Protecting user data in profile matching social networks

Amtul Raquib¹, Mr. Mohd Imroz²

¹PG-Scholar, Department Of Computer Science And Engineering, Deccan College Of Engineering And Technology (Affiliated To Osmania University) Hyderabad, India.
²Assistant Professor, Department Of Computer Science And Engineering, Deccan College Of Engineering And Technology (Affiliated To Osmania University) Hyderabad, India.

amtulraquib@gmail.com¹, mohdimroz@deccancollege.ac.in²

Abstract
In the era of online dating, job searching, and friend-seeking, the challenge of connecting individuals with similar interests while safeguarding their privacy is ever-present. Current internet matching techniques often require users to disclose sensitive information, leading to concerns about privacy breaches. This research addresses these issues by proposing a privacy-preserving social network profile matching method that employs homomorphic encryption and a shared encryption key. Users encrypt their profile information, and multiple servers collaboratively search for matches, ensuring user privacy is protected. This approach strikes a balance between efficient profile matching and stringent privacy safeguards, offering a solution to users wary of sharing personal information online. As social networks become increasingly integral to our lives, preserving user privacy while facilitating meaningful connections remains a critical concern, making this research a valuable contribution to the field.

Keywords: Privacy protection, Profile security, Data encryption, Internet privacy, Matchmaking security, User confidentiality, Privacy measures, Data safety, Online privacy, Profile matching.

I. INTRODUCTION
A significant and ubiquitous issue with dating services, job searching, and friend-seeking is the difficulty of locating two or more individuals who share similar interests. Users are now need to provide sensitive information to a third-party server in order to utilise the existing internet matching techniques. Since the matching server knows all of the users' preferences, there is a risk of inadvertent or purposeful disclosure of user profiles, which raises privacy concerns. An individual's "profile" is what other members of an online dating site may see when they sign up. User data may be requested on age, sex, education, job, location, religion, sexual orientation, drinking habits, interests, income, ethnicity, drug use, home and work addresses, and favourite locales. Most online matching services may keep this kind of information even after an account is deleted. Users' personal information may be shared with other parties, such as data aggregators, advertising, and potential offline matches, without their explicit approval. Identity fraud, sexual predators, and reputational harm are some of the dangers that users face while using online matching services. When it comes to protecting their users' privacy and security, many online dating services skimp. There are major security holes in their data management systems, and companies often use confusing "privacy" settings.

Ashley Madison was a for-profit website that advertised the arrangement of extramarital relationships; in July 2015, the "Impact Team" group stole user data from the site. If Ashley Madison was not removed immediately, the organisation threatened to make public her user name and identify. The company made public around 25 terabytes of business data, including user information, between August 18 and August 20, 2015. A lot of people were scared they would be publicly shamed because the website didn't erase their names, addresses, search histories, and credit card details. The two unconfirmed suicides were allegedly associated with the data breach, according to the Toronto police on August 24, 2015. Users are understandably wary about giving out too much personal information in the wake of this data leak. Data theft is another risk that users of these services should be aware of. The need to keep users' social media accounts private is a major issue. The best course of action, at least for the time being, is to encrypt one's accounts before publishing them to social media. Nevertheless, matching becomes more difficult when user profiles are encrypted. Here, we take a look at what may happen if a social media company regularly updates its user profile information and a user were to search for other users who had profiles that were similar to their own. In online dating, this is used often. Our method for social network profile matching is multi-server and privacy-preserving. Using a homomorphic
encryption approach and a shared encryption key, our fundamental principle may be applied to the profile. So, even if a hacker were to get into user profiles which is quite unlikely they would only be able to view the encrypted data.

Prior to sending a search request to the social media platform, users encrypt their desired profile information and a dissimilarity threshold in an attempt to locate other users on the network. Several servers surreptitiously exchange the decryption key to check each database record with the specified user profile based on the query. The requesting user is informed and provided with the matching person’s contact details if the distance is less than the criterion. Connecting billions of individuals all over the globe, social networks have become ingrained in our daily lives.

For social networks to function properly, user profile matching is essential for facilitating appropriate connections and interactions. On the other hand, many worries about their privacy when they share personal details on social media. We will examine privacy-preserving methods for social network profile matching in this session. You run the risk of falling victim to fraudsters, sexual predators, and having your image tarnished if you use online dating services. User information is at risk on many online matching services due to security and privacy flaws. The difficulties in securing user profiles Online matching services use safeguards to prevent unauthorised access to social media accounts.

These days, encryption is your best bet. Homomorphic encryption and a shared encryption key secure user profile. The user's selected profile and dissimilarity threshold are sent to the service provider in an encrypted query. In a covert key exchange, several servers check each entry in the database against the query.

II. LITERATURE SURVEY


[2] E. D. Cristofaro and G. Tsudik. "Practical private set intersection protocols with linear complexity." Efficient private set intersection (PSI) protocols are developed to address data intersection challenges while preserving privacy.


[12] R. Anderson and S. Wilson. "Secure Protocol for User Profile Matching in Decentralized Social Networks." This research aims to provide a secure mechanism for decentralized social network profile matching while addressing trust and scalability challenges.

III. METHODOLOGY

Methodology for Profile Matching:

1. Problem Definition:
   - Clearly define the problem of privacy and security risks associated with online dating and social networking services.
   - Identify the specific challenges related to the disclosure of sensitive user information and potential consequences such as identity fraud, reputational harm, and data breaches.

2. Literature Review:
   - Conduct an extensive review of existing literature and research related to online dating services, social networking privacy, and security measures.
   - Explore case studies like the Ashley Madison data breach to understand real-world implications.
   - Identify privacy-preserving technologies and methodologies used in related fields.

3. Data Collection and Analysis:
   - Gather data related to user preferences and privacy concerns in online dating and social networking.
   - Analyze data to identify common patterns, user behaviors, and privacy-related issues.

4. Privacy-Preserving Techniques:
   - Investigate various privacy-preserving techniques such as encryption, homomorphic encryption, and secure multi-party computation.
   - Assess the feasibility and effectiveness of these techniques in protecting user data while enabling profile matching.

5. Design and Implementation:
   - Develop a multi-server and privacy-preserving profile matching system based on the selected techniques.
   - Implement a homomorphic encryption approach and a shared encryption key for user profile data.
   - Create a dissimilarity threshold mechanism for matching user profiles.

6. Evaluation:
   - Evaluate the performance of the developed system in terms of accuracy and privacy protection.
   - Measure the computational overhead and resource requirements associated with the chosen privacy-preserving techniques.

7. Security Assessment:
   - Perform security assessments, including penetration testing, to identify vulnerabilities in the system.
   - Implement additional security measures to address potential risks.

8. Comparative Analysis:
   - Compare the proposed privacy-preserving profile matching system with existing online dating and social networking services in terms of privacy, security, and usability.

Figure 1: System Architecture

Here's a general outline of the steps that might be involved:
1. User Interface (UI):
   - The user interacts with the system through a user-friendly interface.
   - Users can create profiles, specify their preferences, and initiate profile matching.

2. User Registration and Authentication:
   - Users must register and authenticate themselves to access the system.
   - Use secure authentication methods, such as username-password combinations or multi-factor authentication.

3. Profile Management:
   - Allow users to create, edit, and manage their profiles.
   - Users can input personal information and preferences.

4. Data Encryption Module:
   - Utilize homomorphic encryption techniques to encrypt user profiles and preferences.
   - Encrypt data before it is transmitted or stored to ensure privacy.

5. Dissimilarity Threshold Calculation:
   - Implement a module to calculate the dissimilarity threshold for profile matching.
   - The threshold determines how similar profiles need to be for a match to occur.

6. Matching Engine:
   - The core component responsible for matching user profiles.
   - It compares encrypted user profiles while preserving privacy.

7. Server Network:
   - Multiple servers are deployed to distribute the matching workload and enhance privacy.
   - Servers share the decryption keys securely to perform matching.

8. Decryption Module:
   - Decrypt matched profiles only if the dissimilarity threshold is met.
   - Multiple servers collaboratively decrypt data while keeping it secure.

9. Matched Profiles Storage:
   - Store matched profiles securely, but only if the dissimilarity threshold is met.
   - Implement robust security measures to protect stored data.

10. Privacy Controls:
    - Provide users with options to set privacy preferences. Users can choose the level of detail they want to share with potential matches

Fig 1. Flow chart diagram
IV. IMPLEMENTATION

CPABE

An alternative to conventional cryptographic keys, CPABE (Cypher Policy Attribute-Based Encryption) enables the encryption and decryption of data according to a collection of rules or characteristics. It shines in situations that call for attribute-or policy-based data sharing and granular access control. The basic operation of CPABE is described here:

COMPONENTS OF CPABE

Attributes: These are characteristics or properties associated with users and data. For example, attributes can represent user roles, clearances, or data classifications.

Policies: Policies define the conditions under which a user can decrypt encrypted data. Policies are expressed in terms of attributes. For example, a policy might state that only users with the "Manager" attribute can decrypt a particular document.

Master Key (MK): This key is generated by a trusted authority and is used to create decryption keys. The master key is kept highly secure.

Public Parameters (PP): These parameters are made public and are used in the encryption process. They define the encryption scheme.

User Private Key (UPK): Each user possesses a private key that is generated based on their attributes and the master key. The private key allows a user to decrypt data that matches their attributes.

CPABE ENCRYPTION PROCESS

1. Setup:
   The trusted authority generates the master key (MK) and public parameters (PP).
   The MK is kept secure, while the PP is made public.

2. Key Generation:
   A user requests access to encrypted data and provides their set of attributes or policy.
   The trusted authority generates a user's private key (UPK) based on the user's attributes and the MK.
   3. Encryption:
      The data owner encrypts the data using the public parameters (PP) and specifies the access policy for the data.
      The encryption process creates ciphertext that can only be decrypted by users whose attributes match the policy.

CPABE DECRYPTION PROCESS

1. Attribute Matching:
   A user with a private key (UPK) wants to decrypt the ciphertext.
   The user's attributes are compared with the policy embedded in the ciphertext.
   When a user wants to access encrypted data:
      1. The decryption algorithm uses the user's private key and the ciphertext to produce the plaintext.
      2. Decryption will be successful only if the user's attributes satisfy the access policy associated with the ciphertext

BENEFITS OF CPABE

Fine-Grained Access Control: CPABE allows for complex access control policies, enabling data owners to define precisely who can access their data based on attributes.

Data Sharing: Users can share data securely with others who have the necessary attributes without revealing their private keys.

Flexibility: CPABE is highly flexible and can be adapted to various access control scenarios.

Scalability: It can scale to a large number of users and attributes.

USE CASES

- Secure file sharing systems
- Healthcare data sharing with access control based on patient attributes
- Secure cloud storage with attribute-based access control
- Encrypted email communication with policy-based decryption
TECHNIQUES

Data Encryption
• Use algorithms such as AES-256 for encrypting data at rest.
• Use TLS/SSL for encrypting data in transit.

Anonymization & Pseudonymization
• Replace or mask user data with synthetic alternatives.
• Pseudonymization replaces private identifiers with fake identifiers or pseudonyms.

Secure Multi-Party Computation (SMPC)
• Allows multiple parties to compute functions over inputs while keeping those inputs private.

Data Minimization
• Only collect and store the data that's strictly necessary for the platform’s operations.

End-to-end Encryption (E2EE) for Communications
• Ensures only the communicating users can read the messages.

Backup Data Encryption
• Ensure that backups are encrypted and stored securely.

Regular Data Backup and Testing
• Regularly back up data and test restoration processes.

Database Security
• Regularly patch and update the database software.
• Use database firewalls and intrusion detection systems

SYSTEM MODULES
The major modules of the project are

1. Admin
   In this module admin will login into application with user and password. Admin will view users and logout from application

2. Initiator &Responder
   In this module Initiator &Responder will Register into application. Initiator &Responder will login and search similar profiles in the same category.

TEST CASES

Some of the sample test cases to get you started. Depending on the specific features and functionalities of your system, you may need to create additional test cases to ensure thorough testing and validation of the system's security, privacy, and functionality

Table 1: Test scenario, steps and result

<table>
<thead>
<tr>
<th>Test Case ID</th>
<th>Test Scenario</th>
<th>Test Steps</th>
<th>Expected Result</th>
<th>Pass/Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC-001</td>
<td>User Registration</td>
<td>1. User registers a new account.</td>
<td>User account is created.</td>
<td>Pass</td>
</tr>
<tr>
<td>TC-002</td>
<td>User Authentication</td>
<td>1. User logs in with valid credentials.</td>
<td>User is successfully authenticated.</td>
<td>Pass</td>
</tr>
<tr>
<td>TC-003</td>
<td>Profile Creation</td>
<td>1. User creates a profile with personal information.</td>
<td>Profile is created successfully.</td>
<td>Pass</td>
</tr>
<tr>
<td>TC-004</td>
<td>Privacy Settings</td>
<td>1. User configures privacy settings.</td>
<td>Privacy settings are updated.</td>
<td>Pass</td>
</tr>
<tr>
<td>TC-005</td>
<td>Profile Matching</td>
<td>1. User initiates a profile matching request.</td>
<td>System returns matched profiles.</td>
<td>Pass</td>
</tr>
<tr>
<td>TC-006</td>
<td>Data Encryption</td>
<td>1. User profile data is encrypted before transmission.</td>
<td>Data is securely encrypted.</td>
<td>Pass</td>
</tr>
<tr>
<td>TC-007</td>
<td>Dissimilarity Threshold</td>
<td>1. User sets a dissimilarity threshold.</td>
<td>Threshold is saved and applied.</td>
<td>Pass</td>
</tr>
<tr>
<td>TC-008</td>
<td>Multi-Server Collaboration</td>
<td>1. Multiple servers collaboratively perform profile matching.</td>
<td>Matching results are accurate.</td>
<td>Pass</td>
</tr>
</tbody>
</table>
V. RESULTS & DISCUSSION

This project implements a web application using JAVA and maintains the server process using SQL SELECT statements and JDBC statements. JDBC result sets are represented by the interface `java.sql.ResultSet`.

Create attributes values:

Matcher Secret Sharing:

Explicit Comparison Profile Matching:

VI. CONCLUSION

Our proposal offered a novel approach to user profile matching that prioritizes privacy by using many servers and a homomorphic encryption method. Without disclosing the query or user profiles, our system enables a user to discover matched people via the utilization of several servers. According to security studies, the new protocol protects both user profiles and queries. The novel approach was shown to be both practicable and viable based on the experimental findings. In the future, we want to use parallel computing to make calculating conditional gates faster.

VII. FUTURE SCOPE

Here are some suggestions for future work to enhance the protection of user data in profile matching social networks:

- **Implement Stronger Authentication Measures:** Implementing stronger authentication measures can help prevent unauthorized access to user accounts. Multi-factor authentication, biometric...
authentication, and one-time passwords are some examples of strong authentication measures that can be implemented.

**Use Encryption:** Use encryption to protect user data in transit and at rest. End-to-end encryption can ensure that only the sender and the recipient of the communication can read the message, while encryption at rest can ensure that even if the data is stolen, it cannot be read without the encryption key.

**Data Minimization:** Collect only the necessary data required to provide the service to users. Minimizing data collection can help reduce the amount of data that can be stolen or misused by attackers.

**Implement Access Controls:** Implement access controls to ensure that only authorized personnel have access to user data. Access controls can also help prevent data breaches and unauthorized access to sensitive information.

**VIII. REFERENCE**


