ISSN: 2321-7782 (Online)
Volume 2, Issue 11, November 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

A Study on Recommendation Generation using Homomorphic Encryption

Shital E. Jarande[1]
PG Student
Department of Computer Engineering
GHRIET, SP University of Pune
Pune, India

Amit Dravid[2]
Assistant Professor
Department of Computer Engineering
GHRIET, SP University of Pune
Pune, India

Abstract: ***The Online recommended system enables personal service to user. We are studying technique that operates on security to sensitive customer data, which can be altered by the service provider. This will protect customer data by encryption of the data and executes recommendations by the encryption. Thus, the service provider observes neither user preferences nor recommendations. The method proposed by author uses homomorphic encryption and secures multi-party computation, which introduce a significant overhead in computational complexity. Author minimizes the introduced overhead by securing data and using cryptographic protocols particularly developed for this purpose.***

Keywords: Recommendation systems, privacy, secure multiparty computation, homomorphic encryption.

I. INTRODUCTION

From many years, we have experienced phenomenal progress in information and communication systems. As a result, online applications have become very popular for millions of people. Personalization is a common approach to attract even more people to web service. The system can suggest personalized services tailored to a particular user based on his preferences. Since the personalization of the services offers high profits to the service providers and poses interesting research challenges, research for generating recommendations, which interests from industry. The techniques for generating recommendations for users strongly rely on the way personal user information is gathered. This information can be provided by the user himself as in profiles, or the service provider can observe users' actions like click logs. On one hand, more user information helps the system to improve the accuracy of the recommendations. On the other hand, the personal information on the users creates a severe privacy risk since there is no solid guarantee for the service provider not to misuse the users' data. It is often seen that whenever a user login in the system, service providers statements the ownership of data provided from user and approves itself to allocate the data to third party. To address the privacy considerations in recommendation systems, in, Canny proposes a system where the private user data is encrypted and recommendations are generated by applying an iterative procedure based on the conjugate gradient. This algorithm calculates characterization of user in subspace. It generates recommendation by computing projections in encrypted domain. This iterative algorithm takes many rounds for convergence and in each round; users need to participate in an expensive decryption procedure which is based on a threshold value, where an important portion of the users is assumed to be online and honest. The output of iteration, which is the characterization matrix, is available in the clear. Canny proposes a method to protect the privacy of users based on a probabilistic factor analysis model by using a similar approach as in. While Canny works with encrypted user data, Polat and Du suggest protecting the privacy of users by using randomization techniques. In their papers, they blind the user ratings with random data assuming that during the aggregation this randomization cancels out and the result is a good estimation of the outcomes. The achievement of the method depends on number of customers participated in the calculation since for the system to work; the users need to be present in vast amounts. This creates a trade-off between accuracy/correctness of the recommendations and the number of users in the system. Moreover, the outcome of the algorithm is also available to the server, which poses a privacy threat to user. The randomization techniques are believed to be

highly insecure. In that paper, the privacy sensitive data of the users of a recommendation system are protected by means of encryption and recommendations are generated by processing the encrypted data. Being highly efficient compared to other techniques in terms of computational and communication costs as well as provably secure, still suffers from a computationally expensive comparison protocol. In this paper, construct an extremely efficient system that does not require the active participation of the user by using a semi trusted party. Private recommendation system uses elgamel algorithm but system is more complex and inefficient. To overcome this drawback proposed system uses ElGamal algorithm. This system is efficient to generate private recommendations in a privacy-preserving manner.

II. Related Work

Paillier's scheme in which the expansion factor is reduced and which allows to adjust the block length of the scheme even after the public key has been fixed, without losing the homomorphic property. It shows that the generalization is as secure as Paillier's original system and proposes several ways to optimize implementations of both the generalized. We build a threshold of the generalized as well as zero-knowledge protocols to show that a given cipher text encrypts one of a set of given plain texts, and algorithms are to verify multiplicative on plaintexts. Then it also shows how these building blocks can be used for applying the scheme to efficient process. This reduces dramatically the work needed to compute the final result of a process, compared to the previously best known schemes. It shows how the basic scheme for process can be easily adapted to casting a vote for up to timeout of L candidates. The same basic building blocks can also be adapted to provide receipt-free elections, under appropriate physical assumptions. The scheme for 1 out of L process can be optimized.

The ElGamal Algorithm provides an alternative to the RSA for public key encryption.

1. Security of the RSA depends on the (presumed) difficulty of factoring large integers.
2. Security of the ElGamal algorithm depends on the (presumed) difficulty of computing discrete logs in a large prime modulus.
3. ElGamal has the disadvantage that the cipher text is twice as long as plaintext. It has benefits that the same text gives different cipher when it gets encrypted.

III. Literature Survey

In recent years, outsourcing large amount of data in system and how to manage the data raises many challenges to privacy. The worries of privacy will be addressed if users encrypted the data in the system. In cryptographic schemes, homomorphic system allows us to performing computations on encrypted data. The research are dealing with homomorphic encryption scheme for maintaining privacy and security in system by detecting the error incurred while transferring data using RSA cryptosystem.[2]

By offering personalized content to users, recommendation systems have become a vital tool in e-commerce and online media applications. Content based algorithm recommended item or product to user that is most similar to those previously purchased. In this paper, author proposes technological mechanisms to protect the privacy of individuals in a recommendation system. Our proposal is based on homomorphic system. This is used for the reserved rating of the customers from the service provider. The user's privacy respected by service providers. This is done by generating recommendations with encrypted customer ratings, the service provider's commercially valuable item-item similarities are protected against curious entities, in turn. Author's proposal explores simple and efficient cryptographic techniques to generate private recommendations using a server-client model, which neither relies on (trusted) third parties, nor requires interaction with peer. The main strength of their contribution lies in providing a highly efficient solution without resorting to unrealistic assumptions [3].

Recommendation systems have become important tools in ecommerce. They combine one user's ratings of product or services with ratings from other users to answer queries such as "Would I like X?" with predictions and suggestions. Users thus receive anonymous recommendations from people with similar tastes. While this process seems innocuous, it aggregates user preferences in ways analogous to statistical database queries, which can be exploited to identify information about a particular user. This is especially true for users with eclectic tastes who rate product across different types or domains in the systems. These straddles highlight the conflict between personalization and privacy in recommendation systems. While straddles enable serendipitous recommendation and information about existence may be used in conjunction with other data sources to uncover identities and reveal personal details. We use a graph-theoretic model to study the benefit from and risk to straddles [4].

Collaborative filtering is one of the most widely adopted and successful recommendation approached. Depends on the consumer and product, COLLABRATIVE FILTERING characterizes consumer and product indirectly by previous operations. The example is to recommend most famous product to consumers. Researchers are advancing COLLABRATIVE FILTERING technologies in such areas as algorithm design, human computer interaction design, consumer incentive analysis, and privacy protection. Despite significant progress in COLLABRATIVE FILTERING research, three main problems limit COLLABRATIVE FILTERING's e-commerce applications. First, most research has focused multi graded rating data that explicitly indicate consumer' preferences. However, the available data about consumer-product interactions in e-commerce applications are typically binary transactional data (for example, whether or not someone purchased an item). You can apply COLLABRATIVE FILTERING algorithms for multi graded rating data to binary data, typically with some modest modifications. But these algorithms can't exploit the special characteristics of binary transactional data to achieve more effective recommendation. The second problem is the lack of understanding of the relative strengths and weaknesses of different types of COLLABRATIVE FILTERING algorithms in e-commerce applications. The need for such comparative studies is evident from the many recent studies that have proposed algorithms but conducted only limited comparisons [5].

IV. METHODS

A Recommendation System now becomes decision maker for the people who lack sufficient personal experience to evaluate the items that are on website. It provides recommendation for specific items such as books, news, tourism services etc. Personalization is common term for improving online services and attracts more users. The previous recommendation system like Amazon, Suggest, e-bay provide suggestion for specific items but drawback that service provider can see the ratings of users.

It may happen that owner can give more ratings for particular item and can increase product popularity. This results in misguiding to user and privacy is violated. In our flexible recommendation system users give ratings for some product in encrypted format, so service provider does not able to decrypt it. In this recommendation system users decision time is saved, because time plays vital role in recommendation system. During less amount of time user should get more rated product.

Notation and Security Model:

In paper[1], author consider only the semi honest model where participants are assumed to be honest-but-curious. Therefore, the privacy-preserving protocol they present in this paper can be proven secure in this model by using existing proofs for the basic building blocks, which it is composed of. In addition, in this model they assume that none of the involved parties, namely the service provider, the PSP, and the users, will collude. In particular, they assume that business driven parties, the service provider, and the PSP, do not collude since such an action will harm their business reputation. They refer readers for a discussion on the practicality of using third parties.

Homomorphic Encryption:

A number of public-key cryptosystems are additively Homomorphic, meaning that there exists an operation on cipher texts such that the result of that operation corresponds to a new cipher text whose decryption yields the sum of the messages. Given messages and, the Homomorphic property of an additively cryptosystem with multiplication being the Homomorphic operation in the encrypted domain. Paillier and DGK scheme are two additively Cryptosystems used.

Paillier cryptosystem is used to encrypt privacy sensitive data of user. DGK cryptosystem is used for sub-protocol. DGK is more efficient in terms of encryption and decryption compared to paillier due to it's smaller message space of a few bits.

Elgamel is semantically secured. This cannot distinguish the encryptions of known messages and messages. This achieved by having many possible ciphers for plain text, and choosing randomly between these. Semantic security is required as the messages to be encrypted in our recommendation system have low entropy and could otherwise be recovered by brute force guessing.

Collaborative Filtering:

To generate recommendations for a particular user in a group of N users and M items, author used a system which is based on collaborative filtering. There are three steps of the system are used:

1. Similarities are computed between that particular user and all other users.
2. The L most similar users are selected by comparing their similarity values with a threshold.
3. Recommendations on all the other items are generated as the average rating of the L most similar users.

System Diagram:

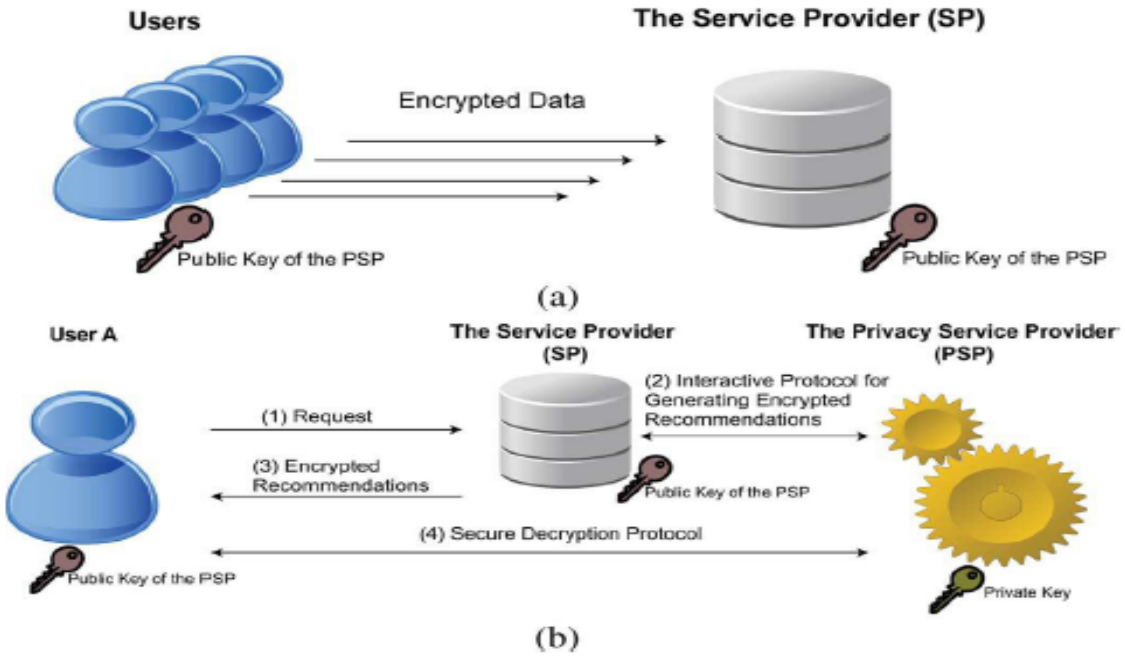

Fig. 1. System model of generating private recommendations.

(a) Encrypted database construction; (b) generating private recommendations.

» **The Service Provider** (SP) has a business interest in generating recommendations for his customers. He has resources for storage and processing.

» **The Privacy Service Provider** (PSP) is a semitrusted third party who has a business interest in providing processing power and privacy functionality. The PSP has private keys for the Paillier and the DGK cryptosystems.

» **Users** are the customers of the service provider. Based on their preferences, in the form of ratings, the service provider generates recommendations for them.

V. CONCLUSION

From the study of svevral papers, conclusions of different authors we came to know that there are different methods and various techniques used for generating recommendations. Cryptographic protocol for generating recommendations based on homomorphic encryption and MPC techniques is used by author. In particular, author proposes to encrypt the privacy sensitive data such as user preferences and similarity values between users and generate recommendation by processing encrypted data. While the homomorphic property allows us to realize linear operations in the encrypted domain, non-linear operations like comparing encrypted values require realizing cryptographic protocols.

References

1. Zekeriya Erkin, Thijs Veugen, Tomas Toft, Reginald L. Lagendijk, "Generating Private Recommendations Efficiently using Homomorphic Encryption and Data Packing", IEEE Transactions on Information Forensics and Security, Vol 7, No. 3, June 2012.
2. Casino, F. Domingo-Ferrer, J. ; Patsakis, C. ; Puig, D. ; Solanas, A. , "Privacy Preserving Collaborative Filtering with k-Anonymity through Microaggregation", e-Business Engineering (ICEBE), 2013 IEEE 10th International Conference on 11-13 Sept.
3. Erkin, M. Beye, T. Veugen, and R. L. Lagendijk, "Privacy enhanced recommendation system," in Proc. ThirtyFirst Symp. Information Theory in the Benelux, Rotterdam, 2010, pp. 35–42.
4. F.McSherry and I. Mironov, "Differentially private recommendation 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.
5. H. Polat and W. Du, "Privacy-preserving collaborative filtering using randomized perturbation techniques.," in Proc. ICDM, 2003, pp. 625–628.
6. Hao Ji, Jinfeng Li, Changrui Ren, Miao He He "Hybrid Collaborative Filtering Model for improved Recommendation" 2013 IEEE.
7. J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, "An algorithmic framework for performing collaborative filtering," in Proc.22nd Ann. Int. ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR'99), New York, NY, 1999, pp. 230–237, ACM.
8. P.Bogetoft, D. L.Christensen, I. Damgård, M. Geisler, T. P. Jakobsen, M. Krøigaard, J. D. Nielsen, J. B. Nielsen, K. Nielsen, J. Pagter, M. I. Schwartzbach, and T. Toft, "Secure multiparty computation goes live," in Proc. Financial Cryptography, 2009, pp. 325–343.
9. P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in Proc. Advances in Cryptology (EUROCRYPT' 99), ser. LNCS, J. Stern, Ed., May 2–6, 1999, vol. 1592, pp. 223–238, Springer.
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. Recommendation Systems (RecSys'09), New York, NY, 2009, pp. 157–164, ACM.
11. R. Cramer, I. Damgård, and J. B. Nielsen, "Multipart y computation from threshold homomorphic encryption," in Proc. Int. Conf. Theory and Application of Cryptographic Techniques (EUROCRYPT'01), London, U.K., 2001, pp. 280–299, Springer-Verlag.
12. Z. Erkin, M. Beye, T. Veugen, and R. L. Lagendijk , " Generating Private Recommendations Efficiently Using Homomorphic Encryption and Data Packing", IEEE Transaction on Information Forensics and Security", Vol. 7,No. 3, JUNE 2012.
13. Z. Erkin, M. Beye, T. Veugen, and R. L. Lagendijk, "Efficiently computing private recommendations," in Proc. Int. Conf. Acoustic, Speech and Signal Processing (ICASSP), Prague, Czech Republic,May 2011, , pp. 5864– 5867, 2011.

