SocialMix: Supporting Privacy-aware Trusted Social Networking Services

Chao Li
School of Information Sciences
University of Pittsburgh
Pittsburgh, USA
Email: chl205@pitt.edu

Balaji Palanisamy
School of Information Sciences
University of Pittsburgh
Pittsburgh, USA
Email: bpalan@pitt.edu

James Joshi
School of Information Sciences
University of Pittsburgh
Pittsburgh, USA
Email: jjoshi@pitt.edu

***Abstract*—Online Social Networks (OSNs) have been one of the most successful web-based communication models. In the recent years, a new category of OSNs namely anonymous social networks are becoming popular. Unlike traditional Online Social Networks, anonymous social networks allow users to communicate without exposing their identity. This paper presents a trusted anonymous social network service that can anonymize user identities during interaction even though the communication happens with the user's own trusted friends and contacts on the social network. A fundamental requirement of such a trusted anonymous social networks is to protect the user's identity under the guarantees of anonymity. However, in existing approaches, even though the user information is anonymized, by continuously aggregating the information from the messages posted by a user, it is possible to re-identify the user with high probability. In this paper, we propose *SocialMix* that anonymizes the users of a trusted social network such that the aggregation of messages can be prevented. We make three original contributions. First, we develop the *SocialMix* model for trusted anonymous social networks so that communication privacy can be protected by k-anonymization. Second, by considering the features of OSNs, we analyze the vulnerabilities of the naive methods that might be exploited to break the privacy. We develop new techniques to improve the attack-resilience of the *SocialMix* approach. Third, we propose intelligent mix node selection methods to significantly reduce the required number of social mix nodes while still keeping high anonymization rate. Our experiments shows that *SocialMix* provides high attack resilience and keeps high anonymization rate with few mix nodes under the trusted social network model.**

***Keywords*-social network; anonymous social network; k-anonymization; social network privacy;**

I. Introduction

With the rapid development of information technology and the huge proliferation of online social interactions, we are witnessing an immense popularity of Online Social Networks (OSNs). Based on the famous 'Six degrees of separation' theory, any two persons in this world can be connected by six steps in both physical and online social network worlds. Due to the low cost and real-time features, OSNs have become a popular tool to easily make and keep interpersonal relationships.

Many Online Social Networks work under a tradeoff between user privacy and communication trust. In conventional online social networks such as *Facebook* and *LinkedIn*, communication messages are typically linked with users through friendship relationship in order to obtain higher trust among the communicating entities. Typically users have to establish the communication trust by associating each message with their identity, photos and other information. During the last few years, many anonymous social networks, like *Whisper*, *Secret*, *Cloaq* and *Rayzit* gathered a large number of users. One of the fundamental objectives of anonymous social networks is to disassociate the communication messages from the real identities of the users. In completely anonymous discussion forums such as *Whisper*, the increased privacy comes only at the cost of reduced trust among the communicating parties, which may be unacceptable for many applications that benefit from a trusted social network. Sometimes users may have a need to communicate with their trusted friends without exposing their identities. For example, Bob may suffer from a sensitive disease and may wish to communicate with his friends anonymously to seek advice. Similarly, a journal peer-review may require anonymous communication between the editors and the reviewers even though they are connected on a trusted underlying social network. Such requirements hasten the development of a new category of OSNs, namely trusted anonymous social networks which is the focus of our work. Thus, the objective of our work is to bring the best of both the worlds: namely ensuring a highly trusted communication network and at the same time, guaranteeing a high degree of user privacy.

In a trusted anonymous social network, even though the user identity may be anonymized during communication, we note that by continuously aggregating the information from the messages posted by a user, it is possible to re-identify the user with high probability. For instance, if each message can be linked with a single user whose ID is anonymized, even though each message may only leak a small amount of personal information, the adversary may successfully re-identify the user when enough messages are collected. Therefore, it is important that the techniques used to protect the communication privacy over anonymous social networks are resilient to such attacks.

In this paper, we propose *SocialMix*, a social network anonymization approach that selects a subset of users from

978-1-5090-2675-3/16 $31.00 © 2016 IEEE
DOI 10.1109/ICWS.2016.23

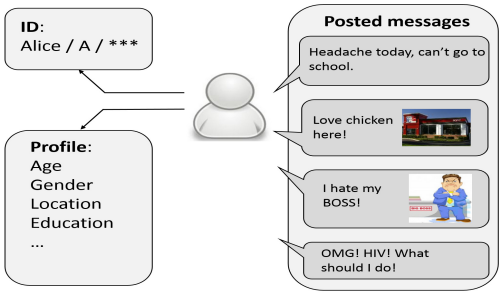


Figure 1: Information in OSNs

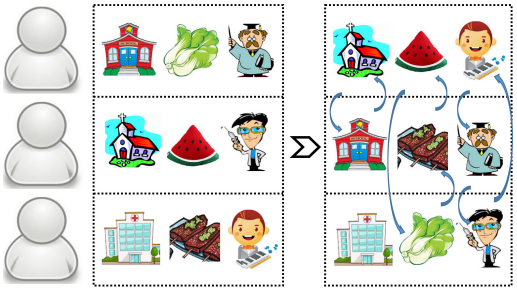


Figure 2: Message perturbation

the social network to work as mix nodes to protect the communication privacy. It supports k-anonymization-based privacy guarantees by perturbing the mapping between the identity of the user and the content of each message so that the posted message of each user cannot be aggregated for inference. We make three original contributions. First, we develop the *SocialMix* model for trusted anonymous social networks so that communication privacy can be protected by k-anonymization. Second, by considering the features of OSNs, we analyze the vulnerabilities of the naive methods that might be exploited to break the privacy. We develop new techniques to improve the attack-resilience of the *SocialMix* approach. Third, we propose intelligent mix node selection methods to significantly reduce the required number of social mix nodes while still keeping high anonymization rate. Our experiments shows that *SocialMix* provides high attack resilience and keeps high anonymization rate under the trusted social network model.

The rest of the paper is organized as follows: Section 2 analyzes the privacy over trusted anonymous social networks, presents the objectives and assumptions of the proposed approach. Section 3 presents the *SocialMix* algorithm, analyzes two types of attacks that needs to be tackled when applying the social mix approach for online social networks. We propose new techniques to improve attack resilience and present a suite of mix-node selection schemes. Section 4 discusses the experimental evaluation. We present the related work in section 5 and finally conclude in section 6.

II. Preliminary

In this section, we first analyze some existing anonymous social networking approaches to show the requirements of a privacy-aware trusted anonymous social network. Then, we present our objectives and describe the adversary knowledge and assumptions. Finally, we introduce the concept of mix nodes for supporting trusted anonymous social networking.

A. Privacy over anonymous social networks

Unlike general social networks that require users to provide real identities for building trusted relationships with each other, the goal of anonymous social networks is to break the relationship between the posted messages and the real identity of users so that users are able to communicate with their trusted contacts in an anonymous manner.

There are two categories of anonymous social networks, namely untrusted anonymous social networks and trusted anonymous social networks. The difference between them is that untrusted anonymous social networks remove (de-identify) user identities while trusted anonymous social networks anonymize user identities. In untrusted anonymous social networks, the relationships among users are weak so the feedback from other users are untrusted. As an example, *Whisper* allows nickname or allows no ID to be provided for communication. Any posted message is public to all with the nickname information, post time and location and can be commented by any other users. In such cases, people have to be watchful for the information received from other users which makes the untrusted anonymous social networks an unreliable and less trust-worthy environment for interaction and information sharing. In contrast, the relationships in trusted anonymous social networks is based on real friendships, which makes the communication trustworthy. The user only knows that the message comes from his friend, but the identity of the poster is anonymized with his other friends. One example is *Secret*, works like a masquerade. The users knows the messages come from other users within two hops, namely 'friends' and 'friends of friends', but the ID of the poster is not displayed. Another example is LEAF [12], which aims to build a platform for survivors of intimate partner violence (IPV) to communication and get help from others anonymously. The users need to communicate with other trusted users to get helpful feedback without exposing their identity.

We note that even though the user ID and profile information have been removed by most trusted anonymous social networks, the contents of posted messages may also expose the user identity. While it makes the adversary harder to infer the identity without an explicit identify information, the personal information contained in the message content may aid the adversary to improve the success rate of re-identification. For instance, in Figure 1, the first message tell us Alice has 'headache'. The second message shows that Alice is eating 'chicken', she 'love' it and the 'location' of her. For an adversary who has background knowledge about

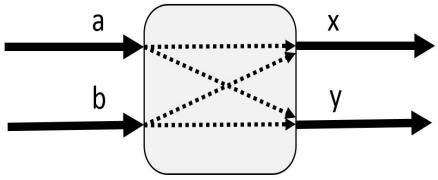


Figure 3: Mix node

Alice, these information may be sufficient to successfully infer her identity even if her ID is hidden in the communication. Once the re-identification is successful, other sensitive information of Alice may also be leaked out. Therefore, a central assumption in most anonymous social network communication is that the exposed anonymous message does not include any personally identifiable information or other information about the user that may increase the adversary's chances of re-identification.

B. Objective and assumptions

In this paper, we aim at ensuring high degree of user privacy while keeping communicating over a trusted over anonymous social network. As discussed before, the user privacy may be broken by re-identification [10] and a powerful re-identification will happen when enough messages posted by a user is collected by an adversary even though the user ID may be replaced with a pseudonym. To conquer such re-identification attacks, the key idea in our work is to shuffle the messages through message aggregation so that the relationship between the content of a message and the poster's identity can be anonymized and perturbed. An example is shown as Figure 2, which shows three messages collected for three anonymous users before and after message perturbation. Before perturbation, it is known that the first user goes to school, loves cabbage and works as a professor, a friend of him (as a potential adversary) may easily guess his identity correctly when the relationships between these messages are exposed. However, after perturbation, we note that an adversary only knows that the first user goes to church, loves watermelon and works as a musician. Since none of the information can be linked with the real identity of first user, the re-identification through aggregating message contents is likely not possible.

We also note that the adversaries in a trusted anonymous social network may be either active or passive. Passive adversaries only wait there to see the nicknames (pseudonyms) of the poster given by the OSN server. Active adversaries can be more powerful by tracking the poster physically (see whether the poster opens the website or app at the post time) or monitor the poster's network to know whether the user sends a https request to the OSN server although the message content is encrypted.

C. Definitions

In our work, a set of users is selected to work as mix-nodes to perturb the messages shared on the trusted anonymous social network. A mix node can be considered as a node assuring k-anonymization for mapping between poster identity and message content. k-anonymization is initially proposed for database publishing [16]. A released database has k-anonymization if each record in the database is indistinguishable from at least $k-1$ other records within the database. For social networks, a node-based scheme is appropriate to achieve k-anonymization.

DEFINITION *1 (*IDEAL MIX NODE*): A mix node N is said to be an ideal mix-node iff*

(1) The node N has at least k messages during perturbation.
(2) The perturbation starts when at least k messages are present and ends when the stored messages is less than k.
(3) The amount of time duration each message stays in a mix-node is completely random.

The first two requirements guarantee the k-anonymization property, while the last requirement is a restriction that can help adversaries to break the protection due to the features of OSNs. In Figure 3, we show a simple mix node with input messages a, b and output messages x, y, where a, b are old user identity and x, y are new user identity. The goal of an adversary is to infer the correct mapping between a, b and x, y. The k is set to two in this case, so during the time a, b enter the node and the time one of them leaves this node first as x or y, this node satisfies the first two requirements. The constraint on spending completely random duration of time inside injects the highest randomness so that the adversary has lowest possibility to infer the correct mapping between a, b and x, y. Let $p_{x\to a}$ denotes the probability assigned by the adversary to the mapping x to a. The last two requirements make sure that $p_{x\to a} = p_{x\to b} = p_{y\to a} = p_{y\to b} = 1/2$. The uncertainty provided by the ideal mix node can be measured by using Shannon's classic entropy [15]:

$$H(x) = -\sum_{i\in A} p_{x\to i} \times log_2(p_{x\to i})$$

$$= -k \times \frac{1}{k} \times log_2(\frac{1}{k}) = log_2(k)$$

This is the upper bound of the uncertainty can be given by a mix node, which is higher for larger parameter k.

However, when the ideal mix node model is applied to a real-world social network, there are some restrictions that provide additional information to adversaries so that the probability distribution assigned by them to the mappings can be skewed, which will minimize the anonymity (entropy) obtained. We will discuss attacks based on them and corresponding solutions in section 3.

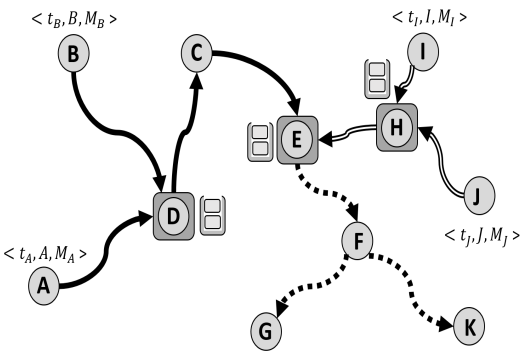


Figure 4: Mix nodes

III. SOCIALMIX

In this section, we first present the *SocialMix* approach to apply mix nodes to anonymous social networks for message perturbation. Then, we analyze the attack models and improve the attack resilience of *SocialMix*. Finally, we present a suite of mix-node selection schemes that deploy social mix nodes on a trusted social network.

A. SocialMix approach

The purpose of applying mix nodes to trusted anonymous social networks is to perturb the mapping between $< poster\ ID, post\ time >$ and the content of the posted message so that the passive adversaries cannot aggregate messages with poster ID and active adversaries cannot aggregate messages with poster ID and post time. An example is shown as Figure 4, where $< t_N, N, M_N >$ means that user N posts a message with content M_N at time t_N. Once a message is posted, it is directly shown to the one-hop friends of poster. The communication privacy should be protected from the beginning so that all the other users, who can read the message through either one-hop relationship with the poster or sharing from these one-hop friends, cannot break the anonymity by aggregating messages. Therefore, in Figure 4, nodes D and H should work as mix nodes for the four shown messages. Once user A posts a message M_A at time t_A, it should not be directly read by her one-hop friend D. Instead, D should work as a buffer until k new messages posted by her one-hop friends fill the buffer to enable a mix node. Therefore, the communication between poster A, B and all other users can be protected by k-anonymization. Similarly, node H protects the communication between poster I, J and other nodes. Both D and H work in an event-driven manner, which perturb the messages generated by their one-hop nodes. Once a user posts a message, all her one-hop nodes also work as event-driven mix nodes.

However, the protection given by event-driven mix nodes alone may not be sufficient for communication privacy over trusted anonymous social networks. First, the possible posters of a message may have close relationships. For example, in Figure 4, any user seen M_A can infer the poster to be either A or B, namely a one-hop friend of D. Second, the trust among the poster and other users reduces rapidly along the number of hops [17], so higher anonymity level is expected for users far from the poster. However, the event-driven mix nodes can only make same anonymity level for all users from hop-1 to hop-n. To solve this problem, we propose intermediate mix nodes which work as switches to make the protection diverse and multi-level. In Figure 4, user E is such an intermediate mix node which perturbs the messages shared by her one-hop friends. It also works as a buffer to guarantee that any message shared by herself may be linked with k messages shared by her one-hop friends to her. By doing this, any message read by F, G, K is protected by higher anonymity level, $k = 4$, and may come from any node of A, B, I, J. The intermediate mix node achieves higher anonymity level for further users in an exponential manner. Before meeting the first intermediate mix node, the anonymity level is k. After passing the first intermediate mix node, it becomes k^2. After $n-1$ intermediate mix nodes, the anonymity level will be k^n. In the best case when all the nodes have enough neighbors, we can get an exponentially growing anonymity level that fits the trust model of OSNs well. Also, now for F, G, K, the possible poster may be either a friend of D or a friend of H, which improves the diversity and makes it harder for the adversaries to break it.

There is a difference between event-driven mix nodes and intermediate mix nodes. For event-driven mix nodes, the mix nodes themselves should also be bewared. For example, node D should not know whether a message is generated by A or B. Since the messages have been perturbed before the node, we call a event-driven mix node as pre-mix node. For intermediate mix nodes, the node first decides whether to share the coming messages and then shuffles the shared messages. Therefore, since the perturbation happens after the node reads the messages, we call it post-mix node.

B. Attack-resilient SocialMix

As we have discussed in section 2, in cases where the distribution of probabilities assigned by the adversary to different mappings is skewed, the communication privacy may be broken, which endangers the privacy of the posters. Some features of social networks may help the adversaries to break the *SocialMix*. In particular, time-based attack and friendship-based attack can be made based on such information.

We first describe the time-based attack. In most trusted anonymous social networks, once a message is posted, it will be shown to other users in a real-time manner. If we consider the mix node as a buffer, it will perform it in a FIFO manner. Therefore, active adversaries can link messages to user identity through time information, even if the poster ID is de-identified. To handle this problem, we need to make sure each message can spend a random duration of time

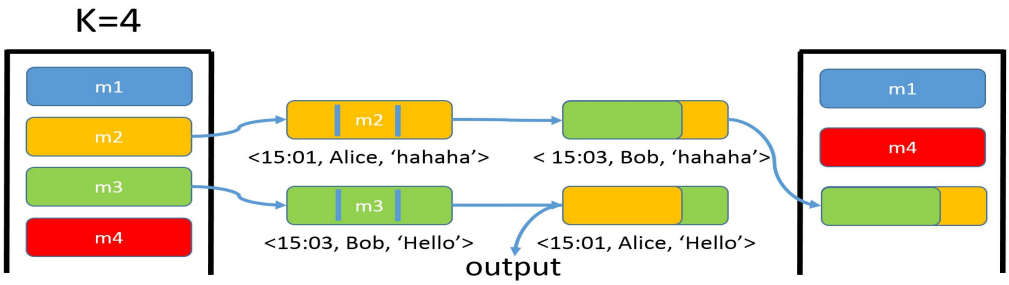


Figure 5: SocialMix

inside the mix node so that the third requirement of ideal mix nodes can be achieved.

Algorithm 1: Attack-resilient SocialMix

```
Input  : mixNodeID, k.
Output: outputTable.
for each timestamp do
    Clear the outputTable;
    for each neighbor of mixNodeID do
        if there is a shared message then
            Put shared message in holdTable;
            if holdTablesize = k then
                Randomly choose two message;
                if two segments are different then
                    Exchange their timestamp, poster ID;
                    Randomly put one to outputTable;
                    Put the other back to holdTable;
                else
                    Put the selected one to outputTable;
                end
            end
        end
    end
end
```

To improve time-based attack resilience, we need to increase the randomness of stay time of messages in buffers. In a social network, since the speed of the messages is instantaneous and the stay time is exactly the time stored in the buffer, which can be fully controlled by the mix node, it is possible to ensure that the messages stay for a random amount of time. An example is shown as Figure 5. The size of the buffer of this mix node is 4, which is also the value of parameter k. Messages $m1, m2, m3, m4$ come one by one and are stored into the buffer. At the time message $m2$ comes, the buffer is filled and the requirements of mix node is satisfied. This mix node randomly selects two messages from the message set of the buffer, which are $m2$ *and* $m3$ in this case. Each message has the structure $\langle timestamp, poster\ ID, message\ content \rangle$. In this case, message $m2$ stands for 'Alice posts the message 'hahaha' at time 15:01' and message $m3$ means 'Bob posts the message 'Hello' at time 15:03'. Therefore, each message is represented by a three-element tuple. To cut the link between $< timestamp, poster\ ID >$ and message content and to perturb the input/output, the first two elements of the tuple of these two selected messages, namely post time and poster ID, are exchanged. After that, one of the revised message is put back to the buffer and the other one is sent out as output. A special case is that we randomly choose the same one after the two rounds. Then the selected message can be directly output. The entire scheme provides three benefits. First, the system is triggered when the buffer has k messages and ends when the buffer has $k-1$ messages. Therefore, next time a new message comes, the system can be triggered again and one message can be output due to this. That means, the output rate and even output pattern is same as input, so there will be no blocking and the adversary cannot infer useful information from the difference between I/O pattern. Second, since the selected two messages may be same or different, each message may be perturbed with other messages or itself with equal chance, which provides the possibility of uniform probability distribution. Third, once a message is stored in the buffer, in each timestamp, it has equal chance to be picked out. That means, it may be the last stored message but been selected first or it may be the first coming message but been selected very late. By doing this, the duration of time inside the mix-zone for each message is completely random so that the mix node gives highest resilience towards time-based attack. The algorithm of *SocialMix* with high resilience to time-based attack for post-mix nodes is shown as Algorithm 1. The algorithm for pre-mix nodes is similar.

The friendship is a special feature of social networks, which might be exploited by adversaries. One user may have higher probability to share the messages of their best friends but have very low probability to share messages coming from somebody they don't like. Therefore, with background knowledge about the friendship of a mix node, the adversary can assign different probability to different neighbors so that the probability distribution is skewed. To prevent this, instead of using all the nodes of the social network as mix nodes, we only select a subset of them with higher resilience towards friendship-based attack. For each node in the network, we can assign the probabilities based on the friendship and calculate the entropy to measure the resilience and then select the top-n nodes with higher entropy or select

the nodes with entropy higher than a threshold to be the mix nodes. By doing this, each selected mix nodes has friendship-based attack resilience higher than a lower bound.

C. Mix Node Placement

Though the pre-mix nodes are event-driven, the post-mix nodes should be pre-determined. There should be a module on OSN server which can regularly select post-mix nodes based on the latest network topology. We propose three placement schemes, namely naive placement, top-n-based placement and centrality-based placement.

Naive placement: A naive method for mix node selection is to randomly select the nodes with higher resilience towards friendship-based attack resilience. However, we want the selected mix-nodes to play an important role in the network, otherwise many information flows cannot pass them and the anonymization rate, which is the ratio between the No. of messages passed at least one post-mix node and the No. of total messages, will be low. The random placement scheme may result in aggregations of post-mix nodes.

Top-n-based placement: Among the nodes with friendship-based attack resilience higher than the lower bound, we can further filter out the n nodes with highest entropy. The reason is that the entropy is highly related to the degree of nodes. Another advantage for this scheme is that the lower bound of friendship-based attack resilience can be further increased.

Centrality-based placement: Centrality is an important measurement for networks which can be used to measure the importance of the role of a user in a network. In this scheme, we select post-mix nodes based on their degree centrality, betweenness centrality and eigenvector centrality. Degree centrality measures the number of one-hop friends of each user. Betweenness centrality measures the proportion of shortest paths passing through each node. Eigenvector centrality assigns weights to one-hop neighbors of each nodes.

IV. Experimental Evaluation

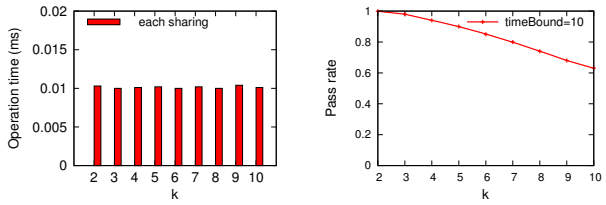
In this section, we experimentally evaluate the performance of the proposed *SocialMix* schemes. Before discussing the results, we first briefly describe the experimental setup.

A. Experimental Setup

In our experiments, we use the 'Zachary's karate club' data set which is a small social network of friendships among 34 members of a karate club at a US university in the 1970s [19]. For each node in it, there are two attributes. The first attribute is 'activity', which represents the frequency of the message generated (posted) by this node. The second attribute is 'friendship', which indicates the probability a message passing this node can be shared by his one-hop friends. The range of both activity and friendship is [1,100]. If a random number generated within the range [1,100] is smaller than the activity and friendship, new message generation and message sharing should be done respectively. The visualization and file conversion is done by 'igraph' of R language while the main simulation is done by Java language.

B. Experimental Results

In our experiments, we first evaluate the performance of *SocialMix* in terms of operation time and pass rate. Then, we measure its resilience towards both time-based attack and friendship-based attacks. First, we evaluate the proposed post-mix node placement schemes.

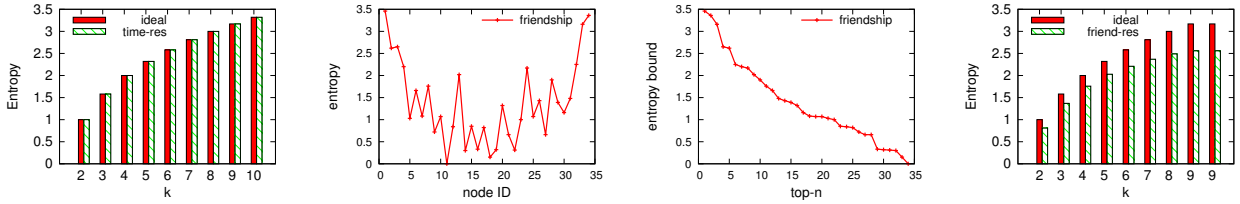


(a) Time performance (b) Delay performance

Figure 6: SocialMix time performance

The first set of experiments evaluates the performance of *SocialMix*. Figure 6(a) measures the operation time of the algorithm with varying anonymity level k. As can be seen, the operation time for each sharing process for a single mix node is stable for varying k. This makes sense because once the system becomes stable, the size of stored message is either $k-1$ or k. Then, whatever the value of k is, the *SocialMix* always randomly selects two messages from the set and exchanges the first two elements between them. The entire process is independent of the value of k. This guarantees a good algorithm scalability in terms of anonymity level. Figure 6(b) shows the delay performance of the mix-node. We do not want a message to be blocked by a mix-node, which means it is stored in the buffer for a long time and cannot be selected for output. Therefore, we set a time bound 10 and measure the probability that a message can pass this mix node within 10 timestamps with varying k. The results show that the pass rate is lower for higher k. A larger k refers to a buffer with larger size, so the probability that one message can be selected in one sharing operation is lower. Based on the requirement, we need to select an appropriate value of k.

The second set of experiments evaluates the attack resilience of *SocialMix* for time-based attack and friendship-based attack. As we have discussed, the *SocialMix* can provide complete random duration time inside the mix node for each message. Without *SocialMix*, each output can be accurately linked with one of the input, if the coming order of messages is known. Therefore, there is no uncertainty and the entropy for that is 0. Figure 7(a) shows the entropy under



(a) Time-based attack (b) Friendship-based attack resilience distribution (c) Entropy bound with varying top-n (d) Friendship based attack

Figure 7: SocialMix attack resilience

ideal situation and time-based attack. As can be seen, the entropy under time-based attack is exactly same as the ideal case, which means the adversary cannot gather any additional information through time-based attack and *SocialMix* can completely defeat time-based attack. Figure 7(b) shows the distribution of friendship-based attack resilience in the test social network, which is measured by entropy. As can be seen, the variation of entropy is very large. Some of the nodes may have high resilience with entropy larger than 3 while node 11 provides no resilience with entropy 0. If we select these low-resilient nodes as mix nodes, the perturbed mapping can be broken with very high probability, which proves the importance of our solution. Our scheme is based on top-n, which selects the n nodes with highest entropy as mix nodes. The evaluation of this top-n scheme is shown in Figure 7(c). The entropy bound is the lowest entropy provided by any selected mix node, which can be seen as the lower bound of the resilience. The two extreme top-n conditions, namely top-1 and top-34, gives 3.46 entropy bound and 0 entropy bound respectively. In practice, based on the demand, a threshold can be set to determine the value of n. Figure 7(d) shows the entropy under ideal situation and friendship-based attack. The results show that even though we have chosen the better nodes with higher resilience, the entropy by performing friendship-based attack is lower, which means that there are still some information leaked out. Theoretically, the entropy under friendship-based attack will be closer to the ideal case if the friendship between each pair of nodes is similar.

The third set of experiments evaluates the pre-mix node placement performance under naive-based, top-n-based and centrality-based schemes. Figure 8 shows the anonymization rate with varying selected number of post-mix nodes. The anonymization rate stands for the proportion of messages that passed at least one post-mix node. As can be seen, the anonymization rate of naive scheme which randomly selects post-mix nodes grows slowly with increasing number of selected post-mix nodes. However, even for random selection case, a subset of 15 nodes among the 34 nodes can already guarantees a very high anonymization rate, which proves that we do not need all the nodes to be mix nodes so that the cost can be significantly reduced. The other four schemes based on either top-n or centrality have better performance. Even two post-mix nodes selected by

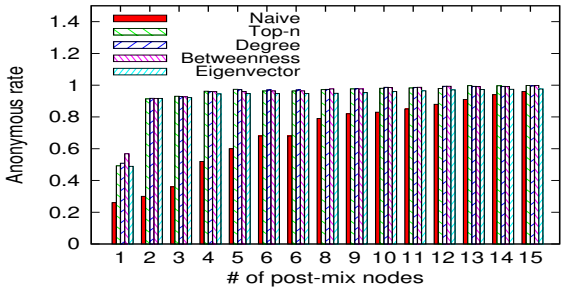


Figure 8: Anonymous rate evaluation

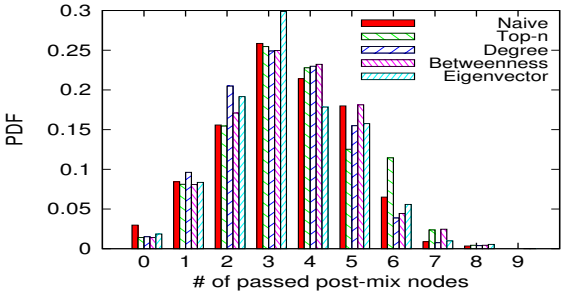


Figure 9: Anonymous level evaluation

them can make anonymization rate higher than 90%, which demonstrates the effectiveness of these enhanced placement schemes. In Figure 9, we measure the PDF of number of passed post-mix nodes by a message under different post-mix placement schemes. If each post-mix node can guarantee anonymity level k, by passing n nodes, the anonymity level for users with low trust who are far away from the poster will be k^n to guarantee stronger protection. Therefore, we do not want the distribution to be skewed to either left or right. A left skewness will under-protect the privacy while a right skewness will over-protect the privacy. As can be seen, the PDF for all the schemes roughly follows normal distribution, which provides an appropriate protection of the privacy.

V. Related Work

There have been many works on user privacy over OSNs. Most works focus on applying access control models to make sure the personal information can only be seen by the users with authorization. Basically, the works on this can be divided into the access control towards the OSN service

providers and third parties (e.g. NOYB [9], Persona [1]) and access control towards other users (e.g. Lockr [11], VisualSec [8]). The anonymity, as a way to protect user privacy, also drew attention of the researchers and some works on the effects of anonymity over OSNs was proposed [4], [14].

During the last few years, a new social communication model, anonymous social networks, emerged due to the need for privacy-aware interactions in social networks. Some examples are *Whisper*, *Secret*, *Cloaq* and *Rayzit*. Few works have been done on anonymous social networks. A recent work [18] collected data from *Whisper*, analyzed the effects of anonymity and lack of links and evaluated a location-based attack. However, the re-identification through message aggregation was not taken into account. To the best our knowledge, the work presented in this paper is the first work which applies k-anonymization-based approach for anonymous social networks to improve the resilience to the attacks based on message aggregation and timing information, thus achieving both trusted communication and high privacy.

The idea of social mix nodes proposed in this work is inspired from the concept of mix-zones, which applies k-anonymization to protect location privacy. Mix-zones refer to regions in space where a set of users enter, change pseudonyms and exit in a manner that the mapping between their old and new pseudonyms are anonymized [2][3][5][6][7][13]. As intersection points of physical or virtual flows represented by pseudonyms like IDs, they shuffle the pseudonyms of users so that the input/output mapping cannot be inferred. The notion of social mix nodes studied in this work applies a similar principle to information flows on a social network so that attackers who track the flows based on the pseudonyms cannot get the correct source information, thus protecting the anonymity of the message sender.

VI. Conclusion

This paper proposes *SocialMix*, an anonymous communication mechanism to support privacy-aware trusted social networking services. *SocialMix* operates by perturbing the mapping between message contents and poster identities under the guarantees of k-anonymization. Two attack models, namely time-based attack and friendship-based attack are analyzed and new mix node construction techniques were designed to improve the attack resilience of the proposed approach. We propose a suite of mix node construction and placement schemes that enhance the attack resilience and anonymization effectiveness of the *SocialMix* approach. Our experimental evaluation shows that *SocialMix* provides high attack resilience for trusted communication over social networks with high anonymization rate.

Acknowledgment

This work was performed under a partial support by the National Science Foundation under the grant DGE-1438809.

References

[1] Baden R, Bender A, Spring N, *et al*. Persona: an online social network with user-defined privacy. ACM SIGCOMM Computer Communication Review. ACM, 39(4): 135-146, 2009.

[2] Beresford A R, Stajano F. Location privacy in pervasive computing. *IEEE Pervasive computing*, (1): 46-55, 2003.

[3] Beresford A R, Stajano F. Mix zones: User privacy in location-aware services, 2004.

[4] Bernstein M S, Monroy-Hernndez A, Harry D, *et al*. 4chan and/b: An Analysis of Anonymity and Ephemerality in a Large Online Community. ICWSM, 50-57, 2011.

[5] Buttyan L, Holczer T, Vajda I. On the effectiveness of changing pseudonyms to provide location privacy in VANETs. *Security and Privacy in Ad-hoc and Sensor Networks*, 129-141, 2007.

[6] Freudiger J, Raya M, Flegyhazi M, *et al*. Mix-zones for location privacy in vehicular networks, 2007.

[7] Freudiger J, Shokri R, Hubaux J P. On the optimal placement of mix zones. *Privacy enhancing technologies*, 216-234, 2009.

[8] Ge M, Lam K, Wang X, *et al*. VisualSec: A secure message delivery scheme for online social networks based on profile images. Global Telecommunications Conference, 39(4): 1-6, 2009.

[9] Guha S, Tang K, Francis P. NOYB: privacy in online social networks. Proceedings of the first workshop on Online social networks. ACM, 49-54, 2008.

[10] Hay M, Miklau G, Jensen D, *et al*. Resisting structural re-identification in anonymized social networks. Proceedings of the VLDB Endowment, 1(1): 102-114, 2008.

[11] Tootoonchian A, Saroiu S, Ganjali Y, *et al*. Lockr: better privacy for social networks. Proceedings of the 5th international conference on Emerging networking experiments and technologies. ACM, 169-180, 2009.

[12] Palanisamy B, Sensenig S, Joshi J, *et al*. LEAF:A Privacy-conscious Social Network-based Intervention Tool for IPV Survivors. *Information Reuse and Integration (IRI), 2014 IEEE 15th International Conference on*, 138-146, 2014.

[13] Palanisamy B and Liu L. Attack-resilient mix-zones over road networks: architecture and algorithms. *IEEE Transactions on Mobile Computing (TMC 2015)*, 14(3), 495-508.

[14] Schoenebeck S Y. The Secret Life of Online Moms: Anonymity and Disinhibition on YouBeMom. com. ICWSM, 2013.

[15] Shannon C E. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1): 3-55, 2001.

[16] L. Sweeney. k-anonymity: A model for protecting privacy. in *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems (2002)*, 557-570.

[17] Trifunovic S, Legendre F and Anastasiades C. Social trust in opportunistic networks. INFOCOM IEEE Conference on Computer Communications Workshops, 1-6, 2010.

[18] Wang G, Wang B, Wang T, *et al*. Whispers in the dark: analysis of an anonymous social network. Proceedings of the 2014 Conference on Internet Measurement Conference, 137-150, 2014.

[19] Zachary C C. An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 33: 452-473, 1977.