large geographical area. Our model illustrate that the query processing costs of our proposed system are lower than those of TKA, KAWCR, Space Twist because in our model the database of the whole region is stored in a place there is no need to split the query based on the dependencies of various databases.

Table-2: Analysis of various query processing system

	Proposed System	**KAWCR**	**TKA**	**Space Twist**
User Privacy	High	Medium	Low	Medium
Storage at Anonymizer	Medium	Medium	Medium	Medium
Query processing cost	Low	High	High	Medium
Communication cost	Low	Low	High	High
Server load	Medium	Medium	High	High
Traceability attack	Low	Medium	High	Medium

Drawback of proposed system is requirement of storage capacity at the anonymizer is high than compare to the previous model.

5. CONCLUSION

In this paper, we propose a new framework to protect privacy in location-based services for the mobile users. Proposed system has better security and has strong defense system from the attack of the attacker. The main strength of the system is that the system gives accurate result quickly and does not store the information about users who has accessed the system so that if intruder hacks the system he does not get any information regarding user and his identity. In this paper we had seen that the proposed system have better performance than compared to the earlier existing system. In future we would like to work on the only drawback of the storage system at the anonymizer.

6. REFERENCES

[1] P. Kalnis, G. Ghinita, K. Mouratidis, and D. Papadias, "Preventing Location-Based Identity Inference in Anonymous Spatial Queries," In *IEEE TKDE*, 2007.

[2] G. Ghinita, P. Kalnis, and S. Skiadopoulos, "Priv´e: Anonymous location-based queries in distributed mobile systems," In *WWW*, 2007.

[3] C. Zhang and Y. Huang, "Cloaking Locations for Anonymous Location Based Services: A Hybrid Approach," In *GeoInformatica*, Vol.3, No.2, pp.159-182, 2009.

[4] M. L. Yiu, C. Jensen, X. Huang, and H. Lu, "Spacetwist: Managing the trade-offs among location privacy, query performance, and query accuracy in mobile services," In *ICDE*, 2008.

[5] T. Wang and L. Liu, "Privacy-aware mobile services over road networks." In VLDB, 2009.

[6] Gartner, "Gartner Identifies the Top 10 Consumer Mobile Applications for 2012" http://www.gartner.com/it/page.jsp?id=1230413, April 2010.