PROTECTING USER DATA IN PROFILE MATCHING SOCIAL NETWORKS

[1]G.Neha,[2]Yasmeen Sulthana
[1]Mtech, Department of Computer Science, VAAGESWARI COLLEGE OF ENGINEERING Thimmapur, Karimnagar, Telangana, INDIA, Ht.no: 20S41D5810, nehagandra78@gmail.com
[2] yasmeen.cse560@gmail.com, Department of Computer Science, VAAGESWARI COLLEGE OF ENGINEERING Thimmapur, Karimnagar, Telangana, INDIA

ABSTRACT: In this paper, we consider a scenario where a user queries a user profile database, maintained by a social networking service provider, to identify users whose profiles match the profile specified by the querying user. A typical example of this application is online dating. Most recently, an online dating website, Ashley Madison, was hacked, which results in disclosure of a large number of dating user profiles. This data breach has urged researchers to explore practical privacy protection for user profiles in a social network. In this paper, we propose a privacy-preserving solution for profile matching in social networks by using multiple servers. Our solution is built on homomorphic encryption and allows a user to find out matching users with the help of multiple servers without revealing to anyone the query and the queried user profiles in clear. Our solution achieves user profile privacy and user query privacy as long as at least one of the multiple servers is honest. Our experiments demonstrate that our solution is practical.

Keywords *– User profile matching, data privacy protection, ElGamal encryption, Paillier encryption, homomorphic encryption.*

1. INTRODUCTION

Matching two or more users with related interests is an important and general problem, applicable to a wide range of scenarios including job hunting, friend finding, and dating services. Existing on-line matching services require participants to trust a third party server with their preferences. The matching server has thus full knowledge of the users' preferences, which raises privacy issues, as the server may leak (either intentionally, or accidentally) users' profiles. When signing up for an online matching service, a user creates a "profile" that others can browse. The user may be asked to reveal details, such as age, sex, education, profession, number of children, religion, geographic location, sexual proclivities, drinking behavior, hobbies, income, religion, ethnicity, drug use, home and work addresses, favorite places. Even after an account is canceled, most online matching sites may retain such information. Users' personal information may be re-disclosed not only to prospective matches, but also to advertisers and, ultimately, to data aggregators who use the data for purposes unrelated to online matching and without customer consent. In addition, there are risks such as scammers, sexual predators, and reputational damage that come along with using online matching services. Many online matching sites take shortcuts with respect to safeguarding the privacy and security of their customers. Often they use counterintuitive "privacy" settings, and their data management systems have serious security flaws. In July 2015, "The Impact Team" group stole user data from Ashley Madison, a commercial website billed as enabling extramarital affairs.

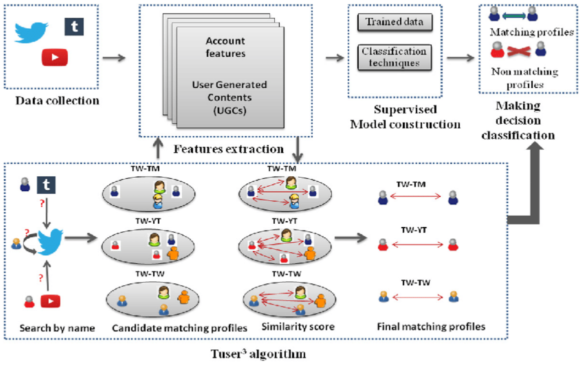


Fig.1: Example figure

The group then threatened to release users' names and personal identification information if Ashley Madison was not immediately shut down. On 18 and 20 August 2015, the group leaked more than 25 gigabytes of company data, including user details. Because of the site's policy of not deleting users' personal information, including real names, home addresses, search history and credit card transaction records, many users feared being publicly shamed. On 24 August 2015, Toronto police announced that two unconfirmed suicides had been linked to that data breach. Such a data breach has raised growing concerns amongst users on the dangers of giving out too much personal information. Users of these services also need to be aware of data theft. A main challenge is thus how to protect privacy of user profiles in social networks. So far, the best solution is through encryption, i.e., users encrypt their profiles before uploading them onto social networks. However, when user profiles are encrypted, it is challenging to perform matching. In this paper, we consider a scenario where a user queries a user profile database, maintained by a social networking service provider, to find out other users whose profiles are similar to the profile specified by the querying user. A typical example of this application is online dating. We give a privacy-preserving solution for user profile matching in social networks by using multiple servers. Our basic idea can be summarized as follows. Before uploading his/her profile to a social network, each user encrypts the profile by a homomorphic encryption scheme with a common encryption key. Therefore, even if the user profile database falls into the hand of a hacker, the hacker can only get the encrypted data. When a user wishes to find people in the social network, the user encrypts his/her preferred user profile and a dissimilarity threshold and submits the query to the social networking service provider. Based on the query, multiple servers, which secretly share the decryption key, compare the preferred user profile with each record in the database. If the dissimilarity is less than the threshold, the matching user' contact information is returned to the querying user.

2. LITERATURE REVIEW

Veneta: Serverless friend-of-friend detection in mobile social networking

Recently, mobile social software has become an active area of research and development. A multitude of systems have been proposed over the past years that try to follow the success of their Internet bound equivalents. Many mobile solutions try to augment the functionality of existing platforms with location awareness. The price for mobility, however, is typically either the lack of the popular friendship exploration features or the costs involved to access a central server required for this functionality. In this paper, we try to address this issue by introducing a decentralized method that is able to explore the social neighborhood of a user by detecting friends of friends. Rather than only exploiting information about the users of the system, the method relies on real friends, and adequately addresses the arising privacy issues. Moreover, we present VENETA, a mobile social networking platform which, among other features, implements our novel friend of friend detection algorithm.

Practical private set intersection protocols with linear complexity

Increasing dependence on anytime-anywhere availability of data and the commensurately increasing fear of losing privacy motivate the need for privacy-preserving techniques. One interesting and common problem occurs when two parties need to privately compute an intersection of their respective sets of data. In doing so, one or both parties must obtain the intersection (if one exists), while neither should learn anything about other set. Although prior work has yielded a number of effective and elegant Private Set Intersection (PSI) techniques, the quest for efficiency is still underway. This paper explores some PSI variations and constructs several secure protocols that are appreciably more efficient than the state-of-the-art.

Efficient robust private set intersection

Computing Set Intersection privately and efficiently between two mutually mistrusting parties is an important basic procedure in the area of private data mining. Assuring robustness, namely, coping with potentially arbitrarily misbehaving (i.e., malicious) parties, while retaining protocol efficiency (rather than employing costly generic techniques) is an open problem. In this work the first solution to this problem is presented.

A public key cryptosystem and a signature scheme based on discrete logarithms

A new signature scheme is proposed, together with an implementation of the Diffie-Hellman key distribution scheme that achieves a public key cryptosystem. The security of both systems relies on the difficulty of computing discrete logarithms over finite fields.

Efficient private matching and set intersection

We consider the problem of computing the intersection of private datasets of two parties, where the datasets contain lists of elements taken from a large domain. This problem has many applications for online collaboration. We present protocols, based on the use of homomorphic encryption and balanced hashing, for both semi-honest and malicious environments. For lists of length k, we obtain O(k) communication overhead and O(k ln ln k) computation. The protocol for the semi-honest environment is secure in the standard model, while the protocol for the malicious environment is secure in the random oracle model. We also consider the problem of approximating the size of the intersection, show a linear lower-bound for the communication overhead of solving this problem, and provide a suitable secure protocol. Lastly, we investigate other variants of the matching problem, including extending the protocol to the multi-party setting as well as considering the problem of approximate matching.

Fully homomorphic encryption using ideal lattices:

We propose a fully homomorphic encryption scheme – i.e., a scheme that allows one to evaluate circuits over encrypted data without being able to decrypt. Our solution comes in three steps. First, we provide a general result – that, to construct an encryption scheme that permits evaluation of arbitrary circuits, it suffices to construct an encryption scheme that can evaluate (slightly augmented versions of) its own decryption circuit; we call a scheme that can evaluate its (augmented) decryption circuit bootstrappable. Next, we describe a public key encryption scheme using ideal lattices that is almost bootstrappable. Lattice-based cryptosystems typically have decryption algorithms with low circuit complexity, often dominated by an inner product computation that is in NC1. Also, ideal lattices provide both additive and multiplicative homomorphisms (modulo a public-key ideal in a polynomial ring that is represented as a lattice), as needed to evaluate general circuits. Unfortunately, our initial scheme is not quite bootstrappable – i.e., the depth that the scheme can correctly evaluate can be logarithmic in the lattice dimension, just like the depth of the decryption circuit, but the latter is greater than the former. In the final step, we show how to modify the scheme to reduce the depth of the decryption circuit, and thereby obtain a bootstrappable encryption scheme, without reducing the depth that the scheme can evaluate. Abstractly, we accomplish this by enabling the encrypter to start the decryption process, leaving less work for the decrypter, much like the server leaves less work for the decrypter in a server-aided cryptosystem.

3. METHODOLOGY

When signing up for an online matching service, a user creates a “profile” that others can browse. The user may be asked to reveal details, such as age, sex, education, profession, number of children, religion, geographic location, sexual proclivities, drinking behavior, hobbies, income, religion, ethnicity, drug use, home and work addresses, favorite places. Even after an account is canceled, most online matching sites may retain such information. Users’ personal information may be re-disclosed not only to prospective matches, but also to advertisers and, ultimately, to data aggregators who use the data for purposes unrelated to online matching and without customer consent. In addition, there are risks such as scammers, sexual predators, and reputational damage that come along with using online matching services.Many online matching sites take shortcuts with respect to safeguarding the privacy and security of their customers. Often, they use counterintuitive “privacy” settings, and their data management systems have serious security flaws.

In this paper, we consider a scenario where a user queries a user profile database, maintained by a social networking service provider, to find out other users whose profiles are similar to the profile specified by the querying user. A typical example of this application is online dating. We give a privacy-preserving solution for user profile matching in social networks by using multiple servers. Our basic idea can be summarized as follows. Before uploading his/her profile to a social network, each user encrypts the profile by a homomorphic encryption scheme with a common encryption key. Therefore, even if the user profile database falls into the hand of a hacker, the hacker can only get the encrypted data. When a user wishes to find people in the social network, the user encrypts his/her preferred user profile and a dissimilarity threshold and submits the query to the social networking service provider. Based on the query, multiple servers, which secretly share the decryption key, compare the preferred user profile with each record in the database. If the dissimilarity is less than the threshold, the matching user’ contact information is returned to the querying user.

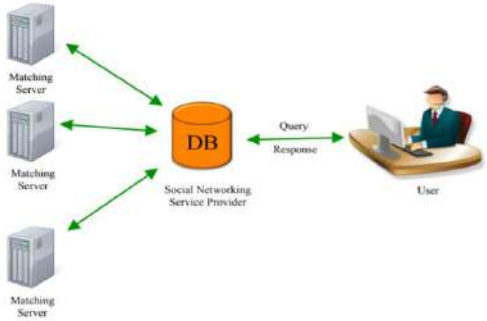
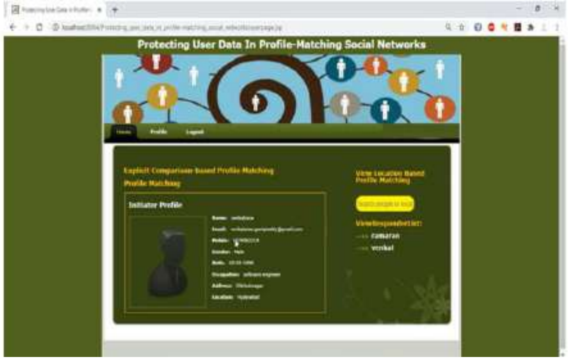


Fig.2: System architecture

4. IMPLEMENTATION

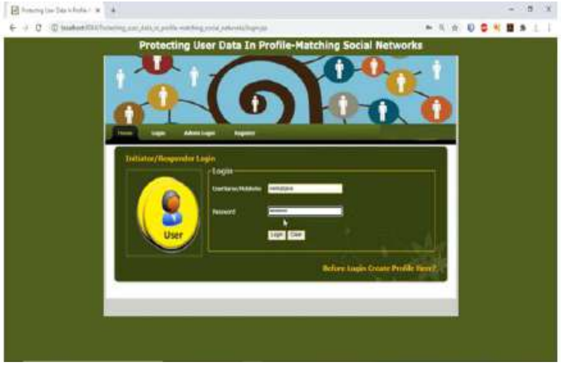
1) We formally define the user profile matching model, the user profile privacy and the user query privacy.
2) We give a solution for privacy-preserving user profile matching for a single dissimilarity threshold and then extend it for multiple dissimilarity thresholds.
3) We perform security analysis on our protocols. If at least one of multiple servers is honest, our protocols achieve user profile privacy and user query privacy.
4) We conduct extensive experiments on a real dataset to evaluate the performance of our proposed protocols under different parameter settings. Experiments show that our solutions are practical and efficient.

5. EXPERIMENTAL RESULTS



Protecting User Data In Profile-Matching Social Networks

Fig.3: Home screen



Initiator/Responder Login

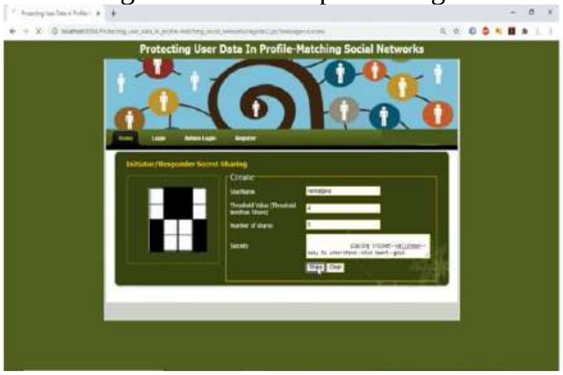
Login

UserName/Mobileno [illegible]

Password ••••••••

Before Login Create Profile Here?

Fig.4: Initiator/Responder login



Initiator/Responder Secret Sharing

Create

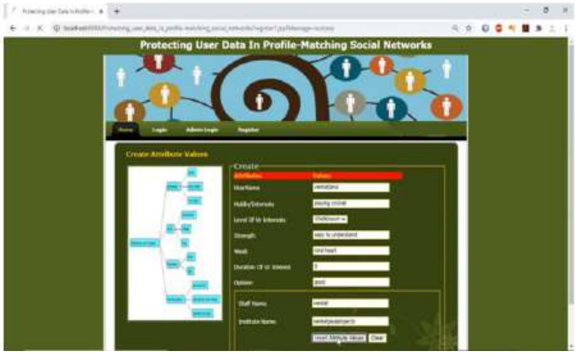
UserName [illegible]

Threshold Value (Threshold lessthan Share) 4

Number of shares 5

Secrets playing cricket--[illegible]-- easy to understand--kind heart--good

Fig.5: Secrete sharing

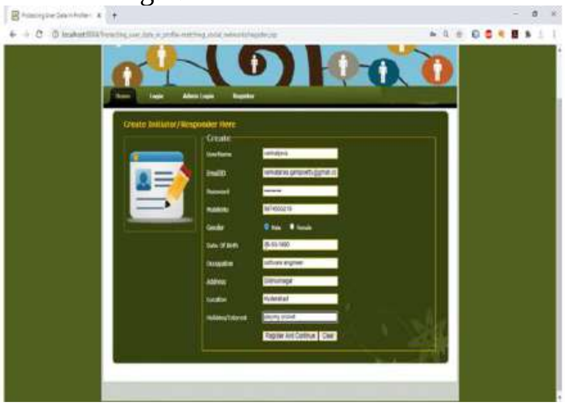


Create Attribute Values

Create

Attributes	Values
UserName	[illegible]
Hobby/Interests	playing cricket
Level Of Ur Interests	[illegible]
Strength	easy to understand
Weak	kind heart
Duration Of Ur Interest	5
Opinion	good
Staff Name	[illegible]
Institute Name	[illegible]

Fig.6: Create attribute values



Create Initiator/Responder Here

Create

UserName [illegible]

EmailID [illegible]@gmail.c

Password ••••••••

MobileNo [illegible]

Gender ● Male ○ Female

Date Of Birth [illegible]

Occupation software engineer

Address [illegible]

Location Hyderabad

Hobbies/Interest playing cricket

Fig.7: Create responder/initiator here

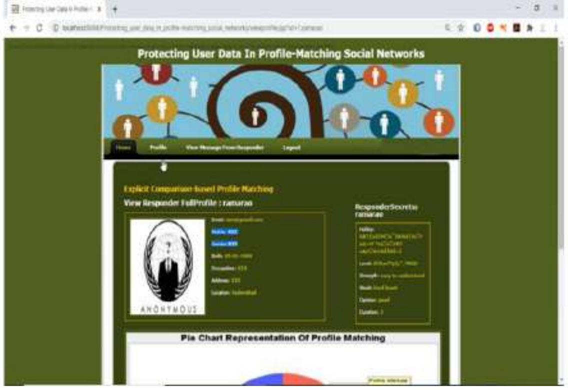

Protecting User Data In Profile-Matching Social Networks
Home
Profile
View Message From Responder
Logout
Explicit Comparison-based Profile Matching
View Responder FullProfile : ramarao
ResponderSecrets: ramarao
ANONYMOUS
Pie Chart Representation Of Profile Matching

Fig.8: Profile matching

6. CONCLUSION

In this paper, we proposed a new solution for privacy-preserving user profile matching with homomorphic encryption technique and multiple servers. Our solution allows a user to find out the matching users with the help of multiple servers without revealing the query and the user profiles. Security analyses have shown that the new protocol achieves user profile privacy and user query privacy. The experimental results have showed that the new protocol is practical and feasible.

Our future work is to improve the performance of computing conditional gates by parallel computation.

REFERENCES

[1] R. Agrawal, A. Evfimievski, and R. Srikant, Information sharing across private databases, in SIGMOD 2003, pp. 86-97.

[2] M. von Arb, M. Bader, M. Kuhn, and R. Wattenhofer, Veneta: Serverless friend-of-friend detection in mobile social networking, in IEEE WIMOB 2008, pp. 184-189.

[3] B. H. Bloom, Space/time trade-offs in hash coding with allowable errors, Communications of the ACM 13 (7): 422-426, 1970.

[4] D. Boneh, E. J. Goh, K. Nissim, Evaluating 2-DNF formulas on ciphertexts, in TCC 2006, pp 325-341.

[5] D. Chaum, Blind signatures for untraceable payments, in Crypto 1982, pp. 199-203.

[6] E. D. Cristofaro and G. Tsudik, Practical private set intersection protocols with linear complexity, in Financial Cryptography and Data Security 2010.

[7] D. Dachman-Soled, T. Malkin, M. Raykova, and M. Yung, Efficient robust private set intersection, in ACNS 2009, pp. 125-142.

[8] T. ElGamal, A public-key cryptosystem and a signature scheme based on discrete logarithms, IEEE Transactions on Information Theory 31 (4): 469-472, 1985.

[9] M. Freedman, K. Nissim, and B. Pinkas, Efficient private matching and set intersection, in EUROCRYPT 2004, pp. 1-19.

[10] C. Gentry, Fully homomorphic encryption using ideal lattices, in STOC 2009, pp 169-178.