Abstract: In this paper, we study user profile matching with privacy-preservation in mobile social networks (MSNs) and introduce a family of novel profile matching protocols. We first propose an explicit Comparison-based Profile Matching protocol (eCPM) which runs between two parties, an initiator and a responder. The eCPM enables the initiator to obtain the comparison-based matching result about a specified attribute in their profiles, while preventing their attribute values from disclosure. We then propose an implicit Comparison-based Profile Matching protocol (iCPM) which allows the initiator to directly obtain some messages instead of the comparison result from the responder. The messages unrelated to user profile can be divided into multiple categories by the responder. The initiator implicitly chooses the interested category which is unknown to the responder. Two messages in each category are prepared by the responder, and only one message can be obtained by the initiator according to the comparison result on a single attribute. We further generalize the iCPM to an implicit Predicate-based Profile Matching protocol (iPPM) which allows complex comparison criteria spanning multiple attributes. The anonymity analysis shows all these protocols achieve the confidentiality of user profiles. In addition, the eCPM reveals the comparison result to the initiator and provides only conditional anonymity; the iCPM and the iPPM do not reveal the result at all and provide full anonymity. We analyze the communication overhead and the anonymity strength of the protocols. We then present an enhanced version of the eCPM, called eCPM+, by combining the eCPM with a novel prediction-based adaptive pseudonym change strategy. The performance of the eCPM and the eCPM+ are comparatively studied through extensive trace-based simulations. Simulation results demonstrate that the eCPM+ achieves significantly higher anonymity strength with slightly larger number of pseudonyms than the eCPM.

Keywords: Mobile Social Network, Profile Matching, Privacy Preservation, Homomorphic Encryption, Oblivious Transfer.

I. INTRODUCTION

Social networking makes digital communication technology tools to expand the social circle of people sharpening. It has already become an integral and important part of our daily lives, allowing us to contact our family and friends in time. As reported by Com Score [2], the social networking sites like Facebook and Twitter have reached 82 percent of the world's online population, representing 1.2 billion users around the world. Meanwhile, driven by the widespread adoption of advanced hand devices and ubiquitous network connections Bluetooth / Wi-Fi / GSM /LTE, the use of mobile social networking (MSN) has exploded. In MSN, users are able to not only surf the Internet, but also communicate with their peers in the vicinity that use the short-range wireless communications [3] - [7]. Due to its geographical nature, the MSN support many promising and innovative applications [8] for example, through the Bluetooth communications, People Net [8] allows searching for effective information between mobile phones neighbors; A message-relay approach is suggested in [9] to facilitate ride sharing and ride sharing in a local region. Realizing the potential benefits presented by MSN, recent research efforts have been made on how to improve the effectiveness and efficiency of communications among users of MSN [10], [12]. They developed specialized routing protocols and data forwarding associated with the social characteristics exhibited by the behavior of users, such as social friendship, [10], social selfishness [12] and social morality. It is encouraging that traditional solutions can be expanded further to troubleshoot MSN, considering the unique social characteristics.

Privacy preservation is an important research topic in social networks. Given that more personal information is shared with the public, violating the privacy of a target user is much easier. Research efforts have been put into the presentation of identity and privacy issues on social networking sites. Gross and Acquisti argued that users are jeopardizing both offline (eg, stalking) and online (eg, identity theft) based on an analysis of the behavior of more than 4,000 students they have joined a popular social network presented a quantitative analysis of identity information disclosure in social network communities and the subjective opinions of students regarding the identity protection and information disclosure. When social networking platforms extend into the mobile environment, users require more extensive privacy preservation because they are unfamiliar with neighbors nearby that can spy, store and correlate their
personal data in different periods and places. Once personal information is correlated with the location information, the user behavior will be fully disclosed to the public. Chen and Rahman studied various mobile social networking applications (SNAS), such as neighborhood exploring applications for mobile and SNAs specific content sharing applications, which do not provide feedback or control mechanisms for users and can cause localization inappropriate and identity information disclosure.

To overcome the violation of privacy on MSN, many techniques of privacy have been taken in MSN [4] applications. For example, when two users are on the MSN, privacy-preserving matching profile acts as a critical first step to help users, especially to strangers, initialize the conversation with others in a manner and privacy preserving distributed. Many research efforts on privacy preserving matching profile have been carried out. The common objective of these works is to allow the handshake between two users found if users meet the requirement of one another, while eliminating unnecessary disclosure of information if they are not. The original idea of, where an agent of the Central Intelligence Agency (CIA) wants to authenticate herself to a server, but does not want to show their CIA credentials unless the server is a shot of the CIA authentic meanwhile, the server does not want to reveal its CIA credentials to anyone but CIA agents. At MSN, we consider a generalized function to support the exchange of information through the use of a matching profile as a metric. Following the above example, two CIA agents are considered with two different priority levels in the system of the CIA, A with a low priority lA and B with a high priority lB. They are known as an agent of the CIA. However, they do not want to reveal their levels of priority among them. B wants to share some messages to A.

The messages are not related to the user profile, and are divided into several categories, for example, messages related to the different regions (New York or Beijing) in different years (2011 or 2012). B shares a message of a certain category T at once. T category is chosen, but the choice is unknown to B. For each category, B makes two self-defined messages, for example, a low confidential message for the CIA at a lower level and high confidential message for the agent on a higher level. Because lA < lB, A eventually obtains the low-confidential message without knowing that it is a low confidential one. In the meantime, B does not know which message A receives. The above function offers both A and B the highest anonymity since neither the comparison result between lA and lB is disclosed to A or neither B nor the category T of A’s interest is disclosed to B. In the following, we refer to A as the initiator uA, B as the responder uB, the attribute used in the comparison (i.e., priority level) as ax, and the category T of A’s interest as T. The attribute values of uA and uB on the attribute ax are denoted by aA and aB, respectively. We first formally describe two scenarios from the bellow examples and as shown in Fig.1.

**Scenario 1:** The initiator wants to know the result of the comparison, that is, if you have a larger, equal, or less than the responder in a specified attribute value.

**Scenario 2:** The initiator expected response actions of a message related to the category of your interest, yet remains unknown to the responder. Meanwhile, the responder wants to share with the originator of a message is determined by the result of the comparison of their attribute values.

![Fig.1.Scenarios.](image-url)

**II. RELATED WORK**

Mobile social networking and emerging social media platforms have attracted much attention recently, and mobile applications have been developed and pervasive practice. In applications of mobile social networks, matching profile acts as a critical first step to help users, especially to strangers, initialize the conversation with others in a distributed manner introduced a mobile communication system distributed, called E-Small Talker, which facilitates social networking in physical proximity. E-Small Talker automatically discovers and suggests common themes among users to facilitate the conversation studied the case of e-healthcare by proposing a scheme to compensate for the symptoms of social networks of mobile health. In his opinion, this game system is valuable for patients with the same symptoms to exchange their experiences, mutual support, and inspiration to others. In general, the matching profile can be categorized based on the formats of the profiles and the types of gambling operations a well-known coincidence profile FNP scheme, where a client and a server to calculate the intersection set so that the client gets the result, while the server learns nothing.

Later Kissner et al applied a matching profile with several operations including set intersection, union, cardinality and excess threshold operations. Moreover, further extended the FNP scheme with a scheme of private correspondence and distributed designed to reduce protocol complexity. All the above solutions to the set intersection operation depend homomorphic encryption. Meanwhile, other studies uses an oblivious pseudorandom function to build their profile matching protocols where communication and computational efficiency is improved implemented matching profile according to three increasing levels of privacy: i) developing
the common attribute set of the two users; ii) disclose the size of the common attribute set; and iii) revealing the size range of the common attribute is set between a user and their neighbors. In his view, a person honest, but-curious (HBC) adversarial model, which assumes that users try to obtain more information than permitted by the inference results that match the profile, but the truth, follows the protocol. Secure multiparty computation applied the scheme of Shamir secret sharing and homomorphic encryption scheme to achieve the confidentiality of user profiles.

In another category matching profile, the profiles can be represented as vectors, and the operation can be matching or distance domestic product. Such matching profile is a special case of computing the two parties insurance, which was initially introduced by Yao and later generalized to secure multiparty computation by specifically, we present two recent works in this category considered user profile consisting of attribute values and the proximity of two user profiles measured using dot product $f_{dot}(u,v)$. An existing scalar product protocol has been enhanced to allow secure verifiable computation. The enhanced protocol reveals only if the dot product is above or below a given threshold. The threshold value is selected by the user who starts the matching profile. They noted the potential risk of their anonymity protocols; an adversary can adaptively adjust the threshold value to quickly narrow the range of values of the profile of the victim. Therefore, it is required that the threshold value must be greater than a pre-defined lower limit (a system parameter) to ensure user anonymity.

The same problem exists in other studies. Furthermore, users must make a commitment on their profiles to ensure consistency in profile, but the profile forgery attack can still occur during the commitment phase. Profile protocols proposed matching items are new since the comparison of attribute values are considered as the corresponding operation. The intuitive idea is inspired by the famous millionaire problem Yao solution. As in other studies we propose three different protocols with different levels of anonymity. ECPM for conditional anonymity, anonymity and offer a detailed show the relationship between change and variation analysis pseudonym anonymity for the ICPM and IPPM with full anonymity, we show that the use of these protocols does not affect the level of user anonymity, and users are able to fully preserve their privacy.

III. EXPLICIT COMPARISON BASED APPROACH

eCPM protocol allows two users to compare their attribute values in a given attribute without revealing the values together. However, the protocol reveals the result of the comparison to the initiator, and therefore provides conditional anonymity. The protocol has a fundamental pre-program phase, where the TCA generates all system parameters, user pseudonyms and inlay materials.

A. Bootstrapping

The protocol has a fundamental bootstrapping phase, where the TCA generates all system parameters, user pseudonyms, and keying materials. Specifically, the TCA runs $\mathcal{F}$ to generate $\langle \rho, q, R, R_\rho, X \rangle$ for initiating the homomorphic encryption. The TCA generates a pair of public and private keys $(\rho^k_{\text{TCA}}, S^k_{\text{TCA}})$ for itself. The public key $\rho^k_{\text{TCA}}$ is open to all users; the private key $S^k_{\text{TCA}}$ is a secret which will be used to issue certificates for user pseudonyms and keying materials, as shown below. The TCA generates disjoint sets of pseudonyms $(\text{pid})$ and disjoint sets of homomorphic public keys $(\text{pk})$ for users $(\text{ui})$. For every pid, and pk, of $\text{ui}$, the TCA generates the corresponding secret keys psk and sk. In correspondence to each pseudonyms pid, it assigns a certificate $\text{cert}_{\text{pid}}$ to $\text{ui}$ which can be used to conform the validity of pid. Generally, the TCA users $S^k_{\text{TCA}}$ to generate a signature on pid, and pk. The TCA outputs $\text{cert}_{\text{pid}}$ as a tuple $(\text{pk}, \text{Sign}_{\text{TCA}}(\text{pid}, \text{pk}))$. The homo-morphic secret key $sk$ is delivered to $\text{ui}$ to gather with $\text{psk}$; pk is tied to pid, and varies as the change of pseudonyms.

IV. IMPLICIT COMPARISON BASED APPROACH

Here the profile corresponding implicit base (ICPM), the adoption of the unconscious transfer cryptographic technique is proposed. It is considered that users have different values for any given attribute. The ICPM consists of three main steps as shown in Fig.2. In the first step, a category of interest by setting element to 1 and other elements to 0 lengths, the vector then encode the vector using the homo-morphic encryption and sends the encryption vector, but it can still be processed in the cipher text. In the second step, calculates the input cipher texts of messages defined for $1\leq \text{message} \leq \text{length}$.

V. IMPLICIT PREDICTABLE BASED APPROACH

Both eCPM and ICPM perform matching profile on a single attribute. For a game involving multiple attributes, must be performed several times, each time for one attribute. In this section, IPPC extends to cases of multiple attributes, without compromising their anonymity property, and obtain a protocol implicitly Profile Matching based on predicates, i.e. IPPM is a shown in Fig.3. This protocol is based on a predicate that is a logical expression made of multiple comparisons covering different attributes and therefore supports sophisticated matching criteria in a single procedure protocol.
leads to yet better anonymity performance than the post-adaptive one. It shows that in case of 5-anonymity and \( \text{th} = 0.0763 \), the pre-adaptive strategy consumes 449(>369) pseudonyms and results in a 0.0445(<0.0514) break ratio. The pre-adaptive strategy consumes slightly more pseudonyms, but achieves significantly shorter anonymity break period.

VII. CONCLUSION

A unique comparison-based profile matching problem in Mobile Social Networks (MSNs) has been investigated, and novel protocols are proposed to solve it. The explicit Comparison based Profile Matching (eCPM) protocol provides conditional anonymity. It reveals the comparison result to the initiator. Considering the k-anonymity as a user requirement; the anonymity risk level in relation to the pseudonym change for consecutive eCPM runs is analyzed. Further an enhanced version of the eCPM, i.e., eCPM+ is introduced, by exploiting the prediction-based strategy and adopting the pre-adaptive pseudonym change. The effectiveness of the eCPM+ is validated through extensive simulations using real-trace data. Two protocols with full anonymity, i.e., implicit Comparison-based Profile Matching (iCPM) and implicit Predicate-based Profile Matching (iPPM) has been devised. The iCPM handles profile matching based on a single comparison of an attribute while the iPPM is implemented with a logical expression made of multiple comparisons spanning multiple attributes. The iCPM and the iPPM both enable users to anonymously request for messages and respond to the requests according to the profile matching result, without disclosing any profile information. In current version of the iCPM and the iPPM, \( \geq \) and \( \leq \) operations for profile matching is implemented. One future work is to extend them to support more operations, such as \( \geq \) and \( \leq \). Currently, the responder needs to transmit the threshold value of the predicate to the initiator. Restricting the disclosure of such parameter will be of significance for advancing comparison-based family of profile matching protocols and warrants deep investigation.

VIII. REFERENCES

[1] Xiaohui Liang, Student Member, IEEE, Xu Li, Kuan Zhang, Rongxing Lu, Member, IEEE, Xiaodong Lin, Member, IEEE, and Xuemin (Sherman) Shen, Fellow, IEEE, “Matching Profile Anonymously in Mobile Social Networks”, IEEE Transactions on Networking Year 2013.
Fully Anonymous Profile Matching in Mobile Social Networks