GRSM: SECURE GEOMETRIC RANGE SEARCH OVER ENCRYPTED SPATIAL DATA

Naveen Kumar C.G[1], Sanjay Pande.M.B[2]

[1]Research Scholar, Department of Computer Science, Bharathiar University, Coimbatore, India.
[2]Professor& Principal, Sampoorna Institute of Technology & Research, Ramanagar, Karnataka, India.

Abstract

Nowadays, outsourcing the data to the cloud server is a natural activity auctioned by several cloud users. The outsourced data may contain sensitive information. The cloud technologies are very much improved that attracts many Location based Services (LBS) companies. The overall theme of the cloud data resides at the distant server is to be managed with minimal computation by data owner and data users. The data is hived away in encrypted form to prevent anonymity activities. Reachability is one of the issues faced between cloud users and LBS companies. A novel and lightweight scheme, named, Geometric Range Search Model (GRSM) that retrieves the search data from ciphertext dataset. The data is considered as points and the group of points denotes the ciphertext database. Bloom filter is the filter that contains all possible combination of search tokens. The proposed GRSM contain three phases, namely, Encryption phase, Token generation phase and Search phase. Each phase serves as input/ output to retrieve the search data. An experimental result shows the effectiveness of the proposed algorithm.

Keywords: *Data Outsourcing, Cloud Technologies, Location Based Services (LBS), Bloom Filter and Search Tokens.*

--***--

1. INTRODUCTION

With advancements in the cloud computing, greater demands arise for the storage and security analysis. The origin of the cloud computing is from the development of parallel computing, distributed computing, grid computing with the evolution of virtualization and utility computing. Generally, the services of cloud computing is segmented into Infrastructure as a Service (IaaS), Software as a Service (SaaS) and Platform as a Services (PaaS) [1]. It works on the basis of 'Pay-per-use-on' model which can conveniently access the shared IT resources via web modules. The cloud technology offers unlimited resources and services to manage the outsourced data. With further improvements, the support of dynamic data is initialized. These features attracted the cloud users to store massive amount of information to this system [2].

Searchable encryption [3] is the stereotype of encryption technique which most recently studied by research community. It executes search operations on the encrypted databases. By doing that, the privacy of the data should be obtained from semi-trusted third party service providers. The cloud users are unaware about their data location. Thus, the client performs search operation over the server and obtains its results. Prior work depicts that the server's execution from the result set of encrypted documents and its security parameters like data dimension and documents. Better the privacy design, better the search operation. Reachability is one of the parameters that depict the available of the resources over given time e.g. social behavior analysis, recommendations model, public services etc. The location of the user is dynamic one. The user's location can notified from Location Based Services (LBS) [4] such as Google Maps, Foursquare etc. This scenario motivates to study about the reachability analysis of cloud data. Since, the data volume increases, the need for LBS companies are increased.

In order to provide better information retrieval process, the security and privacy issues should be devised properly [5]. Though the concept of searchable encryption is examined previously, the security and privacy challenges are not yet accomplished. Some additional security index was used for the data search process. In this paper, we suggest a geometric range search process over the encrypted data, so as to enhance the data privacy. Geometric queries are the queries that deal with spatial data. The data is denoted as 'points' and queries are portrayed as geometric objects like triangle, spheres and rectangles. The rest of the paper is organized as follows: Section II depicts the prior work; Section III depicts the proposed work; Section IV depicts the experimental analysis and concludes in Section V.

2. RELATED WORK

This section describes the prior work carried out by researchers. Previously, data utilization method is performed over the plaintext search. Due to increase of the cloud users, search operation is given importance. Usually, Boolean search operation [6] was performed over the server to yield better results. This search fails to give better security to the cloud data. The data is being stored to cloud using 'inner-product similarity'. Search over encrypted data is still in its infancy.

Initially, multi-keyword ranked search was introduced by Information Retrieval System (IRS). Latent Semantic Analysis (LSA) was used to retrieve the matched data. Latent values between terms and documents were used for finding the association. Further, k-NN classification technique is used for generating the security index. The secure index was obtained from mini-hash sketches that include cryptography, image processing and information retrieval [7]. The schema contains hash functions and inverted visual words. It yields slow performance in inverted visual words. The theme of cryptographic model is to provide secure systems. The method incurs higher storage overhead and not guarantees the security. The author in [8] studied about the issue of CSP towards search operations. Similarly, authentication model was also suggested the data access. Their technique purely works as Public key encryption. The computational cost of decrypting the data was wisely reduced. The author in [9] suggested an authorized data privacy systems. They defined two algorithms which depicts searching efficiency and privacy of the query. It was executed on personal health records of healthcare applications. It fails to support synonyms or morphological variants are used.

In [10], the author suggested an alternate scheme to handle synonyms variants. Cong method drastically lessened the processing cost and network traffic. It eliminates the search barriers in the information retrieval systems. And they also proposed other methods named, ‘Ranked Searchable Symmetric Encryption’, ‘Order persevering symmetric encryption’ and ‘One many order preserving mapping'. Their method solved irrelevant data and traffic issues. It didn’t cooperate for the query matching. The author in [11] framed search method based on Identity based Encryption. Their method supported a single query as well as multi-queries. Any cloud users can access the data but only authorized users permit to edit the data. The author in [12] suggested predictive based encryption technique that supports hierarchical functioning systems. It is computationally expensive.

Confidentiality is one of the security parameters used over the encrypted data in cloud. It wisely makes use of ranking order system. Depending on encrypted queries it ranks the documents and document having most rank will be pushed up using ranked method. The given method is well suited for large documents and also it provides higher accuracy and security. But for this method computational cost is high and protecting communication link is bit difficult task. In [13], the author studied about the attribute based encryption model. The attribute based model is merged with the predictive encryption scheme. In order to makes the system faster, a highly secured retrieval system was framed. Keyword plays an important role in both attribute and predictive based encryption. By doing so, privacy of the data is assured. Still, the challenges like keywords refreshment, channel elimination and multi-keyword processing. Though it maintains the secret data but fails to support integrity and privacy design. The study was further enhanced and introduced locality based searching technique. It is highly effective than the hashing techniques. Since, it’s a one-way process, the resultset doesn’t meet the user’s requirements. The author in [14] framed a privacy preserving model. The search operation is carried out in two phases, namely, Ranked over keyword search, search over structured data. Though confidentiality parameter is achieved, the search over encrypted data was unsuccessful. The author in [15] studied about the confidentiality, verifiability and security of their proposed algorithm. The parameter verifiability was achieved by performing cross-check condition over encrypted data. They also studied about the fuzzy logic system to provide privacy.

3. PROPOSED WORK

The proposed work is purely based on Symmetric Key Encryption scheme. The system model of our scheme is given in Fig.1. The system model consists of three entities, namely, data owner, data user and cloud server. The task of data owner is to preserve the data at cloud server, eventually focus on reducing the local cost. The outsourced data will be searched by the data user. The task of cloud server is to provide services to the data owner and data users. Since, the cloud server is semi-trusted, the cloud service is reliable. The learning of range queries over the private information is a challenging task. The data owner stores the data in encrypted form, to preserve the spatial dataset.

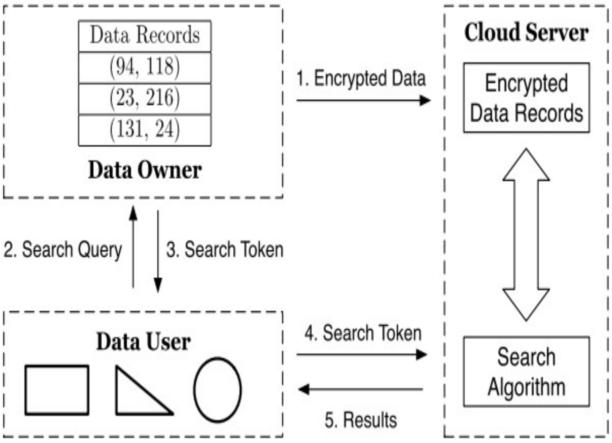


Fig.1 System Architecture

Our proposed algorithm supports different and continuous range queries. The different geometric data is preprocessed and then preceded in the ciphertext data. The proposed algorithm eliminates the multiple rounds of communication between server and client. Firstly, the points are denoted for data records and then range queries are determined from the set of geometric points. The proposed algorithm is explained as follows:

i) Each record is symbolized as geometric points.
ii) Given the input 1^{λ}, the data owner generates the secret keys.

$$SK \leftarrow SSW.\ Setup\ (1^{\lambda}) \quad (1)$$

iii) Along the secret key, bloom filters are generated and outputs as $\{m, h_1 \ldots h_k\}$ where m is the bloom filter length and h is the hash functions. In fact, the bloom filter contains all possible combination of ciphertext, which is further used as search token.

iv) *Encryption phase:* Afforded with secret key SK and dataset D, the data owner encrypts the data as follows:

$$BFD_i: =BF.\ Init\ (m) \quad (2)$$

$$BFD_i :=BF.\ Add\ (D_i, BFD_i) \quad (3)$$

The eqn. (2) and (3) will processed for all data points and the ciphertext C_i will estimated as:

$$C_i \leftarrow SSW.\ Enc\ (SK, U_i) \quad (4)$$

Then, the encrypted dataset is $C= (C_1 \ldots C_n)$

v) *Token Generation phase (S):* The search token is generated from secret key SK and geometric query Q, the data owner calculates

$$S= \{S_1 \ldots S_t\}:= EnumerateInsidePoints\ (Q)$$

$$BF_Q :=BF.\ Init\ (m)$$

BF_Q: = BF. Add (S_i, BF_Q), for $1 \leq i \leq t$, t is the possible points of Q. The search token, TK computes as

$$TK \leftarrow SSW.GenToken\ (SK, \vec{v})$$

vi) Search phase: Afforded with TK and C, the cloud server returns the search results, I_Q

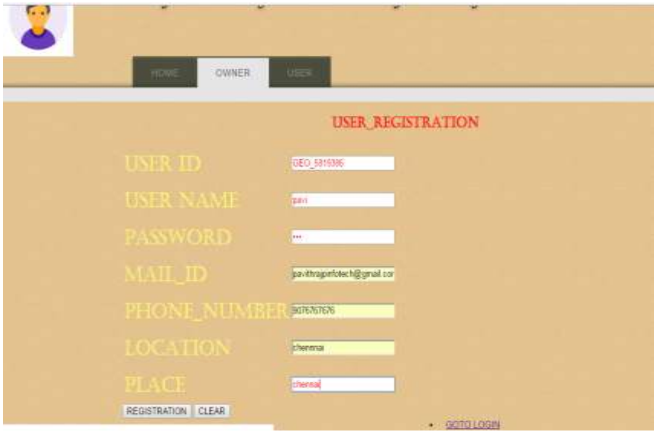
$$Flag_i \leftarrow SSW.\ Query\ (TK, C_i)\ for\ 1 \leq i \leq n.$$

For each flag, the identifier I_i is added to the set I_Q.

The proposed algorithm works in tree structure in order to improve the search complexity. By analyzing the size pattern, search pattern and access pattern, information leakage is reduced in tree structure.

4. EXPERIMENTAL RESULTS

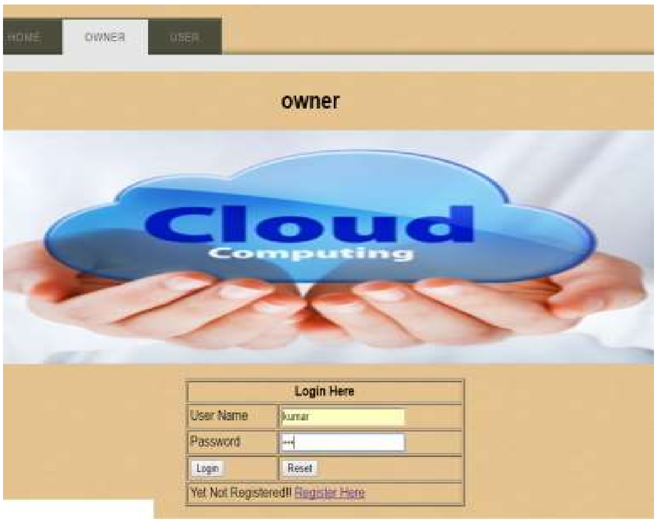
In this section, we analyze the security of our scheme in the semi-trusted cloud server. Our design affects only data manipulations that avoid data leakage. The implementation of proposed algorithm is explained via screenshots.



USER_REGISTRATION

USER ID: GEO_5816386
USER NAME: pavi
PASSWORD: ***
MAIL_ID: pavithrajpinfotech@gmail.cor
PHONE_NUMBER: 9076767676
LOCATION: chennai
PLACE: chensai

Fig.2. User registration form

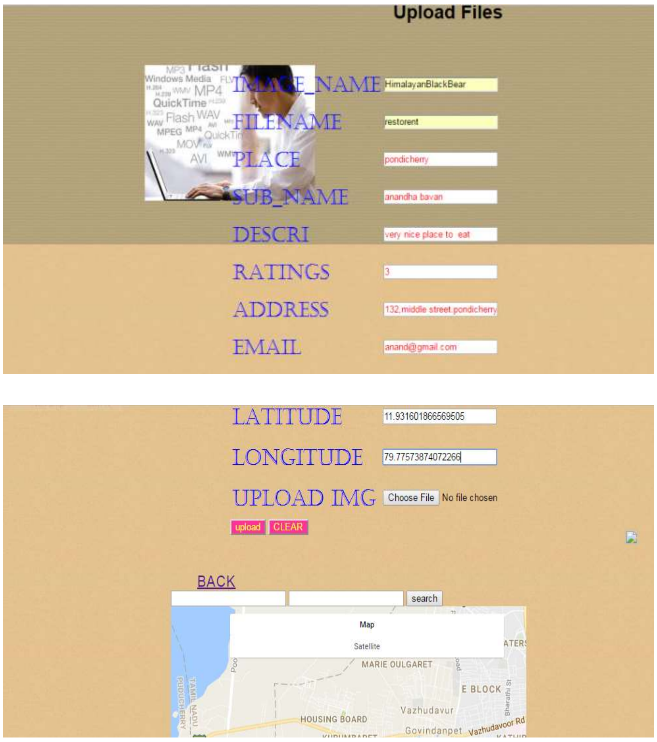


Login Here

User Name: kumar
Password: ***

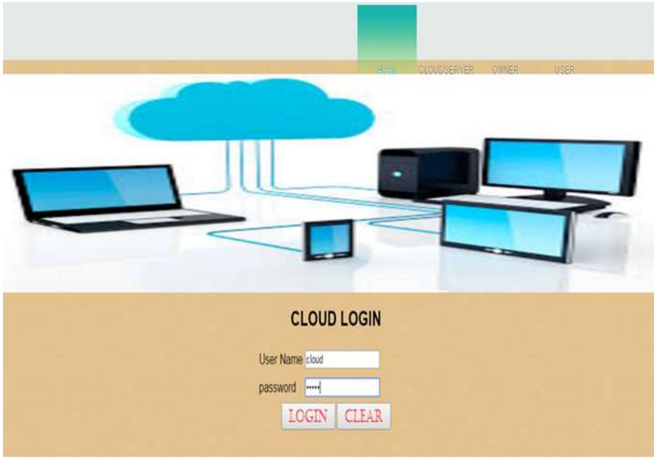
Yet Not Registered!! Register Here

Fig.3. Owner's login



IMAGE_NAME: HimalayanBlackBear
FILENAME: restorent
PLACE: pondicherry
SUB_NAME: anandha bavan
DESCRI: very nice place to eat
RATINGS: 3
ADDRESS: 132,middle street pondicherry
EMAIL: anand@gmail.com
LATITUDE: 11.931601866569505
LONGITUDE: 79.77573874072266
UPLOAD IMG: Choose File No file chosen

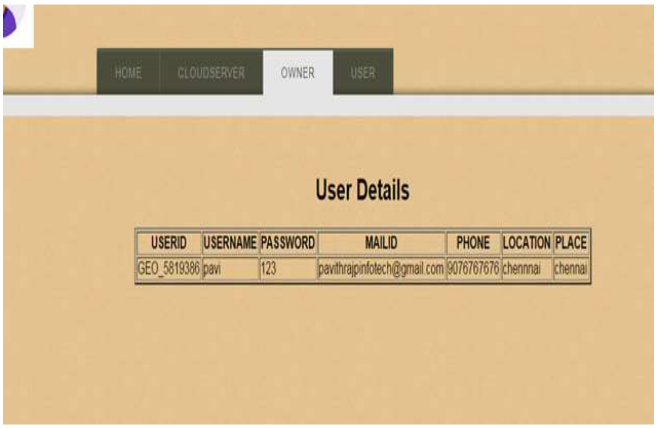
Fig.4. File uploading process



CLOUD LOGIN

User Name: cloud
password: ****

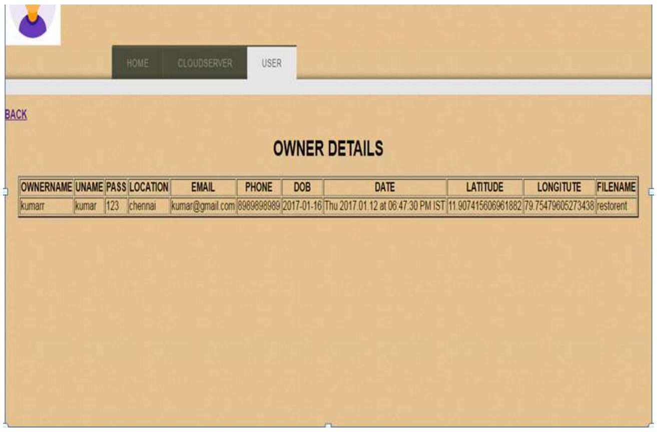
Fig.5. Cloud's login



User Details

USERID	USERNAME	PASSWORD	MAILID	PHONE	LOCATION	PLACE
GEO_5819386	pavi	123	pavithrajpinfotech@gmail.com	9076767676	chennnai	chennai

Fig.6. Viewing user's details

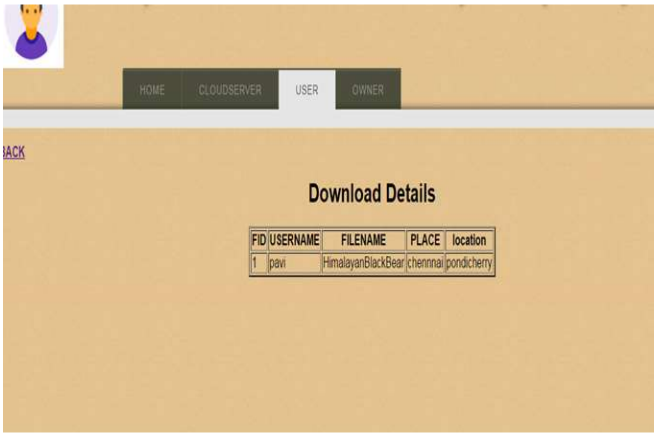


BACK

OWNER DETAILS

OWNERNAME	UNAME	PASS	LOCATION	EMAIL	PHONE	DOB	DATE	LATITUDE	LONGITUTE	FILENAME
kumarr	kumar	123	chennai	kumar@gmail.com	8989898989	2017-01-16	Thu 2017.01.12 at 06:47:30 PM IST	11.907415606961882	79.75479605273438	restorent

Fig.7. Viewing owner's details

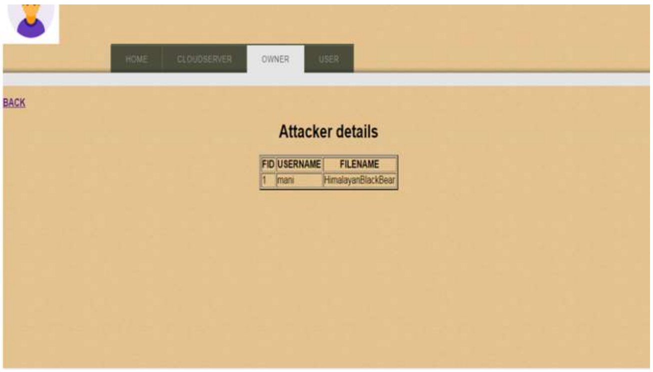


BACK

Download Details

FID	USERNAME	FILENAME	PLACE	location
1	pavi	HimalayanBlackBear	chennnai	pondicherry

Fig.8. Viewing downloads details

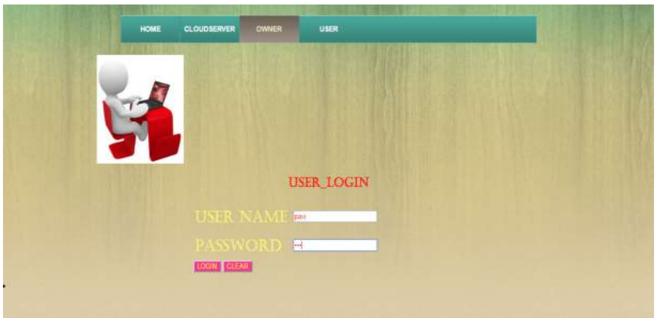


BACK

Attacker details

FID	USERNAME	FILENAME
1	mani	HimalayanBlackBear

Fig.9. Viewing attacker details

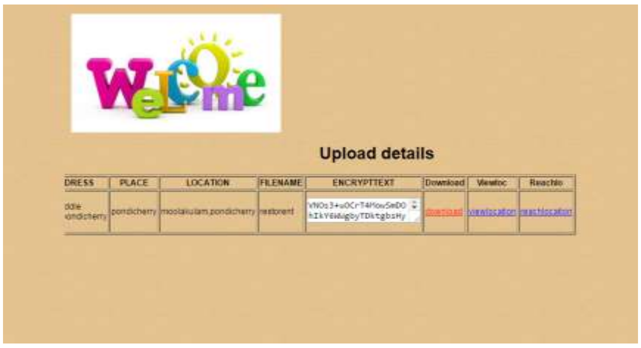


USER_LOGIN

USER NAME pavi

PASSWORD ---

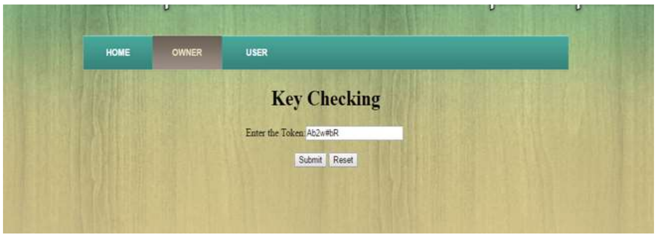
Fig.10. Owner's login



Upload details

DRESS	PLACE	LOCATION	FILENAME	ENCRYPTTEXT	Download	Viewloc	Reachlo
ddle ondicherry	pondicherry	moolakulam,pondicherry	restorent	VNOs3+uOCrT4MowSmDO hIkY6WgbyTDktgbsHy	download	viewlocation	reachlocation

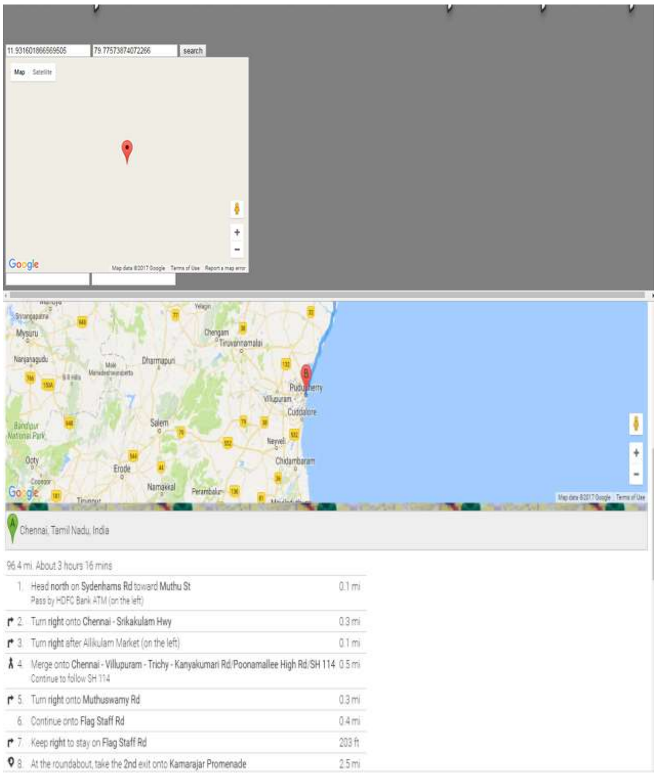
Fig.11. Viewing uploaded file details



Key Checking

Enter the Token: Ab2w#bR

Fig.12. Key checking process



11.931601866569606 79.77573874072266

Chennai, Tamil Nadu, India

96.4 mi. About 3 hours 16 mins

1.	Head north on Sydenhams Rd toward Muthu St Pass by HDFC Bank ATM (on the left)	0.1 mi
2.	Turn right onto Chennai - Srikakulam Hwy	0.3 mi
3.	Turn right after Allikulam Market (on the left)	0.1 mi
4.	Merge onto Chennai - Villupuram - Trichy - Kanyakumari Rd/Poonamallee High Rd/SH 114 Continue to follow SH 114	0.5 mi
5.	Turn right onto Muthuswamy Rd	0.3 mi
6.	Continue onto Flag Staff Rd	0.4 mi
7.	Keep right to stay on Flag Staff Rd	203 ft
8.	At the roundabout, take the 2nd exit onto Kamarajar Promenade	2.5 mi

Fig.13. Location search over the encrypted data.

5. CONCLUSION

In this paper, we examined the reachability issue over encrypted spatial data. We designed a novel and lightweight scheme, Geometric Range Search Model (GRSM) that retrieves the data in reduced search space complexity. Moreover, the anonymity detection using location based services is a troublesome task. Generally, the four spatiotemporal points are enough to identify the location of the individual. In view of third party cloud, there is a chance of revealing the location of the individual to the anonymity. To resolve this issue, the researcher insisted to provide an end-to-end encryption. Each record is considered as the points and the set of points constitutes for range search. Based on the security parameter, the proposed algorithm consists of three important phases, encryption phase, token generation phase and search phase. All the three phases are interlinked with each other to perform search over ciphertext dataset. The proposed algorithm is executed over semi-trusted cloud server. An experimental result shows the efficacy of the proposed GRSM.

REFERENCES

[1]. Li Chen, Xingming Sun, Zhihua Xia ,Qi Liu,” An Efficient and Privacy-Preserving Semantic Multi-Keyword Ranked Search over Encrypted Cloud Data”, International Journal of Security and Its Application, 2014 .

[2]. Wenjun Lu, Ashwin Swaminathan, Avinash L. Varna, and Min Wu, “Enabling search over encrypted multimedia databases,” in proc. of SPIE Media Forensics and Security,09,2009.

[3]. D. X. Song, D. Wagner and A.Perrig, ”Practical Techniques for Searches on Encrypted Data”, Proceedings of the 2000 IEEE Symposium on Security and Privacy.

[4]. Qin Liuy, Guojun Wangyz, and Jie Wuz, ”Secure and privacy preserving keyword searching for cloud storage services”, ELSEVIER Journal of Network and computer Applications, March 2011

[5]. Ming Li et al.,”Authorized Private Keyword Search over Encrypted Data in Cloud Computing, IEEE proc. international conference on distributed computing systems, June 2011.

[6]. Cong Wang et al.,”Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data”, IEEE Transactions on parallel and distributed systems, vol. 23, no. 8, August 2012

[7]. Boneh, D., Di Crescenzo, G., Ostrovsky, R., Persiano, G, “Public key encryption with keyword search”, In: Advances in Cryptology-Eurocrypt 2004.

[8]. Li, M., Yu, S., Cao, N., Lou, W.: Authorized private keyword search over encrypted data in cloud computing. In: Distributed Computing Systems (ICDCS), 2011 31st International Conference on distributing computing systems, pp. 383–392.

[9]. Swaminathan A, Mao Y, Su G-M, Gou H, Varna AL, He S, M. Wu, Oard D, “Confidentiality Preserving Rank-Order search”, IEEE transactions on Storage security and survivability 2007-Conference Papers (Proceedings of the 2007 ACM workshop on Storage security and survivability) pp.7-12.

[10]. Dan Boneh, Giovanni Di Crescenzo, Rafail Ostrovsky, Giuseppe Persiano “Public Key Encryption with Keyword Search” Springer Berlin Heidelberg; EUROCRYP’04, Volume: 3027 of LNCS- 2004 pp.506-522

[11]. Y.-C. Chang, M. Mitzenmacher, “Privacy preserving keyword searches on remote encrypted data” conference 2005.

[12]. Mehmet Kuzu, Mohammad Saiful Islam, Murat Kantarcioglu, “Efficient Similarity Search over Encrypted Data “, IEEE transaction, 2012.

[13]. Ming Li et al.,”Toward Privacy-Assured and Searchable Cloud Data Storage Services”, IEEE Transactions on Network, 27 (4), 2013

[14]. Jianfeng Wang et al., “Efficient Verifiable Fuzzy Keyword Search over Encrypted Data in Cloud Computing” , Journal of Computer Science and Information system, volume 10, Issue 2, April 2013

[15]. Wenhai Sun et al., "Privacy-Preserving Multi-keyword Text Search in the Cloud Supporting Similarity-based Ranking", the 8th ACM Symposium on Information, Computer and Communications Security , Hangzhou, China, May 2013.