www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 4 Issue 6 June 2015, Page No. 12849-12852

Generating Private Recommendations Using ElGamal Homomorphic Encryption

Patil Maulik Y [1], Manjusha Yeola[2]

[1]Department of Computer Engineering, Alard College of Engineering and Management,
Savitribai Phule Pune University, Pune, India.
Patilmaulik11@gmail.com

[2]Department of Computer Engineering, Alard College of Engineering and Management,
Savitribai Phule Pune University, Pune, India.
Manjusha.amritkar@gmail.com

Abstract: ***In E-commerce, Recommender systems have become an important tool since it is used to personalize the online services and products. Users are concerned about the privacy as personal information can be misused easily. Privacy-sensitive data plays an important role in generating recommendations in online services. Conventional data protection mechanisms provide security only against malicious third parties, as their focus is on access control and secure transmission, but it does not secure by service provider. This creates a major privacy risk for the users. In this paper, we aim to protect the private data from the service provider with functionality of the system is preserved; Private data is encrypted and processed under encryption to generate recommendations. This mechanism becomes highly efficient system by introducing a semi trusted third party, so that system does not require the active participation of user. We have proposed a homomorphic encryption using ElGamal cryptosystem as it is faster, requires considerable less time to decrypt and easier to do distributed key generation. We also aim one comparison protocol, for the comparison of multiple values that are packed in one encryption. According to this work, system can generate private recommendations in privacy preserving manner.***

Keywords: Collaborative filtering, Homomorphic encryption, Privacy, Privacy Service Provider (PSP), Recommender system

Introduction

When many of people are accessing online services for daily activities which involve sharing personal information with the service provider [1]. The example of such online services are social networks, online shopping, IP-TV. In social networks, people acquire in touch with other people, and also create as well as share data which includes personal information, images and videos. The provided contents of the user can be access by service providers and they have the right to buildup the collected data and issue them to third parties. Collaborative filtering technique [2] is to generate recommendations in social networks, a very common service is provided for finding groups, new friends and events. The techniques for generating recommendations for users strongly rely on the information gathered from the user. Collaborative filtering algorithm is collected data from different resources such as users' profiles and its behaviors. Also in online shopping, to find services and products appropriate for a particular user, collected user data similar to user preferences and click logs process by service provider. It increases the possibility of a purchase by providing personalized suggestions to their customers. In web-based activities such as e-commerce, electronic retailers and product providers always offer a large number of products or content items which users are often enforced to choose from. The most important challenge in web based activity is matching consumers with most appropriate products and helps them in decision making process. It is helpful for recommendation system. A modified recommendation for products that suit a user's taste can not only enhance user satisfaction and loyalty, but also increase conversions and profits for electronic retailers. Internet leaders are increasingly adopting product recommendation engine for personalized recommendation, such as Amazon, Google, Netflix, TiVo and Yahoo. Recommender systems are flattering an extensive technology used to promote cross-selling. Collaborative filtering is the standards employed to offer users recommendations. Though most collaborative filtering methods require explicit user feedback, such as ratings, it is an entrenched fact that users rate only a small portion of all available products. Consequently, the rating system often acquires insufficient precise feedback which leading to disappointing recommendations.

Our goal in this paper is to present a privacy-preserving version of a content-based recommender system within a realistic business model, which is practical for real-world use. In our scenario, a target customer provides his/her ratings to the service provider, which possesses an item-item similarity matrix. A recommendation for a target product is then generated as a weighted average of the products that the customer rated in the past. While the ratings of the customer are privacy-sensitive, the item-item similarity matrix of the service provider is commercially valuable, and thus, both should be kept private for their respective owners. Our proposal is to use homomorphic encryption [11] to realize linear operations on the encrypted data. Using homomorphic Encryption provides privacy for the customer as his/her private data become inaccessible to the service provider, which does not have the decryption key. The service provider can still generate recommendations, but does this blindly, by performing homomorphic operations on the encrypted data. Because working in the encrypted domain introduces an overhead due to data expansion and expensive operations on large numbers.

Moreover, recommender systems are usually classified into the following category, based on how recommendations are made [9]:

1. Content-based recommendations: In Content-based recommendations, the user will be recommended items similar to the ones the user preferred in the past.

2. Collaborative recommendations: In this recommendation technique, the user will be recommended items that people with similar tastes and preferences liked in the past.

3. Hybrid approaches: In Hybrid approaches, collaborative and content-based methods are combined.

1. LITERATURE SURVEY

Polat and Du [3] in and [4] suggest hiding the personal data statistically, which has been proven to be an insecure approach. McSherry and Mironov proposed a method using differential privacy, which has a similar trade-off between accuracy and privacy [5]. Cissée and Albayrak present an agent system where trusted software and secure environment are required [6]. Canny also presented cryptographic protocols [7] to generate recommendations, which suffer from a heavy computational and communication overhead. 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. Polat use randomized perturbation (RP) technique which protects users' privacy during producing accurate recommendations. Anonymous techniques allow users to reveal their personal information without disclosing their identities but the major problem is that there is no guarantee on the quality of the dataset. So it propose a new scheme, in which each user first disguises his/her personal data, and then sends to a central place where as the data collector cannot derive the truthful information about a user's private information. [4] Distributed method for users to enhance their profiles and protect from an entrusted server, with minimum loss on the accuracy of the recommender system. It addressed the problem of protecting the users' privacy in the existence of an entrusted central server, where the server has direct access to users' profiles. To avoid privacy risk, it proposed a mechanism where users store an offline profile on their own side which hidden from the server and an online profile on the server from which the server generates there commendations. The online profiles of different users are frequently synchronized with their offline versions in an independent and distributed way [8]. Erkin introduces Homomorphic encryption schemes and secure multiparty computation (MPC) techniques for privacy enhanced recommender system. The complexity analysis, the overhead initiate by working inthe encrypted domain is reduced significantly by packing data and using the DGK cryptosystem. Proposed system cannot compare with previous system because of space problem [2].The distributed generation of an RSA private key required by a Threshold Paillier Cryptosystems much more complex than the simple independent partial private key generation possible with the ElGamal encryption algorithm. The private key is a factorization secret in Paillier encryption whereas the distributed key generation is extremely inefficient as in ElGamal is much more efficient in voting scheme. In Paillier, each multiplication is performed modulo N2where N is the product of two large primes. In comparison with ElGamal, each multiplication is performed modulo p a large prime. If N and p should have same length then multiplication in Paillier is more costly than ElGamal. In private recommendation the privacy sensitive data such as user preferences and similarity values between users were to be encrypted and generate recommendation by processing those data. As the Homomorphic property permit us to realize linear operations in the encrypted data. Efficiency plays an important role in the success of cryptographic protocols. But because of large data system becomes costly. Multiparty computation is used to keep secret everything which is not to be public, all parties can agree on this security policy, but the multiparty computation is time-consuming as well as expensive.

2. System Model

In this section, we introduce the privacy-preserving version of the recommender system In our setting, we have three roles [8]:

- **Service Provider (SP):** is always interested in generating recommendations for his customers. He has resources for storage and processing.
- **Privacy Service Provider (PSP):** is a semi-trusted third party who has a business interest in providing processing power and privacy functionality. The PSP has private keys for cryptosystems.
- **Users:** Users are the customers of the service provider. Based on their preferences, in the form of ratings, the service provider generates recommendations for them.

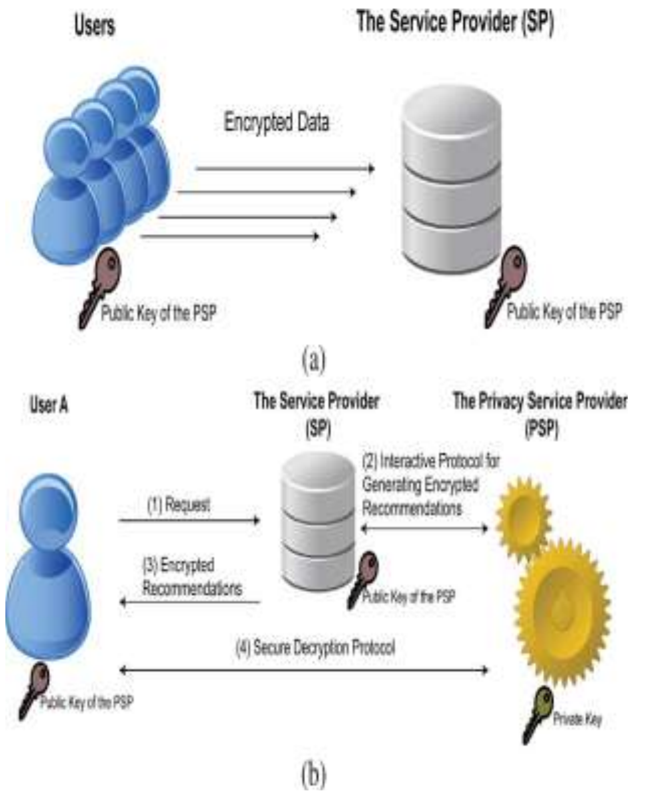
Users
The Service Provider (SP)
Encrypted Data
Public Key of the PSP
Public Key of the PSP
(a)
User A
The Service Provider (SP)
The Privacy Service Provider (PSP)
(1) Request
(2) Interactive Protocol for Generating Encrypted Recommendations
(3) Encrypted Recommendations
Public Key of the PSP
(4) Secure Decryption Protocol
Public Key of the PSP
Private Key
(b)

Figure 1: System model of generating private recommendations a) Encrypted database construction. b) Generating private recommendations

3. Implementation Details

Current systems need active participation of user which becomes privacy risk. To overcome this problem eliminate the need for active participation of users using a semi trusted third party, that is the Privacy Service Provider (PSP), who is trusted to perform the assigned tasks properly, but is not allowed to examine the private data. Encryption and Decryption are doing using additive Homomorphic encryption algorithm such as ElGamal and DGK algorithm. Using this PSP users upload their encrypted data to the service provider and the recommendations are generated by using a collaborative filtering technique between the service provider and the PSP, without interrelate with the users.

4.1 Construction of Database

Before constructing database, system is computing the similarities between particular user and all other user. This similarity stored in vector V. To construct the encrypted database, the users encrypt their data before sending them to the service provider using ElGamal algorithm.

4.2 Generating Recommendations

To generate recommendations, we need two inputs from each user: the densely rated vector to compute the similarity values between users, and the partly rated vector to generate recommendations as the average rating of the top most similar users.

These vectors are highly privacy-sensitive and thus, they will be stored in the encrypted form by the service provider. The service provider does not have the decryption key, thus preventing it from accessing the users' private data.

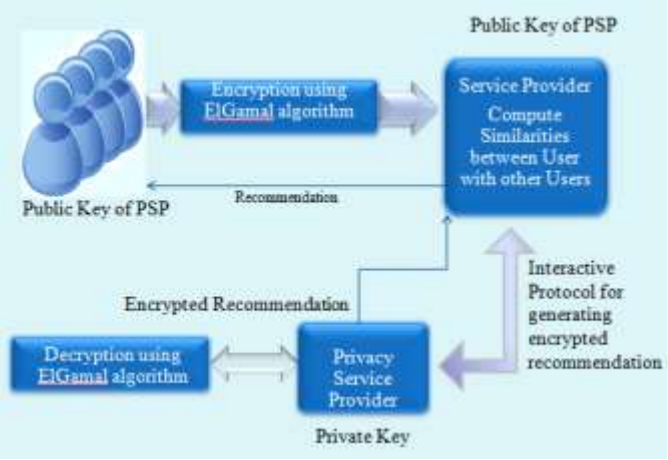
Public Key of PSP
Encryption using ElGamal algorithm
Service Provider
Compute Similarities between User with other Users
Public Key of PSP
Recommendation
Interactive Protocol for generating encrypted recommendation
Encrypted Recommendation
Decryption using ElGamal algorithm
Privacy Service Provider
Private Key

Figure 2: System Architecture

4.3 Algorithm:

```
1. Data from user
2. Encrypt the data using ElGamal algorithm
   a. Choose a large prime p with 150 digits
   b. choose two random integers 1 = q, x < p
   c. Calculate y = $q$x mod $p$
   d. Public key: $p, q, y$; private key: x
   e. Encryption of a data R:  choose a random t and compute
      $a$ = $q$t mod $p$, $b$ = $yR$ mod $p$
   f. Cipher Text c= (c1,c2t)
3. Send cipher text to service provider
4. Calculate Similarities between particular users with all
   other user
5. Send similarities to privacy service provider
6. Decrypt similarities
7. Compute recommendation
   a. Finding similar users
   b. Computing the number L and sum of ratings of most
      similar users
   c. Computing Recommendation
8. Send recommendation to user.
```

4. Result

5.1 Result: Graph shown below gives the comparison between the proposed system and existing system.

5.2 Data set:

U= (U1, U2 …) set of users, I= (I1, I2……..)Set of items, R= (R1, R2…..) Set of densely rated items

This figure gives the average runtime of Homomorphic encryption using Paillier algorithm and Homomorphic encryption using ElGamal algorithm to generate recommendations.

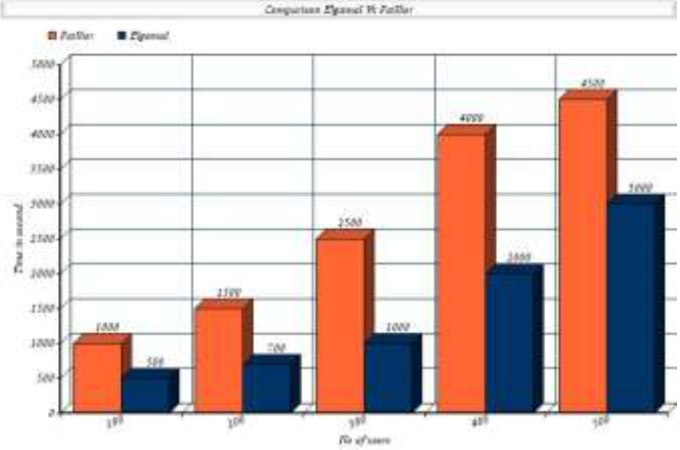

Figure 3: Comparison between HEUP and HEUE

5. Conclusion

In this research, 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. In this construction we use ElGamal cryptosystem rather than Pallier as Elgamal is faster, require minimum decryption time and easier to do distributed key generation and not patented.

We also assume our recommender system is static, meaning that the two sets of items we consider in this work are fixed. Dynamic behaviours such as updating the set of items that are used for similarity computations and removing items from the dataset will cause the service provider and the PSP to run the privacy preserving protocol from the beginning with the new data. We while the experimental results show a significant improvement of the state-of-the-art due to the new setting with a semitrusted third party, and fine-tuned cryptographic protocols, the scalability of the proposed system is not the challenge.

References

[1] G. Eason, List of Social NetworkingWebsites 2009 [Online].Available: http://en.wikipedia.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] H. Polat and W. Du, "Privacy-preserving collaborative filtering using randomized perturbation techniques" in *Proc. ICDM*, 2003, pp.625–628.

[4] H. Polat and W. Du, "SVD-based collaborative filtering with privacy,"in *Proc. 2005 ACM Symp. Applied Computing (SAC'05)*, New York, NY, 2009, pp. 627–636, ACM, 2005, pp. 791–795, ACM Press.

[5] NYF.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, 2005, pp. 791–795, ACM Press.

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

[7] J. F. Canny, "Collaborative filtering with privacy.," in *IEEE* Symp. Security and Privacy, 2002, pp. 45–57.

[8] Zekeriya Erkin, Thijs Veugen, Tomas Toft, and 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. (book style)

[9] Ron Rothblum, "Homomorphic Encryption: from Private-Key to Public-Key", Electronic Colloquium on Computational Complexity, Report No. 146, Sept21,2010

Author Profile

<Author Photo>

Patil Maulik Y. received his Bachelors degree B.E. in Computer Engineering from MIT, Pune in 2012, now pursuing Master of Engineering (ME) in Computer Engineering from Savitribai Phule Pune University, Pune.

<Author Photo>

Manjusha Yeola is an Assistant Professor in Department of Computer Engineering, Alard College of Engineering and management, Pune, Savitribai Phule Pune University, Pune. She has over 10 years teaching experience.