Y. Bhargav *et al*, International Journal of Computer Science and Mobile Applications,
Vol.1 Issue. 5, November- 2013, pg. 7-17 **ISSN: 2321-8363**

Homomorphic Recommendations for Data Packing

Y. Bhargav[1], P. Sreenivasa Moorthy[2]

[1]M.Tech. Student, CSE Dept, CMR Institute of Technology, Hyderabad, A.P
Email-id: bhargav.y9959@gmail.com
[2]Associate Professor, CSE Dept., CMR Institute of Technology, Hyderabad, A.P
Email-id:moorthypsm@gmail.com

Abstract

Recommender systems have become an important tool for personalization of online services. Generating recommendations in online services depends on privacy-sensitive data collected from the users. Traditional data protection mechanisms focus on access control and secure transmission, which provide security only against malicious third parties, but not the service provider. This creates a serious privacy risk for the users. In this paper, we aim to protect the private data against the service provider while preserving the functionality of the system. We propose encrypting private data and processing them under encryption to generate recommendations.

Index Terms

Homomorphic encryption; privacy; recommender systems; secure multiparty computation

I. Introduction

Millions of people are using online services for various daily activities [1], many of which require sharing personal information with the service provider. Consider the following online services.

Social Networks: People use social networks to get in touch with other people, and create and share content that includes personal information, images, and videos. The service providers have access to the content provided by their users and have the right to process collected data and distribute them to third parties. A very common service provided in social networks is to generate recommendations for finding new friends, groups, and events using collaborative filtering techniques.[2]

Online Shopping: Online shopping services increase the likelihood of a purchase by providing personalized suggestions to their customers. To find services and products suitable to a particular customer, the service provider processes collected user data like user preferences and clicklogs.

In all of the above services and in many others, recommender systems based on collaborative filtering techniques that collect and process personal user data constitute an essential part of the service. On one hand, people benefit from online services. On the other hand, direct access to private data by the service provider has potential privacy risks for the users since the data can be processed for other purposes, transferred to third parties without user consent, or even stolen In all of the above services and in many others, recommender systems based on collaborative filtering techniques that collect and process personal user data constitute an essential part of the service. On one hand, people benefit from online services. On the other hand, direct access to private data by the service
provider has potential privacy risks for the users since the data can be processed for other purposes, transferred to third parties without user consent, or even stolen In all of the above services and in many others, recommender systems based on collaborative filtering techniques that collect and process personal user data constitute an essential part of the service. On one hand, people benefit from online services. On the other hand, direct access to private data by the service provider has potential privacy risks for the users since the data can be processed for other purposes, transferred to third parties without user consent, or even stolen In all of the above services and in many others, recommender systems based on collaborative filtering techniques that collect and process personal user data constitute an essential part of the service. On one hand, people benefit from online services. On the other hand, direct access to private data by the service provider has potential privacy risks for the users since the data can be processed for other purposes, transferred to third parties without user consent, or even stolen.[3]

II. Related Work

The need for privacy protection for online services, particularly those using collaborative filtering techniques, triggered research efforts in the past years. Among many different approaches, two main directions, which are based on data perturbation [5] and cryptography [6], have been investigated primarily in literature. Polat and Du in [7] and [8] suggest hiding the personal data statistically, which has been proven to be an insecure approach [9]. Shokri et al. present a recommender system that is built on distributed aggregation of user profiles, which suffers from the trade-off between privacy and accuracy [10]. McSherry and Mironov proposed a method using differential privacy, which has a similar trade-off between accuracy and privacy [11]. Cissée and Albayrak present an agent system where trusted software and secure environment are required [12]. Atallah et al. proposed privacy-preserving collaborative forecasting and benchmarking to increase the reliability of local forecasts and data correlations using cryptographic techniques [13]. Canny also presented cryptographic protocols to generate recommendations, which suffer from a heavy computational and communication overhead [14], [15]. Erkin *et al.* Also propose protocols based on

cryptographic techniques, which are computationally more efficient than their counterparts in [16] and [17]. However, in their proposals the users are actively involved in the computations, which makes the overall construction more vulnerable to time-outs and latencies in the users' connections. Moreover, the computations that a single user has to perform involve encryptions and decryptions in the order of thousands, which makes the system expensive to run for the users.

III. Method

This method protects the private data such as users' ratings against the services provider while processing the functionality of the system. It does their by to encrypt private data and processing them under encryption to generate recommendations. The output of the cryptographic protocol as well as the intermediate values in the algorithm is also private and not accessible to the service provider. This method provides a more efficient privacy processing recommender system by improving the techniques further.

By introducing a semi-trusted third party known as privacy service provider (psp), the need for active participation of users in the computation is eliminated. Psp is trusted to perform the assigned tasks correctly, but is prevented to observe private data. The psp has private keys for the PAILLIER and the DGK cryptosystem. The users, who use an applet or a browser plug-in for the service, upload their encrypted data to the service provider. Then the recommendations are generated by running a cryptographic protocol between the service provider and the psp, with no interaction with the users. Encrypted data is processed using homomorphic crypto system for secure multiplication and decryption protocols. This introduces a significant amount of additional computational over head to the system. This computational and communication cost can be significantly reduced by data packing, In which several numerical values are packed in a compact way prior to encryption.

IV. Results

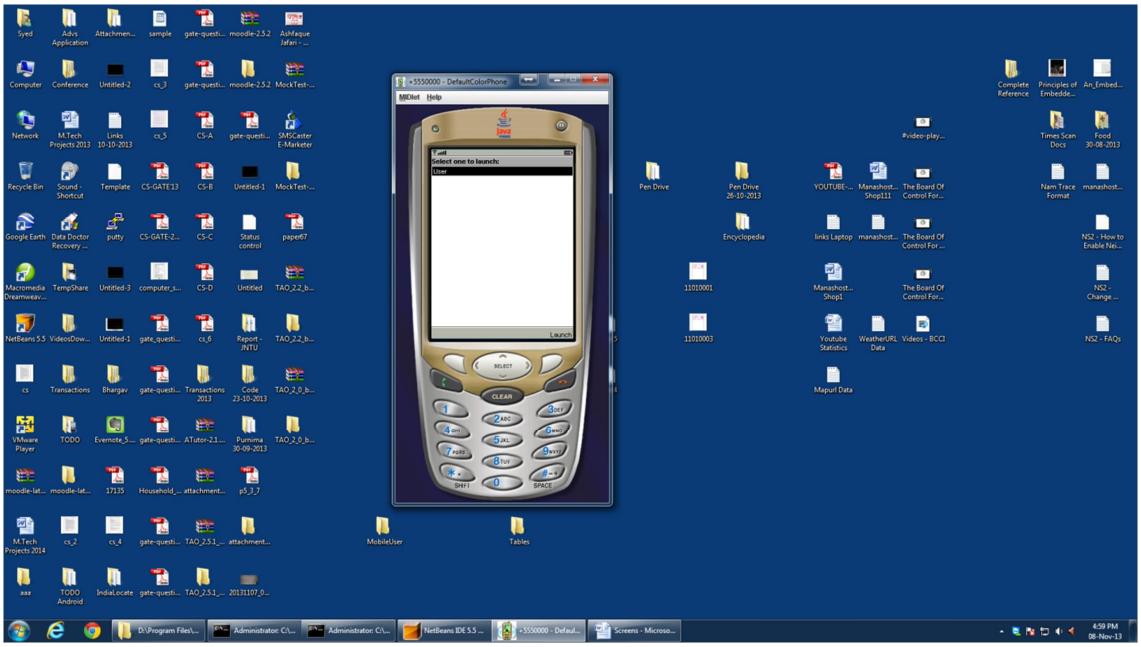

Fig. 1

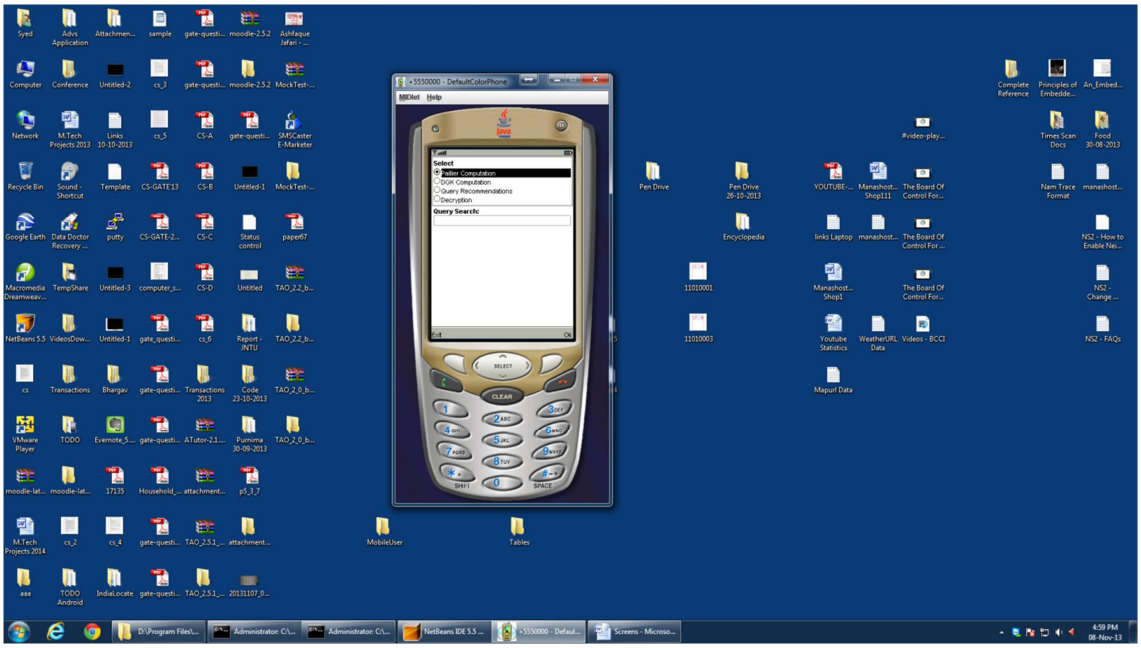

Fig. 2

Fig. 1 & 2: Showing the Online Service Provider

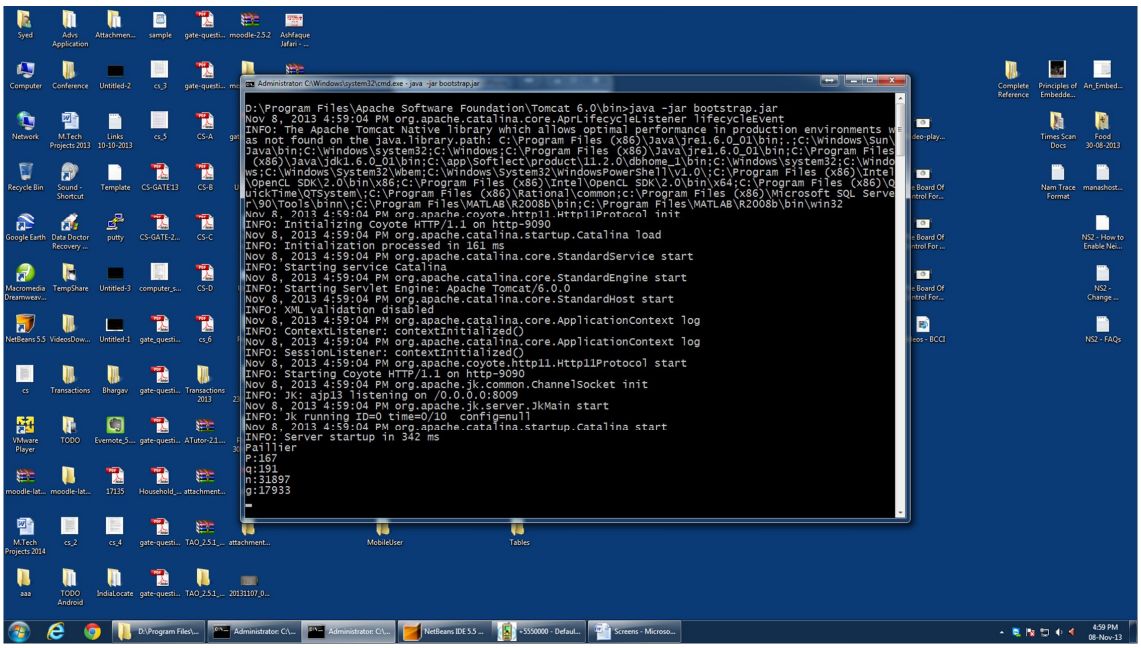

Fig. 3

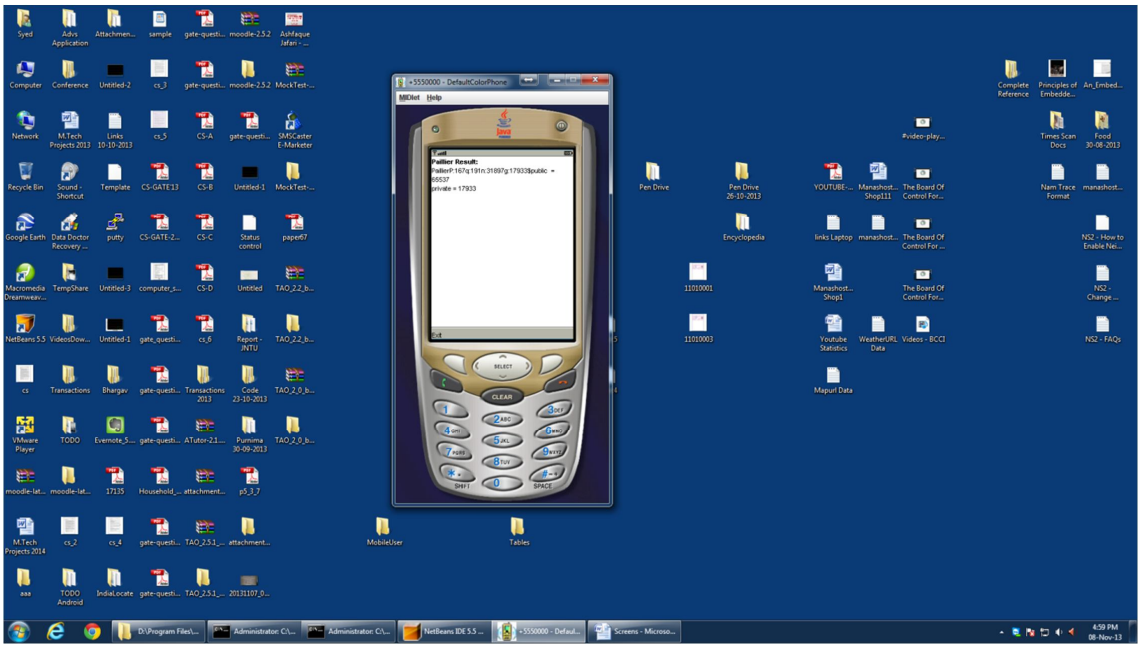

Fig. 4

Fig. 3 & 4: Generating Prime Numbers for Paillier.

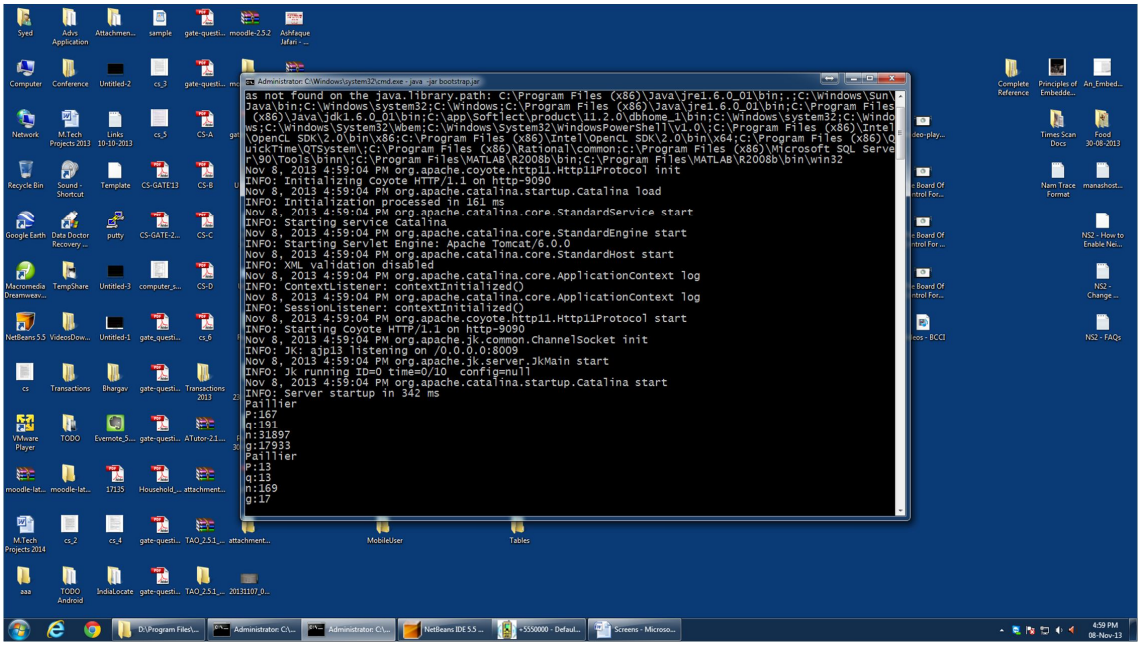

Fig. 5

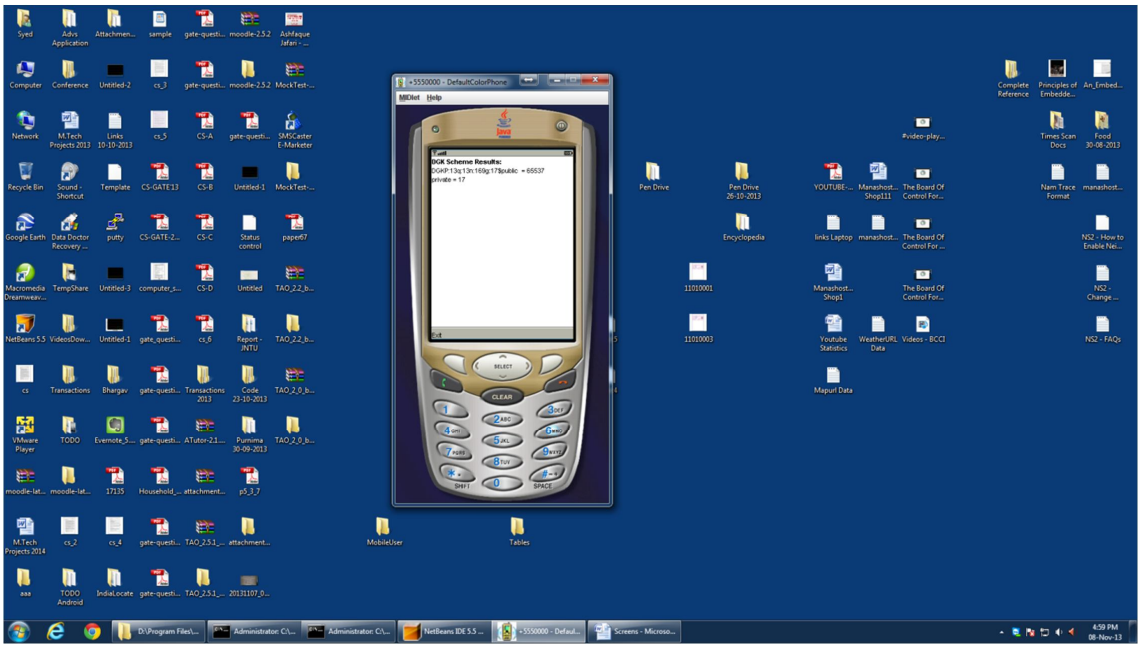

Fig. 6

Fig. 5 & 6: Generating Prime Numbers for DGK.

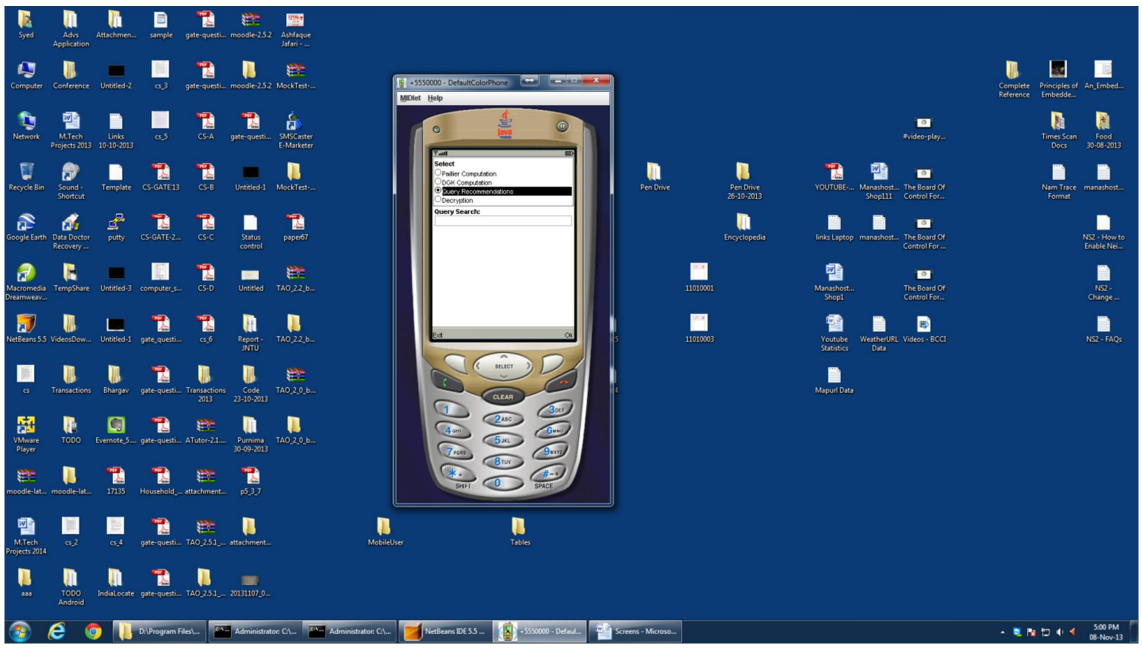

Fig. 7

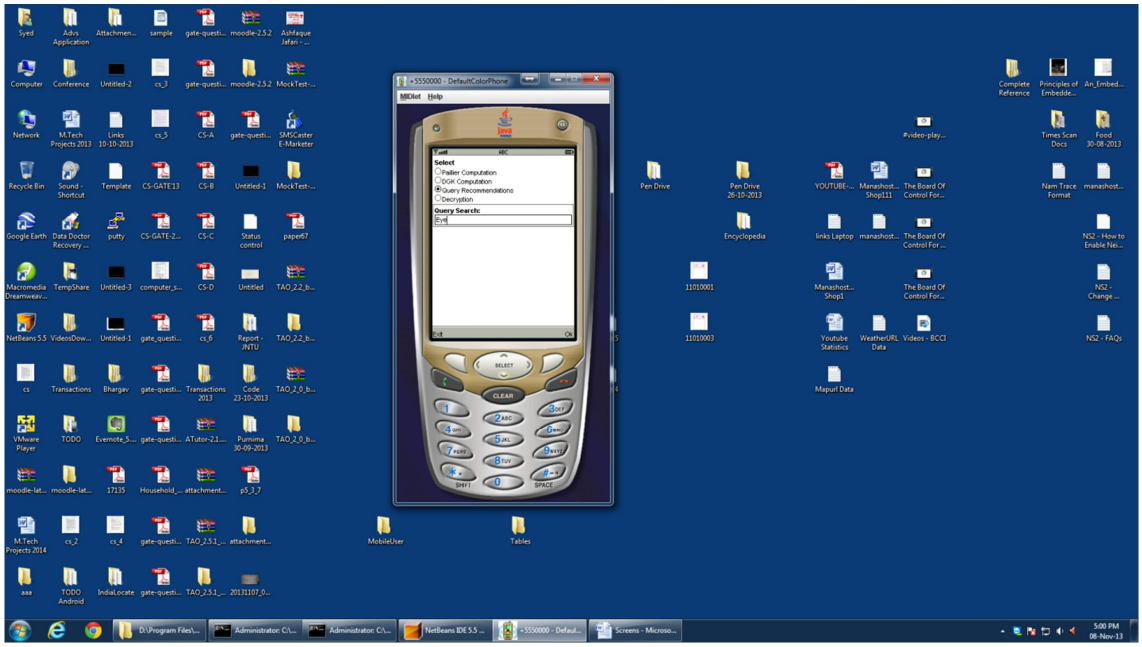

Fig. 8

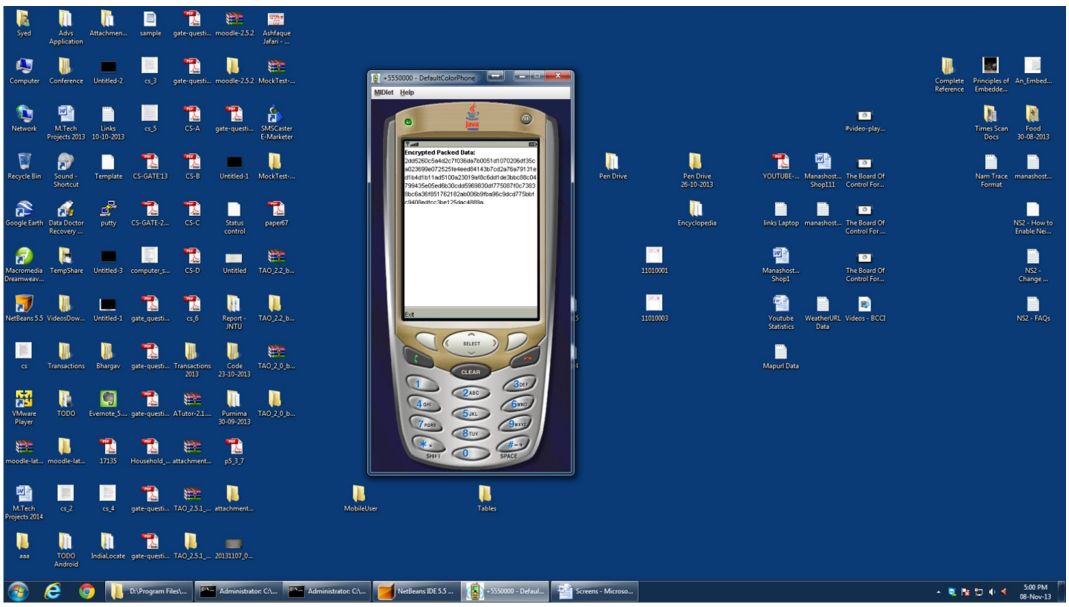

Fig. 9

Fig. 7,8 & 9: Generating Query Recommendations

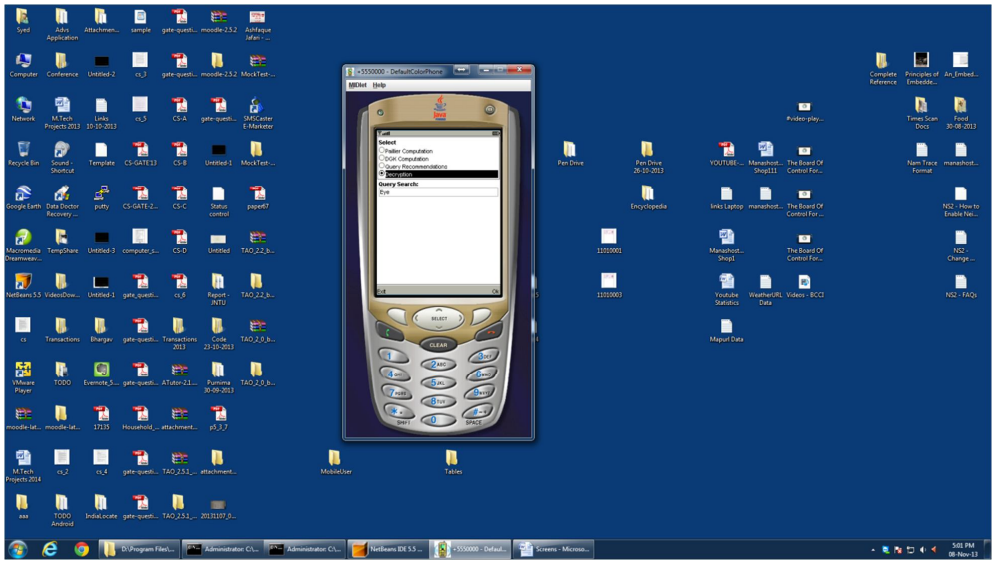

Fig. 10

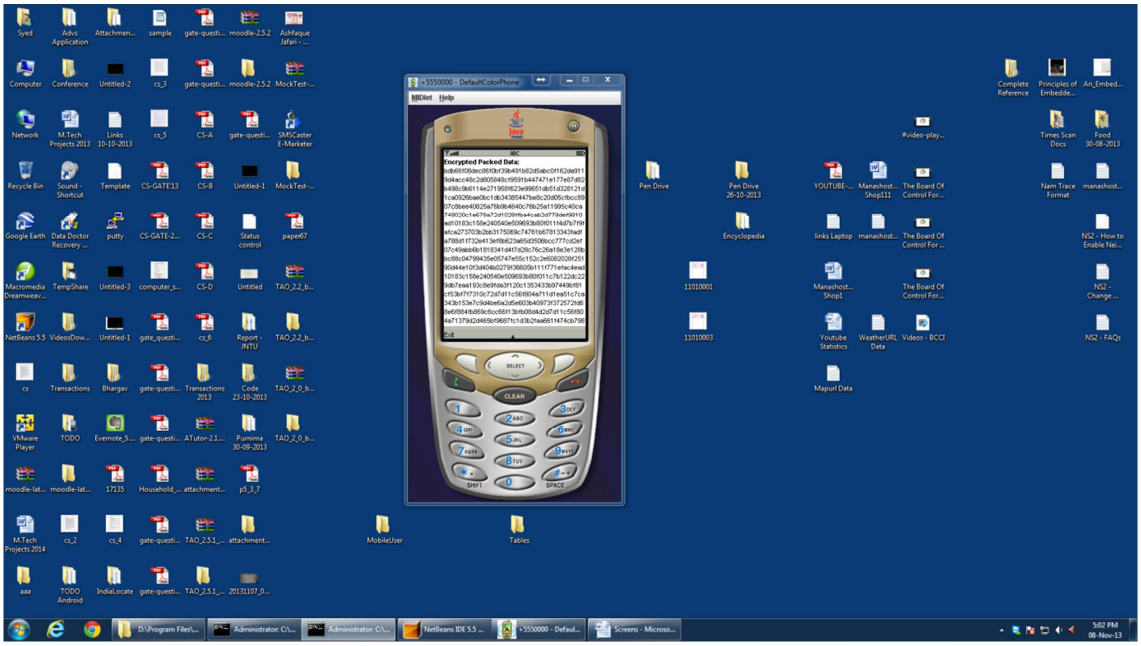

Fig. 11

Fig. 10 & 11: Decrypting the Query Recommendations

V. Conclusions & Future work

In our proposed system, the number of users, who have a similarity value exceeding, is available to user in plain text, which might leak information on the recommender system. In order to hide from the user, it should be kept encrypted. In this case, the recommendations have to be computed by running a secure division protocol, which divides encrypted total by encrypted. Such an approach will introduce additional overhead on the user's side. Notice that even if secure division is chosen, the encrypted has to be compared with, which introduces additional overhead again. Unfortunately, at the end of the comparison, both the service provider and user will know the outcome of since the service provider needs this information to validate the generated recommendations. Since in either scenario a certain amount of information is leaked, we permit users to know the number of similar users for efficiency reasons.

We have presented a highly efficient, privacy-preserving cryptographic protocol for a crucial component of online services: recommender systems. Our construction with a semitrusted third party, the PSP, ensures a protocol where user participation in the heavy cryptographic operations is no longer needed.We also employ data packing to ease the computational and communication burden between the service provider and the PSP. A cryptographic protocol particularly developed for comparing packed and encrypted values, enables us to compare multiple encrypted data elements in a single operation.

References

[1] List of Social NetworkingWebsites 2009 [Online]. Available: http://enwikipedia.org/wiki/List_of_social_networking_websites

[2] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734–749, Jun. 2005.

[3] N. Ramakrishnan, B. J. Keller, B. J. Mirza, A. Y. Grama, and G. Karypis, "Privacy risks in recommender systems," *IEEE Internet Comput.*, vol. 5, no. 6, pp. 54–63, Nov./Dec. 2001.

[4] N. Kroes, Digital agenda, Brussels, May 19, 2011.

[5] R. Agrawal and R. Srikant, "Privacy-preserving data mining," in *Proc. SIGMOD Rec.*, May 2000, vol. 29, pp. 439–450.

[6] Y. Lindell and B. Pinkas, "Privacy preserving data mining," *J. Cryptol.*, pp. 36–54, 2000, Springer-Verlag.

[7] H. Polat and W. Du, "Privacy-preserving collaborative filtering using randomized perturbation techniques.," in *Proc. ICDM*, 2003, pp. 625–628.

[8] H. Polat andW. Du, "SVD-based collaborative filtering with privacy," in *Proc. 2005 ACM Symp. Applied Computing (SAC'05)*, New York, NY, 2005, pp. 791–795, ACM Press.

[9] S. Zhang, J. Ford, and F. Makedon, "Deriving private information from randomly perturbed ratings," in *Proc. Sixth SIAM Int. Conf. Data Mining*, 2006, pp. 59–69.
[10] R. Shokri, P. Pedarsani, G. Theodorakopoulos, and J.-P. Hubaux, "Preserving privacy in collaborative filtering through distributed aggregation of offline profiles," in Proc. Third ACM Conf. Recommender Systems (RecSys'09), New York, NY, 2009, pp. 157–164, ACM.

[11] F.McSherry and I. Mironov, "Differentially private recommender systems:Building privacy into the net," in Proc. 15th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD'09), New York, NY, 2009, pp. 627–636, ACM.

[12] R. Cissée and S. Albayrak, "An agent-based approach for privacypreserving recommender systems," in Proc. 6th Int. Joint Conf. Autonomous Agents and Multiagent Systems (AAMAS'07), New York, NY, 2007, pp. 1–8, ACM.

[13] M.Atallah, M. Bykova, J. Li,K. Frikken, andM. Topkara, "Private collaborative forecasting and benchmarking," in Proc. 2004 ACM Workshop on Privacy in the Electronic Society (WPES'04), New York, NY, 2004, pp. 103–114, ACM.

Authors' Biography

Y.Bhargav had B.Tech from Guru Nanak Engineering College, Ibrahimpatnam, Hyderabad. He is an M.Tech. student in CSE Department of CMR Institute of Technology, Hyderabad. He is currently working for her M.Tech. research project work under the guidance of Mr.P.S.Murthy. His areas of interest include Network Security, Computer Networks, and Programming languages.

P.Sreenivasa Moorthy M.E., (Ph.D) he is currently working as Associate Professor in CSE Department of CMR Institute of Technology, Hyderabad. His areas of interest are Image Processing, DataBases, Networking and Security.

