Privacy-Preserving Zero Knowledge Scheme for Attribute-based Matchmaking

Solomon Sarpong
Department of Biological, Physical and Mathematical Science,
School of Natural Resources and Environmental Sciences,
Somanya, Ghana.
ssarpong@uesd.edu.gh

***Abstract*- Making friends with common attributes is a characteristic of some persons. This characteristic is also extended to matchmaking on social networks. In some of the existing matchmaking protocols, the users are match-paired without considering the number of attributes they have in common. Furthermore, the bane of the existing proposed matchmaking protocols has been how to preserve the users' privacy and this has been considered as the key security issue for such applications. In order to prevent malicious users from gaining extra information, the attributes in this protocol will be certified. Hence, certification authority ensures that, a user actually possesses the input attributes and binds them to him/her. With the use of certified sets and zero knowledge proofs, users can adequately find a matching-pair whilst keeping his/her input set private. Furthermore, in this proposed protocol, a person can find a best match among potential candidates by finding the one who has the maximum number of common attributes. At the end of the protocol, the match-pair can exchange their attributes without any other person knowing the number or the type of attributes they have in common.**

***Keywords*- Certification, malicious, matchmaking, validation, zero-knowledge.**

I.INTRODUCTION

There are scenarios where two or more parties with private set of sensitive information need to find the intersection between these sets. Due to the sensitivity of their private sets, each individual needs to know only the content of the intersection but nothing else. This has necessitated the need for private set intersection.

Usually, in private set intersection one or both users in the protocol obtain the intersection (if an intersection exists). It has become a necessity to control the sharing of sensitive information in recent times. Most often than not, the sharing of information involves two parties; one seeking the information and the other maybe willing or under compulsion to share it.

In such scenarios, the dilemma is impossible to solve unless one party sacrifices some privacy. In this case, the considerations are; (1). Can the information be shared such that nothing else can be learnt apart from what the party is entitled to know? (2). practically, how can it be done? [1].

1. (Authorised) Private Set Intersection:

Private Set Intersection (PSI) is a cryptographic protocol that allows two users to find the intersection of their sets such that neither party can infer elements in the other party's set that are not in the intersection. Thus the parties learn only the content of the intersection but nothing else [2]. However, PSI protocols allow each party to place any element in their own set. Hence, malicious persons in the protocol can claim ownership of any private set of attributes so as to know more information about the other. Authorized Private Set Intersect (APSI) and its variants [3], [4], [5] [6], [7], and [8] are necessary to solve this problem associated with PSI.

In APSI, the elements of each person in the intersection are certified. The certification prevents an individual to claim to have some attributes s/he does not have. The goal of authorization of the private set is to restrict their input and hence reduce the strength of a malicious participant.

Authentication of the attributes is of more importance in distributed systems. Usually, in the distributed system there is mutual authentication of each other – the system verifies the identity of the user and the user also verifies the identity of the system [9].

Basically, PSI consists of two algorithms namely Setup and Interaction. In the Setup stage, the parameters are selected. During the Interaction stage the intersection between the private set of the client and server are computed, only the client gets to know the intersection. In addition to the client and server in PSI, APSI has a CA (off-line).

On the other hand, APSI is a tuple of three algorithms; Setup, Authorize and Interaction. In the Authorize stage, the CA signs each of the elements in the set of the client [10].

2. Security Properties:

There are security requirements concerning [5]. These security requirements include correctness, server and client privacy, and server and client unlinkability. Correctness – the scheme is correct if the client outputs the exact content of the intersection (may be empty) of the two users when the protocol ends. Server Privacy – the client gets to know only the size of the intersection set but nothing else from the server. Client Privacy – the server learns nothing from the client apart from the size of the intersection set. Client unlinkability – a malicious server cannot know the relationship between two computed intersection sets. Also, server unlinkability – a malicious client cannot know the relationship between two computed intersection sets. However, in APSI, Correctness means the client outputs the intersection of the two sets and each of them has been authorized by the CA.

3. Zero-knowledge Proof:

Zero-knowledge proofs (ZKP) are of considerable theoretical and practical interest to mathematicians and cryptographers alike. ZKP’s achieve the seemingly contradictory goals of proving a statement without revealing anything other than the fact that the statement is indeed true [11]. Thus, a zero-knowledge proof is a way that a “prover” can prove possession of a certain piece of information to a “verifier” without revealing it. This is done by manipulating data provided by the verifier in a way that would be impossible without the secret information in question. It should be noted that a third party, reviewing the transcript created, cannot be convinced that either “prover” or “verifier” knows the secret.

ZKP must satisfy three properties; completeness, soundness, and zero-knowledge. Completeness property – an honest prover convinces the honest verifier that a statement is true, if it is true. Soundness property – the probability that a cheating prover can convince an honest verifier that a false statement is true is very small. Zero knowledge property – nothing else is learnt by a cheating verifier if a statement is true.

Most ZKP are three pass protocols. This means that three messages are transmitted between a “prover” and a “verifier”. These are commitment, challenge, and response. Randomness and timing is a couple of other properties typically associated with ZKP. Randomness in the commitment and challenge are used to hide the secret information. Timing is used to prevent an adversary from taking a long time to calculate an answer. A ZKP can also be interactive or non-interactive. Practical applications of ZKP include [9], Guillou-Quisquarter (GQ) and Schnorr [12] protocols.

The main contribution of this work is use APSI with zero-knowledge proof to formulate a matchmaking protocol that is very efficient and secure. In the proposed protocol, not only does the initiator find a matching pair, but the best pair. Our proposed protocol is secure against malicious attacks. In order to prevent this form of attack, there is the use zero knowledge proof in our proposed protocol. Furthermore, apart from the matched-pair that is privy to their common attributes, no one else does.

The paper is organized as follows; the introduction is in section one. This is followed by related works in section two. The proposed protocol is in section three. The section four contains the conclusion.

II. RELATED WORKS

When two or more persons having individual attributes want to find how many attributes they have in common, the use of private set intersection becomes necessary. This intersection technique will enable them find only the common attributes between them. The quest to know how many attributes they have in common will enable them knows how compatible they may be as friends. This is the underlying motive for most matchmaking protocols.

Private set intersection and its variants [3], [4], [5], [8], and [13] can be used to find the intersection between data sets that are on different sources. They enable the intersection between the sets to be found whilst maintaining the privacy of the sets. In the quest to match-pair individuals by computing the intersection of their attributes, there is the use of; (1). Trusted central server [14]. (2).

The fully distributed system [15], [16], [17], and [18]. The hybrid technique [19], and [20], [24, 25, 26]. This protocol is a combination of the first two. Other protocols used in privacy preserving attribute matchmaking include [21].

III. THE PROTOCOL

This protocol basically tries to help users find the most appropriate pair in a matchmaking protocol. To this end, the protocol first allows the initiator (Alice) to look for an individual with the appropriate number of common attributes to qualify as a matched-pair. Alice and the would-be pair then go on to find their intersection.

Basically, our matchmaking protocol comprises Alice finding the appropriate number of same matching attributes for a potential match and the matched candidates exchanging their common elements. In this matchmaking protocol, malicious participants are prevented from manipulating the input set with the use of APSI-CA protocol. The authorization of each input set is done by a certification authority, CA.

This proposed protocol based on APSI-CA protocol will be secured against a malicious sever and semi-honest client. For the sake of simplicity and clarity of the protocol, we assume that Alice and Bob are the only participants in the protocol. In our protocol, Alice has input. The CA authenticates this input by computing for each attribute. Hence, Alice's input becomes. Bob inputs, the CA authenticates his attributes by computing for each attribute.

Hence, his input becomes A lice chooses a random number and after calculating, broadcasts. Bob on receiving computes and sends back to Alice. Alice then goes on to compute and the absolute intersection | |. Upon, calculating the intersection, she is able to know the number of attributes she has in common with Bob. If the number of common attributes is sufficient enough for Bob to qualify as a matched-pair, she then continues with the next stage of the protocol.

In order for Alice and Bob to exchange their common attributes. Alice chooses, calculates and sends to Bob. Alice and Bob undertake zero knowledge proofs to verify. If Alice is not able to prove to Bob that is valid, Bob terminates the protocol. On the other hand, if Alice is able to prove to Bob that is valid, he continues the protocol. Bob then goes on to calculate and send and to Alice. Bob and Alice undertake zero knowledge proofs to verify. If Bob is not able to prove to Alice that is valid, Alice terminates the protocol.

On the other hand, if Bob is able to prove to Alice that is valid, she continues the protocol. Alice then computes and sends to Bob. Both Alice and Bob output. At this point, though both Alice and Bob know the intersection of their inputs, they do not actually know the elements they have in common.

To know the elements they have in common, Alice sends to the CA. Bob also sends to the CA. The CA then checks if the from Alice is the same as that from Bob. If they are the same, the CA sends confirm messages to notify them that their matchmaking is successful. At this point, both Alice and Bob will be able to know the actual elements they have in common.

IV. SECURITY

The certification authority, CA, in our protocol only certifies the input sets of the members in the protocol but does not take part in the matchmaking. This step is important in the protocol as it prevents members from inputting fake elements and it is done just once. When the security parameter k is input, the CA generates an RSA modulus where and picks random elements.

The components of the RSA are chosen such that, e is a small prime, and of. The CA keeps (p, q, and d) as the secret key and the public parameters are. The CA also chooses hash functions In order to certify the input set of a member of the protocol; the CA does the following computation;

- **Correctness:** For any held by Alice, and held by Bob, if: (1) is genuine, CA signature on and (2), hence, we obtain.
- **Privacy:** Bob inputs and Alice input at the end of the protocol, both learn the size of the intersection.
- **Efficiency:** The APSI-CA protocol in this match making involves linear computation.

The communication complexity performed by the client is modular exponentiations. The server also does modular exponentiations; the exponents are taken from the RSA settings. Furthermore, this protocol is characterized by its security in a malicious model in ROM under the RSA and DDH assumptions.

Hence, with the hardness of the RSA and DDH problems and being zero-knowledge proofs, the protocol is a secure computation in ROM. The algorithm is secure in the presence of malicious attacks from participants in the protocol. In our proposed protocol, external attacks are not considered as it is robust against would be external attacks. When the protocol ends, both Alice and Bob know the intersection. In order to know the actual attributes between them, they contact the CA. This to a large extent prevents information asymmetry.

V. CONCLUSION

Mobile social networks is gaining popularity hence, it has become necessary to develop efficient, secure and practical protocols that will allow users feel confident when using them. Based on zero-knowledge proofs and authorised private set intersection with cardinality, we present a new privacy preserving matchmaking protocol. By this, proposed protocol, users can find match-pairs without leaking any private information.

The proposed protocol will enable an individual to find the best match from among many potential candidates. They can then exchange their common intersection attributes. This can be done without any other user knowing any information about the matched-pair. Our protocol for matchmaking preserves users' information from unnecessary leaks in mobile social networking applications.

REFERENCES

[1] L. Kissner and D. Song, "Privacy-Preserving Set Operations," in Advances in Cryptology -- CRYPTO 2005, 2005, pp. 241–257.
[2] G. Ateniese, E. De Cristofaro, and G. Tsudik, "(If) Size Matters : Size-Hiding Private Set Intersection ∗," 2011.

[3] J. Camenisch and G. M. Zaverucha, "Private Intersection of Certified Sets," in Financial Cryptography and Data Security, 2009, pp. 108–127.

[4] E. De Cristofaro, S. Jarecki, J. Kim, and G. Tsudik, "Privacy-Preserving Policy-Based Information Transfer," in Privacy Enhancing Technologies, 2009, pp. 164–184.

[5] E. De Cristofaro and G. Tsudik, "Practical Private Set Intersection Protocols with Linear Computational and Bandwidth Complexity ∗," pp. 1–17, 2010.

[6] E. De Cristofaro, P. Gasti, and G. Tsudik, "Fast and Private Computation of Cardinality of Set Intersection and Union," in Cryptology and Network Security, 2012, pp. 218–231.

[7] E. Stefanov, E. Shi, and D. Song, "Policy-Enhanced Private Set Intersection: Sharing Information While Enforcing Privacy Policies," in Public Key Cryptography -- PKC 2012, 2012, pp. 413–430.

[8] E. De Cristofaro, J. Kim, and G. Tsudik, "Linear-Complexity Private Set Intersection Protocols Secure in Malicious Model," in Advances in Cryptology - ASIACRYPT 2010, 2010, pp. 213–231.

[9] A. Fiat and A. Shamir, "How To Prove Yourself: Practical Solutions to Identification and Signature Problems," in Advances in Cryptology --- CRYPTO' 86, 1987, pp. 186–194.

[10] Y. Sang, H. Shen, Y. Tan, and N. Xiong, "Efficient Protocols for Privacy Preserving Matching Against Distributed Datasets," in Information and Communications Security, 2006, pp. 210–227.

[11] A. Mohr, "A survey of zero-knowledge proofs with applications to cryptography," South. Illinois Univ. Carbondale, pp. 1–12, 2007, [Online]. Available: http://austinmohr.com/Work_files/zkp.pdf.

[12] M. Bellare and A. Palacio, "GQ and Schnorr Identification Schemes: Proofs of Security against Impersonation under Active and Concurrent Attacks," in Advances in Cryptology --- CRYPTO 2002, 2002, pp. 162–177.

[13] S. Jarecki and X. Liu, "Efficient Oblivious Pseudorandom Function with Applications to Adaptive OT and Secure Computation of Set Intersection," in Theory of Cryptography, 6th Theory of Cryptography Conference, TCC 2009, San Francisco, CA, USA, March 15-17, 2009. Proceedings, 2009, vol. 5444, pp. 577–594, doi: 10.1007/978-3-642-00457-5_34.

[14] C. Meadows, "A More Efficient Cryptographic Matchmaking Protocol for Use in the Absence of a Continuously Available Third Party," Apr. 1986, p. 134, doi: 10.1109/SP.1986.10022.

[15] Q. Xie and U. Hengartner, "Privacy-Preserving Matchmaking For Mobile Social Networking Secure Against Malicious Users," 2011.

[16] N. Eagle and A. Pentland, "Social Serendipity ," Time, 2005.

[17] M. Li, N. Cao, S. Yu, and W. Lou, "Find U: Privacy-preserving personal profile matching in mobile social networks," Proc. - IEEE INFOCOM, no. 1, pp. 2435–2443, 2011, doi: 10.1109/INFCOM.2011.5935065.

[18] Z. Yang, B. Zhang, A. C. Champion, D. Li, D. Xuan, and J. Dai, "E-SmallTalker: A Distributed Mobile System for Social Networking in Physical Proximity," in 2010 IEEE 33th International Conference on Distributed Computing Systems, Jun. 2010, pp. 468–477, doi: 10.1109/ICDCS.2010.56.

[19] S. Sarpong and C. Xu, "Privacy-preserving attribute matchmaking in proximity-based mobile social networks," Int. J. Secur. its Appl., vol. 9, no. 5, pp. 217–230, 2015, doi: 10.14257/ijsia.2015.9.5.22.

[20] S. Sarpong, C. Xu, and X. Zhang, "PPAM: Privacy-preserving attributes matchmaking protocol for mobile social networks secure against malicious users," Int. J. Netw. Secur. vol. 18, no. 4, pp. 625–632, 2016.

[21] E. De Cristofaro, A. Durussel, and I. Aad, "Reclaiming privacy for smartphone applications," 2011 IEEE Int. Conf. Pervasive Comput. Commun. PerCom 2011, no. March, pp. 84–92, 2011, doi: 10.1109/PERCOM.2011.5767598.