ISSN No:-2456-2165

Spam: Secure and Privacy-Preserving Attribute-based Matchmaking

Solomon SARPONG
Department of Physical and Mathematical Sciences
University of Environment and Sustainable Development (UESD) Somanya, Ghana

Abstract:- Naturally, friendships are formed between persons with common interests. Since the Internet became ubiquitous and the proliferation of social media platforms, friendship has progressed from being entirely physical to virtual. The bane of social media platforms has been the issue of privacy and security of users' information. Most of the existing schemes where users either broadcast their information or send attributes to a central server so as to find the best match pair has got some privacy issues. In these platforms, users' sensitive personal information can easily be compromised. Also, attacks from semi-honest or malicious adversaries are difficult to prevent. In order to prevent malicious attacks, this paper proposes protocols based on privacy-preserving scalar product computation and authenticated Diffie-Hellman protocol. The use of this platform can help users find a perfect match without compromising their information.

Keywords:- *Proliferation, Ubiquitous, Privacy Preservation, Minimum Threshold, Scalar Product Computation*

I. INTRODUCTION

The study of social networks has gone on for about a century [1]. The attention of computer scientists, psychologists, security personnel and other agencies has been drawn to social media engagements since their usage became ubiquitousness [1]. It has been observed that, the security of users' information has always been an issue in online services. Hence, even users of open online services are averse to their data being used or viewed by unintended parties.

How best to protect the privacy of users' attributes in matchmaking has been a very pertinent challenge. In the quest to address this issue, these types of protocols have been proposed: (i) the use of a trusted Server to perform the matchmaking. The Server knows all the attributes of the persons, the common attributes of the matched-pair and their location. The bane of this protocol is that there are privacy concerns of users' information as the Server knows all of users' attributes. Also, a problem with the Server will cause the whole matchmaking protocol ineffective. In matchmaking social services proposed by [2] this protocol is used. (ii) the other protocol enables users to broadcast their personal information. The intersection set of their attributes is computed by each user after receiving the other's attributes. This is known as the fully distributed platform. On this platform, there is no need for a Server being contacted for every step of the matchmaking protocol. This platform being as good as it is, leaks personal private information. Also, it is not easy to prevent attacks from malicious adversary. This type of platform is used on [3], [4] [5] and (iii) there is also the hybrid protocol where a server is only needed in the initial phase for management and verification but not in the matchmaking. This protocol ensures that the matchmaking is very efficient, leaks no user private information and security of the matched-pair information is guaranteed [6], [7].

This paper proposes a matchmaking protocol based on a variant of [8]. The following modifications were made to enhance the security of the protocol (i) there is no need for a Body Sensor Network (BSN) in our proposed protocol, (ii) after a suitable pair has been found, the protocol does not end. Further communication continues hence, authenticated Diffie-Hellman key exchange is incorporated to guard against attacks.

The main contributions of this paper are: (i) An enhanced matchmaking protocol for matchmaking as only users that meet a certain minimum threshold number of common attributes qualifies to be match-paired, (ii) common attributes to both users are disclosed to the users only. As the users have the same attributes, they will not leak it, (iii) in addition, the protocol can resist semi-honest and malicious attacks.

The rest of the paper is categorized as follows: Section II outline of related works. Our proposed protocol is in Section III. The security of the protocol and conclusion are in Sections IV and V respectively.

II. RELATED WORKS

Since the inception of matchmaking, the features that need addressing has been collecting individual information, making the match pair after the necessary parameters have been fulfilled and the dissemination of results after a pair has been made. These types of protocols have been proposed to achieve this: (i). the use of third party [9] [10], [11], [12]; (ii). fully distributed system [3], [4], [13], [14], [15]; (iii). hybrid system [16], [17], [18], [19], [20], [21], [22], [23], [24].

In asymmetric matchmaking protocol as proposed by [25], when the protocol ends only one of the persons gets to know the intersection set. The person with the intersection set can discontinue the protocol. The protocol in [22] is an improvement on [25] by the removal of the capabilities of malicious attack by persons involved in the protocol. Furthermore, in [22] the intersection set is computed by both persons in the protocol and so long as a person is having the Bluetooth activated, there is the likelihood that matchmaking can be made without the consent of the person.

The matchmaking protocol proposed in [6] defined a best match-pair among many other match-seekers as the user that has the highest number of attributes in common with the initiator. Does the best match-pair have enough attributes to make them a good pair? This question is addressed in the proposed protocol in this paper.

As an improvement, this proposed protocol sets a minimum threshold number of common attributes that qualifies persons to be match-paired. In light of the inherent difficulties in the up-mentioned previous works, this proposed protocol has mechanisms to guard against these difficulties. These mechanisms include; (i) there is privacy preservation as only the common attributes are exchanged between the paired match. Hence, users' personal information are not disclosed to unrelated participants during the matchmaking, (ii) The trusted server knows nothing about the intersection set. (iii) also, high scalability is achieved as the server does not take part in the matchmaking phase. This to a large extent reduces the workload on the server. (iv) furthermore, the server does not know who has been matched to who, (v) authenticated Diffie-Hellman key exchange is implemented to enhance security and enable further communication between the match-pair.

III. THE PROTOCOL

The protocol in [8] has been modified so as to meet the security requirements of the protocols in this paper. The protocol in this paper has two phases: the initial phase and the matchmaking phase.

A. Initial Phase

The proposed protocol consists a trusted authority (TA) and a number of l registered persons $\boldsymbol{U} = (U_1, U_2, \ldots, U_l)$ looking for match-pairs. The TA is a trusted and powerful entity which is mainly responsible for management of the protocol and it is equipped with a vector matrix of attributes. Each user sends the attributes together with the identity to the TA. Furthermore, a user generates an RSA key pair and sends the public key to the TA. Hence, the TA identifies a user by the identity used, together with the attributes and the public key. The TA then identifies and authenticates a user by issuing an identifying certificate. This certificate consists of the RSA number and the user's chosen ID. This process is usually done once by a user. Also, each device being used in the matchmaking has its owners' attributes configured in it. Let $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ be a set of $U_i's$ attributes, each $a_i \in \boldsymbol{a}$ is an attribute of U_i.

B. Matchmaking

The initiator sets a minimum threshold number of attributes T_m needed to be match-paired. In a bid to for a person, say U_i to find a pair, U_i forms a binary vector of the interests/hobbies. Assume the binary vector is given by $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$. Hence, if $a_i = 1$ then U_i has that interest/hobby on the other hand, of $a_i = 0$ otherwise. Another person looking for a match-pair U_j also does a similar thing with the binary vector $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$. Likewise, if $b_i = 1$, the U_j has that interest if $b_i = 0$ then otherwise. U_i then broadcasts the binary vector when looking for a matching-pair. When U_i and U_j are within the range of each other, they will both receive the binary vector of attributes from each other. Both U_i and U_j undertake a privacy-preserving scalar product computation $\vec{\boldsymbol{a}} * \vec{\boldsymbol{b}}$ as depicted algorithm 1. After computation, if $\vec{\boldsymbol{a}} * \vec{\boldsymbol{b}} < T_m$, then they are not qualified to be match-pairs. However, if $\vec{\boldsymbol{a}} * \vec{\boldsymbol{b}} \geq T_m$, then they qualify to be a match-pair. If the minimum threshold is met, from that point onwards, U_i and U_j can start communicating. At this point, the match-pair know only the number of attributes they both have in common.

The matched-pair will continue communicating hence a secure communication protocol is needed. As a result, they undertake an authenticated Diffie-Hellman [9] protocol. This is necessary so that communication can be done securely to prevent eavesdropping. After the matched-pair have established an authenticated Diffie-Hellman key exchange, then they can exchange their common interests.

C. Algorithm

➢ *Input:*

$U_i's$ binary vector $\vec{\boldsymbol{a}} = (a_1, a_2, \ldots, a_n)$ and $U_j's$ binary vector $\vec{\boldsymbol{b}} = (b_1, b_2, \ldots, b_n)$

- *Step 1:*

  The following steps are performed by U_i

  Two large primes

  α and β are chosen; where α is of the length $|\alpha| =$ 256 bits

  and $\beta > (n+1) * \alpha^2$

Let $K = 0$ and select random positive numbers

$(c_1, c_2, \dots, c_n)$ such that $\sum_{i=0}^{n} c_i < \alpha - n$

For each element $a_i \in \vec{a}$

U_i does the following;

A random number r_i is chosen and $r_i * \beta$ is computed such that $|r_i * \beta| \approx 1024$ bits; calculate $k_i = r_i * \beta - c_i$

If $a_i = 1$ then

$C_i = \alpha + c_i + r_i * \beta, \quad K = K + k_i$

Else, if $a_i = 0$, then

$C_i = c_i + r_i * \beta, \quad K = K + k_i$

keep (β, K) secret, and send $(\alpha, C_1, C_2, \dots, C_n)$ to U_j

- *Step 2:*

U_j then executes the following

for each element $b_i \in \vec{b}$

if $b_i = 1$ then

$$D_i = \alpha * C_i = \begin{cases} \alpha^2 + c_i * \alpha + r_i * \alpha * \beta, & \text{if } a_i = 1 \\ c_i * \alpha + r_i * \alpha * \beta, & \text{if } a_i = 0 \end{cases}$$

Else if $b_i = 0$ then

$$D_i = C_i = \begin{cases} \alpha + c_i + r_i * \beta, & \text{if } a_i = 1 \\ c_i + r_i * \beta, & \text{if } a_i = 0 \end{cases}$$

$D = \sum_{i=i}^{n} D_i$ is computed and sent to U_i

- *Step 3:*

U_i does the following;

Compute $E = D + k \bmod \beta$

The computation of $\frac{E - (E \bmod \alpha^2)}{\alpha^2}$ gives the scalar product $\vec{a} * \vec{b} = \sum_{i=0}^{n} a_i * b_i$

- Step 4:

Authenticated Diffie-Hellman Protocol

U_i and U_j choose random numbers U_i^A and U_j^B respectively; U_i^A and $U_j^B \in Z_q^*$

U_i computes $g^{U_i^A} = Enc\left(g^{U_i^A} || U_i\_ID\right)$ and sends it to U_j

U_j computes and send $g^{U_j^B} || sign_{U_j}\left(g^{U_i^A} || g^{U_j^B} || U_{i\_ID}\right)$ to U_i

U_i computes and send $sign_{U_i}\left(g^{U_i^A} || g^{U_j^B} || U_j\_ID\right)$ to U_j

U_i computes $\left(g^{U_i^A}\right)^{U_j^B}$ and U_j computes $\left(g^{U_j^B}\right)^{U_i^A}$

IV. SECURITY

The proposed protocol is user-centric privacy preserving in finding the common attributes between the users. The minimum threshold T_m ensures that a match-pair is made only when they have enough attributes in common. In the calculation of $\vec{a} * \vec{b}$ only the interests/hobbies that are common to both persons are displayed for both users. As shown in the algorithm, for each $a_i \in \vec{a}$, the computation of C_i from either $\alpha + c_i + r_i\beta$ or $c_i + r_i\beta$ is indistinguishable. This will make a malicious person unable to distinguish which computation produced a particular C_i. In addition, a malicious person cannot link C_i and C_i' since $(c_i, r_i\beta)$ are random numbers. The computations of $D = \sum_{i=1}^{n} D_i$, and $\sum_{i=1}^{n} c_i + r_i\beta$ in steps 18 and 21 make the computation of $b_i \in \vec{b}$ is also privacy preserving during the scalar product computation.

Therefore, each $a_i \in \vec{a}$ is privacy-preserving during the scalar product computation. On the other hand, for each $b_i \in \vec{b}$, we have $D_i = \alpha C_i$ when $b_i = 1$ and $D_i = C_i$ when $b_i = 0$. Obviously, this operation cannot directly hide α. However, with the computation of $D = \sum_{i=1}^{n} D_i$, the unknown $\sum_{i=1}^{n} c_i + r_i\beta$ will hide the operation on each D_i. As a result, each $b_i \in \vec{b}$ is also privacy preserving during the scalar product computation. When $\vec{a} * \vec{b}$ is computed and the result is more than the threshold, U_j is a possible match pair of U_i.

Furthermore, the common attributes of both users are the only information known by the matched-pair and since they have the same interests, so there is the assurance that none of them will leak the personal information. This proposed protocol is resistant to malicious and semi-malicious attackers. In addition, the authenticated Diffie-Hellman protocol guarantees is a secured means of communication between the match-pair after a match has been made.

V. CONCLUSION

With increasing popularity of match-paring on social networks, it is imperative to develop secure and efficient protocols to enable users to effectively and securely interact. This paper has presented a matchmaking protocol that preserves users' information from malicious and honest-but-malicious users. In this protocol, only the matched-pair know the number and actual attributes they have in common.

REFERENCES

[1]. A. Narayanan and V. Shmatikov, "De-anonymizing social networks," in *Proceedings - IEEE Symposium on Security and Privacy*, 2009, pp. 173–187. doi: 10.1109/SP.2009.22.

[2]. C. Meadows, "A More Efficient Cryptographic Matchmaking Protocol for Use in the Absence of a Continuously Available Third Party," Los Alamitos, CA, USA: IEEE Computer Society, Apr. 1986, p. 134. doi: 10.1109/SP.1986.10022.

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

[4]. 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*, Los Alamitos, CA, USA: IEEE Computer Society, Jun. 2010, pp. 468–477. doi: 10.1109/ICDCS.2010.56.

[5]. Y. Y. Shieh, F. Y. Tsai, A. Anavim, M. Shieh, M. D. Wang, and C. M. C. Lin, "Mobile healthcare: The opportunities and challenges," *Int J Electron Healthc*, vol. 4, no. 2, pp. 208–219, 2008, doi: 10.1504/IJEH.2008.019793.

[6]. Y. Wang, T. T. Zhang, H. Z. Li, L. P. He, and J. Peng, "Efficient privacy preserving matchmaking for mobile social networking against malicious users," in *Proc. of the 11th IEEE Int. Conference on Trust, Security and Privacy in Computing and Communications, TrustCom-2012 - 11th IEEE Int. Conference on Ubiquitous Computing and Communications, IUCC-2012*, 2012, pp. 609–615. doi: 10.1109/TrustCom.2012.142.

[7]. S. Sarpong, C. Xu, and X. Zhang, "An Authenticated Privacy-preserving Attribute Matchmaking Protocol for Mobile Social Networks," 2015.

[8]. R. Lu, X. Lin, and X. Shen, "SPOC: A secure and privacy-preserving opportunistic computing framework for mobile-healthcare emergency," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 3, pp. 614–624, 2013, doi: 10.1109/TPDS.2012.146.

[9]. J. Kjeldskov and J. Paay, "Just-for-us: a context-aware mobile information system facilitating sociality," in *Proceedings of the 7th international conference on Human computer interaction with mobile devices \& services*, 2005, pp. 23–30.

[10]. K. A. Li, T. Y. Sohn, S. Huang, and W. G. Griswold, "Peopletones: a system for the detection and notification of buddy proximity on mobile phones," in *Proceedings of the 6th international conference on Mobile systems, applications, and services*, 2008, pp. 160–173.

[11]. N. Eagle and A. Pentland, "Social Serendipity: Mobilizing social software," *IEEE Pervasive Computing, Special Issue: The Smartphone*, pp. 28–34, 2005.

[12]. A.-K. Pietiläinen, E. Oliver, J. LeBrun, G. Varghese, and C. Diot, "Mobiclique: middleware for mobile social networking," in *Proceedings of the 2nd ACM workshop on Online social networks*, 2009, pp. 49–54.

[13]. J. Bosco, A. Kanpogninge, Q. Xia, and B. Klugah-brown, "Distributed Privacy Preservation Matchmaking protocol in Mobile Social Networks," vol. 4, no. 5, pp. 7–17, 2015.

[14]. L. Zhang, X.-Y. Li, and Y. Liu, "Message in a sealed bottle: Privacy preserving friending in social networks," in *2013 IEEE 33rd International Conference on Distributed Computing Systems*, 2013, pp. 327–336.

[15]. B. Wang, B. Li, and H. Li, "Gmatch: Secure and privacy-preserving group matching in social networks," in *GLOBECOM - IEEE Global Telecommunications Conference*, 2012, pp. 726–731. doi: 10.1109/GLOCOM.2012.6503199.

[16]. K. Arthi and M. Chandramouli Reddy, "A secure and efficient privacy-preserving attribute matchmaking protocol for mobile social networks," *International Journal of Network Security*, vol. 19, no. 3, pp. 421–429, 2017, doi: 10.6633/IJNS.201703.19(3).11.

[17]. Y. Wang, J. Hou, Y. W. Tan, and X. Nie, "A recommendation-based matchmaking scheme for multiple mobile social networks against private data leakage," in *Procedia Computer Science*, Elsevier B.V., Jan. 2013, pp. 781–788. doi: 10.1016/j.procs.2013.05.100.

[18]. S. Sarpong, C. Xu, and X. Zhang, "PPAM: Privacy-preserving attributes matchmaking protocol for mobile social networks secure against malicious users," *International Journal of Network Security*, vol. 18, no. 4, pp. 625–632, 2016.

[19]. S. Sarpong and C. X. Xu, "A secure and efficient privacy-preserving attribute matchmaking protocol in proximity-based mobile social networks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture*

Notes in Bioinformatics), vol. 8933, pp. 305–318, 2014, doi: 10.1007/978-3-319-14717-8_24.

[20]. W. Dong, V. Dave, L. Qiu, and Y. Zhang, "Secure friend discovery in mobile social networks BT - INFOCOM, 2011 Proceedings IEEE," pp. 1647–1655, 2011.

[21]. S. Sarpongp and C. Xup, "A collusion-resistant Privacy-preserving Attribute Matchmaking for Mobile Social Networks," 2015. [Online]. Available: www.ijiset.com

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

[23]. S. Sarpong and C. Xu, "Privacy-preserving attribute matchmaking in proximity-based mobile social networks," *International Journal of Security and its Applications*, vol. 9, no. 5, pp. 217–230, 2015, doi: 10.14257/ijsia.2015.9.5.22.

[24]. S. Sarpong, C. Xu, and X. Zhang, "An Authenticated Privacy-preserving Attribute Matchmaking Protocol for Mobile Social Networks," 2015.

[25]. R. Agrawal, A. Evfimievski, and R. Srikant, "Information sharing across private databases," in *Proceedings of the 2003 ACM SIGMOD international conference on Management of data*, 2003, pp. 86–97.