CONFIDENTIALITY PROTECTIVE USER PROFILE MATCHING IN PUBLIC NETWORKS

Maddala Mounika[1], K. Tulasi Krishna Kumar Nainar[2]

[1]Student-CSE Department, Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andhra Pradesh.

[2]Associate Professor -CSE Department, Sanketika Vidya Parishad Engineering College, Visakhapatnam, AP

Article DOI: https://doi.org/10.36713/epra8438
DOI No: 10.36713/epra8438

ABSTRACT

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. Here, 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.

KEY WORDS: *User profile matching, data privacy protection, ElGamal encryption.*

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 cancelled, most online matching sites may retain such information.

OBJECTIVE OF THE PROJECT

Therefore, the adversary advantage in guessing b is negligible. According to Definition, our protocols (including the extended protocols) have user profile privacy. Based on the query, multiple servers, which secretly share the decryption key, compare the preferred user profile with each record in the database. We formally define the user profile matching model, the user profile privacy and the user query privacy.

PURPOSE OF THE PROJECT

Our work is also closely related to homomorphic encryption, a form of encryption that allows computation on cipher texts, generating an encrypted result which, when decrypted, matches the result of the operations as if they had been performed on the plaintext. The purpose of homomorphic encryption is to allow computation on encrypted data. Typical homomorphic encryption schemes include the ElGamal, Goldwasser- Micali , Paillier schemes, which support either addition or multiplication on encrypted data, and the Boneh-Goh- Nissim scheme, which supports arbitrary additions and one multiplication (followed by arbitrary additions) on encrypted data. Fully homomorphic encryption schemes, e.g., the Gentry scheme, supports arbitrary computation on encrypted data, but has not become practical so far.

LITERATURE REVIEW

Information sharing across private databases

Literature on information integration across databases tacitly assumes that the data in each database can be revealed to the other databases. However, there is an increasing need for sharing information across autonomous entities in such a way that no information apart from the answer to the query is revealed. We formalize the notion of minimal information sharing across private databases, and develop protocols for intersection, equijoin, intersection size, and equijoin size. We also show how new applications can be built using the proposed protocols.

Server less 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 neighbourhood 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.

Space/time trade-offs in hash coding with allowable errors

In this paper trade-offs among certain computational factors in hash coding are analysed. The paradigm problem considered is that of testing a series of messages one-by-one for membership in a given set of messages. Two new hash-coding methods are examined and compared with a particular conventional hash-coding method. The computational factors considered are the size of the hash area (space), the time required to identify a message as a non-member of the given set (reject time), and an allowable error frequency. The new methods are intended to reduce the amount of space required to contain the hash-coded information from that associated with conventional methods. The reduction in space is accomplished by exploiting the possibility that a small fraction of errors of commission may be tolerable in some applications, in particular, applications in which a large amount of data is involved and a core resident hash area is consequently not feasible using conventional methods.

Evaluating 2-DNF formulas on cipher texts

Let be a 2-DNF formula on Boolean variables. We present a homomorphic public key encryption scheme that allows the public evaluation of ψ given an encryption of the variables. In other words, given the encryption of the bits, anyone can create the encryption of more generally, we can evaluate quadratic multi-variate polynomials on cipher texts provided the resulting value falls within a small set. We present a number of applications of the system.

Blind signatures for untraceable payments

Automation of the way we pay for goods and services is already underway, as can be seen by the variety and growth of electronic banking services available to consumers. The ultimate structure of the new electronic payments system may have a substantial impact on personal privacy as well as on the nature and extent of criminal use of payments. Ideally a new payments system should address both of these seemingly conflicting sets of concerns.

Practical private set intersection protocols with linear complexity, in Financial Cryptography and Data Security

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.

METHODOLOGY

We formally define the user profile matching model, the user profile privacy and the user query privacy. We give a solution for Confidentiality Protective User Profile Matching for a single dissimilarity threshold and then extend it for multiple dissimilarity thresholds. 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. 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.

PROCESS

The System Design Document describes the system requirements, operating environment, system and subsystem architecture, files and database design, input formats, output layouts, human-machine interfaces, detailed design, processing logic, and external interfaces.

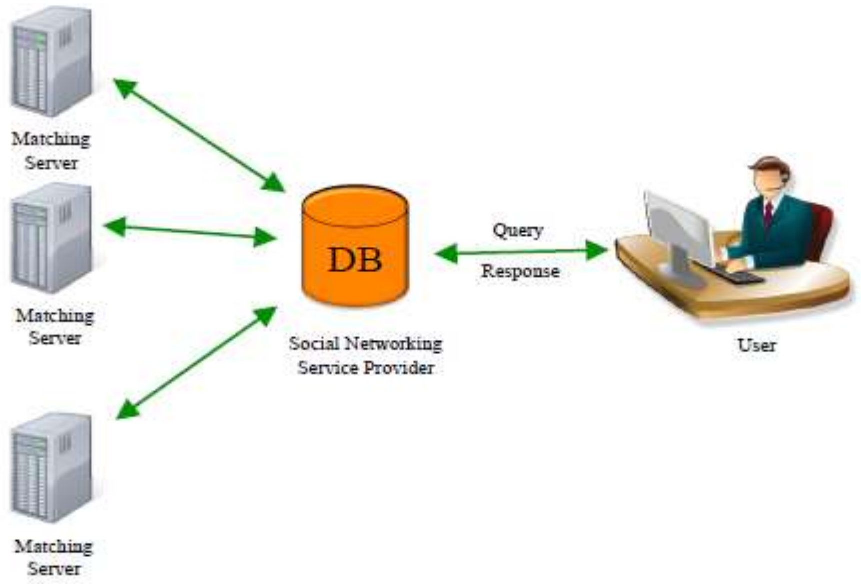


Fig 1. System Architecture

COLLABORATIVE DECRYPTION ALGORITHM

In the collaborative decryption algorithm there is only one decryption. The algorithm is designed so that the decryption result is either 0 or 1 and it is hard for the adversary to retrieve Ask1 0 from (B0=Ask1 0)rk , where rk is randomly chosen by the first server. Thus this algorithm also does not help the adversary win the game. In the collaborative encryption algorithm (Algorithm 5), the private sharing algorithm is used to share the contact information CIi of the user Ui at first. As we have shown, Algorithm 1 does not help the adversary win the game. After that, the first server encrypts CI(1) i with the public key PKU of the query user U.

The following implementation modules are a big part of the research work.

1. Model for Confidentiality Protective User Profile Matching
2. User profile matching
3. ElGamal encryption

Model for Confidentiality Protective User Profile Matching

Our model considers a social networking service environment with users and servers. Our model with one user, one database (DB) server and n matching servers is illustrated in our model, all users stores their profiles in the DB server in the service provider. User profile attributes are either sensitive or insensitive. We consider protection of sensitive attributes only. In addition, user profile attributes are either numeric (e.g., income) or categorical (e.g., address). We consider numeric attributes only in this paper except in Section.

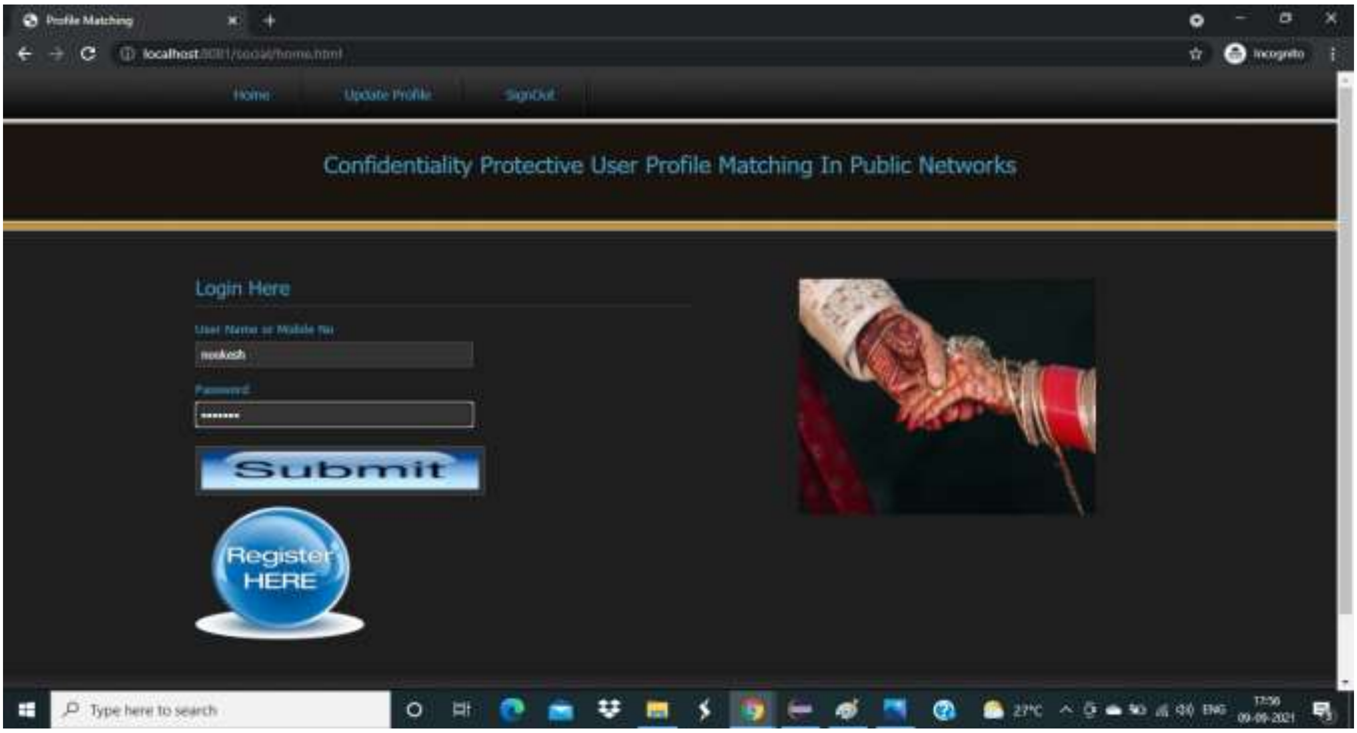
User Profile Matching

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.

ElGamal encryption

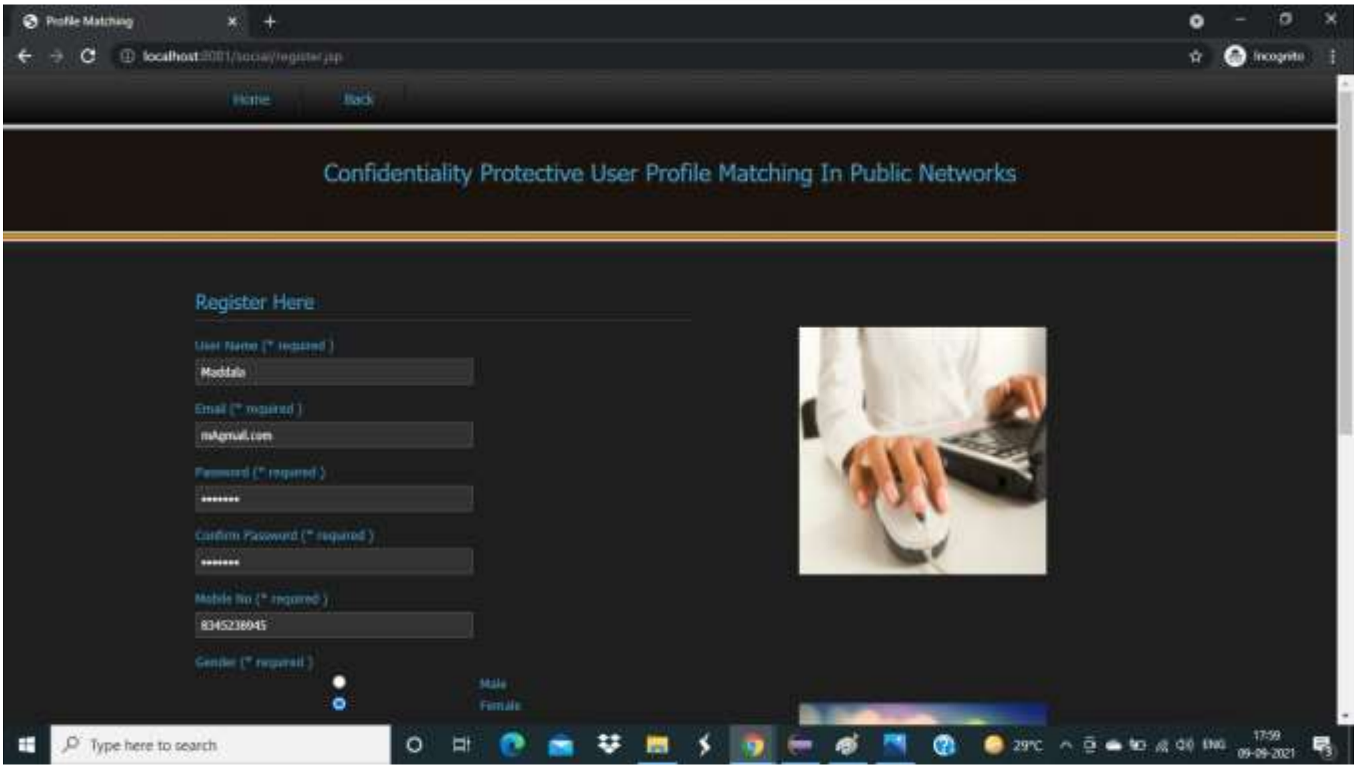
In our previous work, we use a variant ElGamal encryption scheme as the underlying homomorphic encryption scheme, which assumes the two prime factors of the modulus are public parameters. Has found a security flaw in the encryption scheme, that is, an attacker may decrypt the ciphertexts without the decryption key. In this paper, we fix the security flaw by keeping the factorization of the modulus secret. In our previous work, the user profile data is shared by all matching servers and therefore each matching server is required to maintain a user profile database. In this paper, we keep the user profile data in the social service provider only and therefore each matching server does not need to maintain any user profile database. In our previous work, a query user can specify only one dissimilarity threshold for user matching. In this paper, we allow the query user to specify multiple dissimilarity thresholds for user matching.

SAMPLE SCREENSHOTS


Profile Matching
localhost:8081/social/home.html
Home
Update Profile
SignOut
Confidentiality Protective User Profile Matching In Public Networks
Login Here
Submit
Register HERE

User Name or Mobile No: mukesh
Password: ••••••••

FIG 2. This is the login screen for this application.


Profile Matching
localhost:2001/social/register.jsp
Home
Back
Confidentiality Protective User Profile Matching In Public Networks

Register Here

User Name (* required)
Maddala

Email (* required)
m@gmail.com

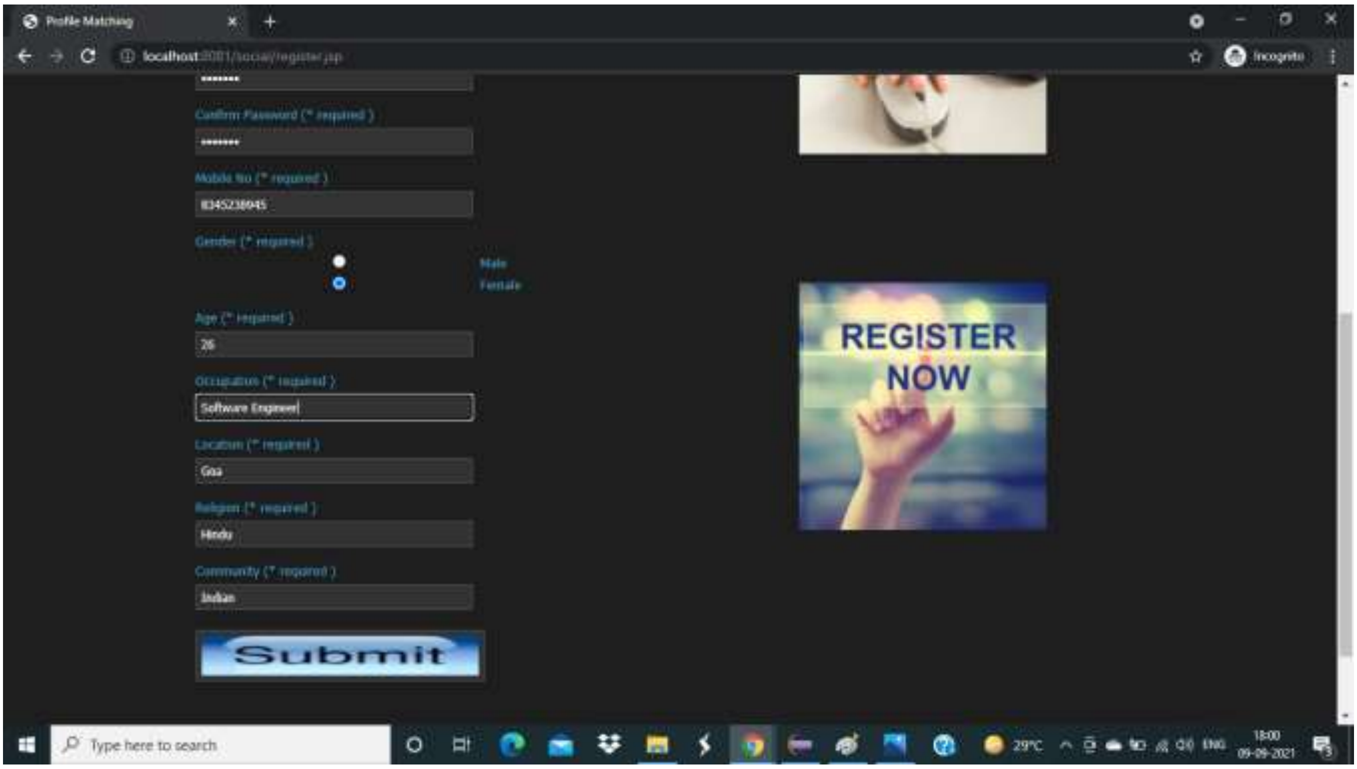
Password (* required)
****

Confirm Password (* required)
****

Mobile No (* required)
8345238945

Gender (* required)
○ Male
● Female

FIG 3. New user can register here with their details.



****

Confirm Password (* required)
****

Mobile No (* required)
8345238945

Gender (* required)
○ Male
● Female

Age (* required)
26

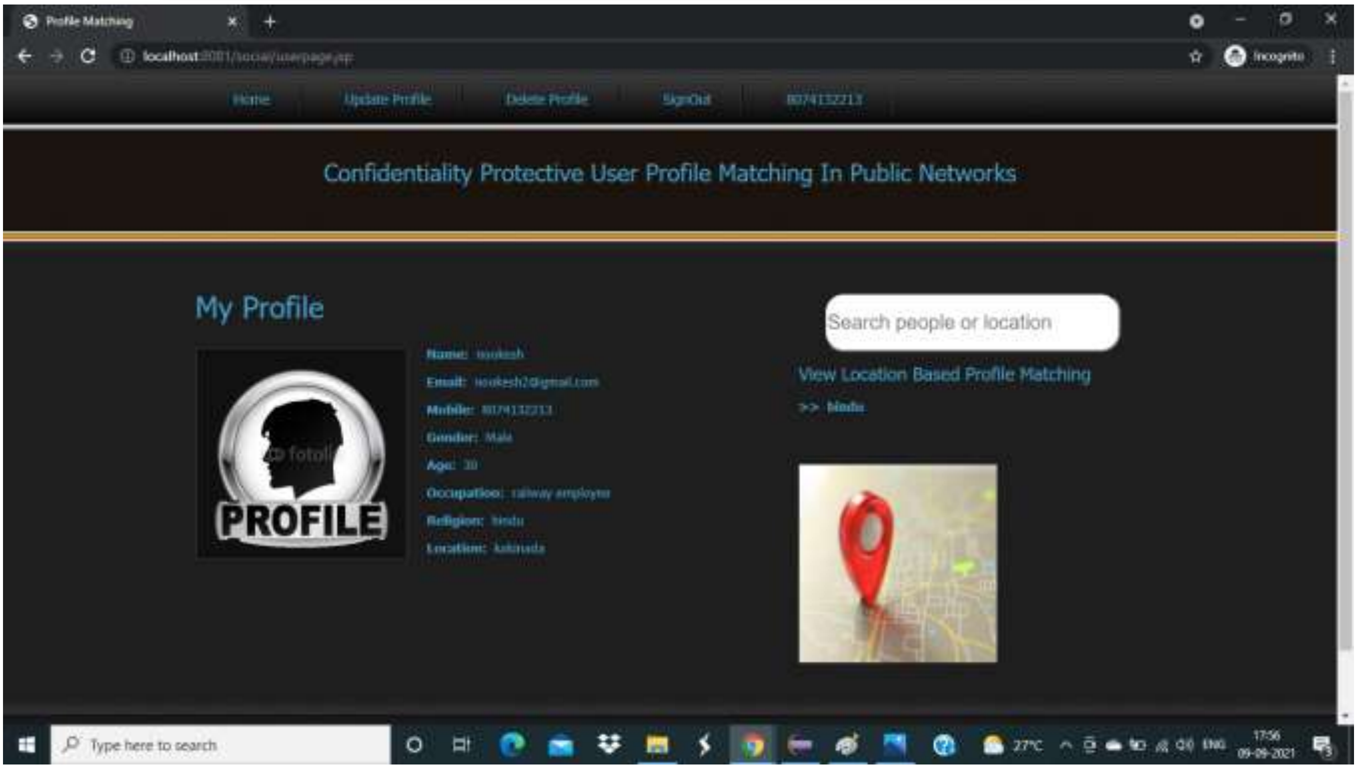
Occupation (* required)
Software Engineer

Location (* required)
Goa

Religion (* required)
Hindu

Community (* required)
Indian

FIG 4. Additional details can be provided here.



Home | Update Profile | Delete Profile | SignOut | 8074132213

Confidentiality Protective User Profile Matching In Public Networks

My Profile

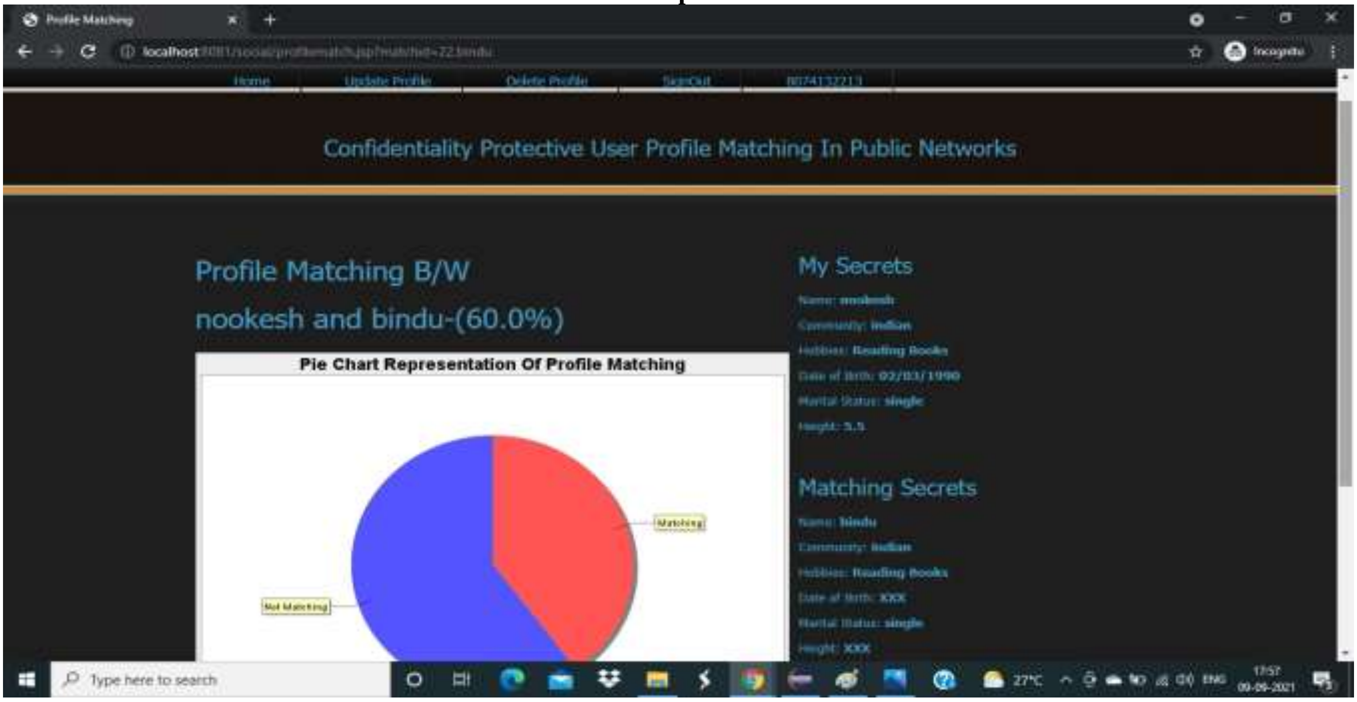
Name: nookesh
Email: nookesh2@gmail.com
Mobile: 8074132213
Gender: Male
Age: 30
Occupation: railway employee
Religion: hindu
Location: kakinada

Search people or location

View Location Based Profile Matching

>> bindu

FIG 5. After completing your registeration we can see our profile and we can search our related profiles based on their location or profile name.



Home | Update Profile | Delete Profile | SignOut | 8074132213

Confidentiality Protective User Profile Matching In Public Networks

Profile Matching B/W

nookesh and bindu-(60.0%)

Pie Chart Representation Of Profile Matching

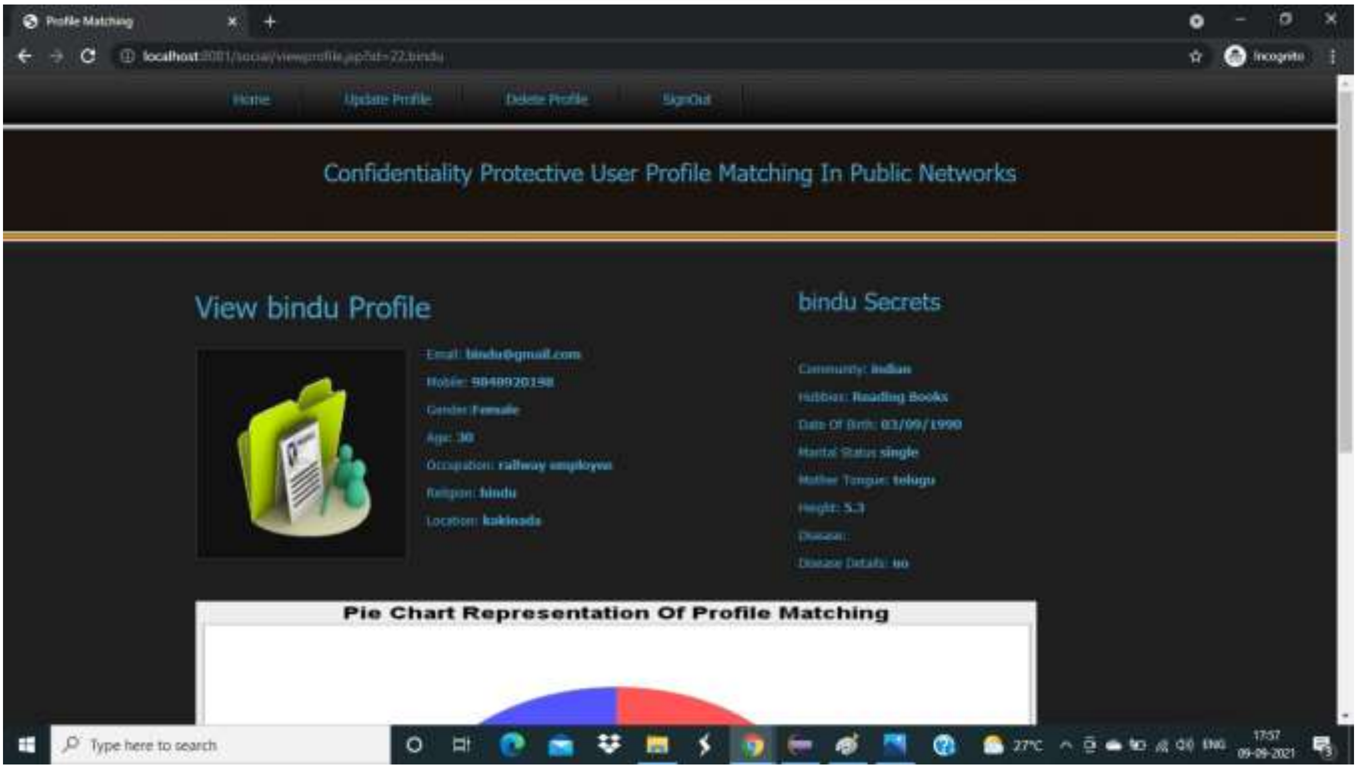
My Secrets

Name: nookesh
Community: indian
Hobbies: Reading Books
Date of Birth: 02/03/1990
Marital Status: single
Height: 5.5

Matching Secrets

Name: bindu
Community: indian
Hobbies: Reading Books
Date of Birth: XXX
Marital Status: single
Height: XXX

FIG 6. Now can find some profiles which are related to us . Then we can see either they have our matching interests or not and we can see the match percentage, if the percentage is above 50%then only we can see their interests or else we can’t.



Confidentiality Protective User Profile Matching In Public Networks

View bindu Profile

Email: bindu@gmail.com
Mobile: 9849920198
Gender:Female
Age: 30
Occupation: railway employee
Religion: hindu
Location: kakinada

bindu Secrets

Community: indian
Hobbies: Reading Books
Date Of Birth: 03/09/1990
Marital Status single
Mother Tongue: telugu
Height: 5.3
Disease:
Disease Details: no

Pie Chart Representation Of Profile Matching

FIG 7. And if we like that profile we can view that profile and see further more details.

CONCLUSION

In this paper, we proposed a new solution for Confidentiality Protective 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 169-17*