Friend in the Middle (FiM): Tackling de-anonymization in social networks

Conference Paper · March 2013

DOI: 10.1109/PerComW.2013.6529495

Filipe Beato
KU Leuven

Mauro Conti
University of Padova

Friend in the Middle (FiM): Tackling De-Anonymization in Social Networks

Filipe Beato
ESAT/COSIC - KU Leuven and iMinds
Leuven, Belgium
filipe.beato@esat.kuleuven.be

Mauro Conti
University of Padua
Padua, Italy
conti@math.unipd.it

Bart Preneel
ESAT/COSIC - KU Leuven and iMinds
Leuven, Belgium
bart.preneel@esat.kuleuven.be

***Abstract*—With the large growth of Online Social Networks (OSNs), several privacy threats have been highlighted, as well as solutions to mitigate them. Most solutions focus on restricting the visibility of users information. However, OSNs also represent a threat for contextual information, such as the OSN structure and how users communicate among each other. Recently proposed de-anonymization techniques proved to be effective in re-identifying users in anonymized social network. In this paper, we present Friend in the Middle (FiM): a novel approach to make OSNs more resilient against de-anonymization techniques. Additionally we evaluate and demonstrate throughout experimental results the feasibility and effectiveness of our proposal.**

***Index Terms*—Online Social Networks, Security, Privacy.**

I. Introduction

Online social networks (OSNs) have become a large success. With the increase of users' privacy awareness, OSNs also start to generate privacy concerns, due to leakages of private information, unwanted viewers, etc. Some OSNs offer privacy controls that allow users to hide content or relationships from others in the network. However, the OSNs manager still have access to all published information. Hence, OSNs might share this information with target advertising partners and other third parties services. Several solutions were proposed to protect user's data and personal information using privacy preserving mechanisms, such as restrict the visibility of user's information: non authorized users or the OSN manager access only fake or encrypted information.

Most privacy solutions focus on content privacy (e.g. confidentiality of a message exhanced) rather than contextual privacy (e.g. privacy about information as the social network structure, or the fact the some users communicate among them). Even when the user identity is anonymized (VPSN [4]), recently proposed de-anonymization techniques has been shown to be effective to re-identify users. Thus, even when the user hides her real identity, contextual information as OSN graph can be leaked. This problem has clearly been shown to have a significant impact. In [18], [19], [20] authors present models and manage to de-anonimize users that are registered on different networks also based on its connections. In addition, in [13] authors presented a way to classify the sexual orientation of a users based on user's connections.

Our contributions. In this paper, we focus on contextual privacy, and we introduce a novel approach, named *Friend in the Middle* (FiM) to make it more difficult for an adversary to re-identify an anonymized user in a OSN. The main idea is that a OSN profile acting as "Friend in the Middle" helps other two users to be "connected", while them not enjoying a direct connection from the OSN point of view. Furthermore, from the topological point of view, we consider the possibilities of connecting two profiles via only one or more FiMs. To assess and evaluate our proposal we perform different extensions to a real OSN dataset by applying varying FiM approaches to produce different datasets. Later we run experimental evaluations on those datasets where the outcome supports our claims on the feasibility of our approach.

Organization. In Section II, we discuss the state of the art for the problem we address in this paper. In Section III, we introduce the model considered for our study. Subsequently, we introduce the Friend in the Middle approach in Section IV. In Section V we provide a thorough experimental evaluation. Finally, we conclude our work in Section VI.

II. Related Work

While OSNs become widely used by a relevant fraction of the world population, their security and privacy issues also come to the light. Avoiding the disclosure of information to non authorized users or to the OSN manager without the user willingness remains one of the important issues. Some researchers highlighted security issues due to vulnerabilities of the implementation of OSN. In [17] the authors described how an adversary can have access to information of a victim having a profile in Facebook, without the victim willingness to share information with the adversary. Others addressed attacks more inherently related to the nature of OSNs. In [14] authors addressed the problem of an adversary creating a profile with "cloned" information from an existing profile on the OSN. A more general problem of an adversary impersonating a victim has also been recently considered [5].

In general, solutions which aim to tackle the non authorized disclosure of information focus on: i) proposing new "privacy-aware" social network architectures, or ii) extending

Filipe Beato is supported by the FCT Grant SFRH/BD/70311/2010. Mauro Conti is supported by a Marie Curie Fellowship funded by the European Commission for the PRISM-CODE project (Privacy and Security for Mobile Cooperative Devices) under the agreement n. PCIG11-GA-2012-321980.

978-1-4673-5077-8/13/$31.00 ©2013 IEEE

the current OSNs with mechanism to restrict access to information. An example of the first type is Safebook [7], a peer to peer OSN that applies the concept of *matrioshka* to separate the data published among friends of the user, also referred as connections in this work. Solutions like FaceCloak [16], Scramble! [1], VPSN [4], and Hummingbird [6], extend current OSNs and allow users to protect confidentiality and integrity of their published data. We observe that the threat model considered in these works focuses on hiding the content to non authorized users, while not much attention has been paid to the contextual privacy: e.g. the fact the two users are connected, or that they are exchanging messages. The motivation for this work was that if a user is anonymized using mechanisms like [4], the user should not care whether his social network graph is known or they are known the peers with whom the user communicates.

Unfortunately, this is not the case! In fact, de-anonymization techniques [18], [19], [20] for social network have been also proposed. Hence, users anonymized with solutions as VPSN [4] might be easily re-identified. In particular, the work in [19] presents a general way to de-anonymize social networks on a large scale, based on auxiliary data that can be publicly crawled. The authors manage to map the nodes on an anonynimized social network (Facebook) to a non-anonymized social network (LinkedIn)—the latter one being considered to be auxiliary data to which the adversary has access. The de-anonymization attack presented is done in two stages. First, on the input of the two network graphs, where one is anonymized and the other represents a non-anonymized subset. Then, the de-anonymization algorithm identifies k nodes that are represented in both sets. Second, the algorithm propagates and assign scores based on the weight and importance of the node in the network. A recent work [10], also related to this technique, proposes a framework for the detection of multiple identities on social networks based on machine learning techniques that identify common patterns used by the users.

The aim of this paper is to propose a mechanism to make anonymized OSN stronger against de-anonymization technique. At a high level, the idea can be seen as a graph anonymization technique. Other techniques of this type has been already proposed for generic graphs. In particular, some of the solutions for graph anonymization are based on $k - anonymity$ techniques [12], [22] that aim at having for different nodes a similar number of links. Other approaches like [21] address the problem by removing and adding nodes to the graph. Those solution techniques require a topology change of the OSN graph. In the other hand, our solution can be applied to any existing OSN and does not rely on $k - anonymity$. The work in [11] discusses an anonymization technique based on identity separation as proposed in [3], similar to Google+ circles.

III. Model

In the remaining part of the paper, we consider an online social network $\mathcal{S}$, to be described by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ whose vertices represent the users $u \in \mathcal{S}$ and the edges the undirected connections between users. Each u establishes a set of relationships $\mathcal{R}_u \in \mathcal{V}$ that contains all users to which u is connected with. Formally, $(u, v) \in \mathcal{E}$, or represents a valid connection, if and only if $v \in \mathcal{R}_u$.

Further, we present the adversarial capabilities, and then we overview the de-anonymization attack [19] based on comparing OSN graph structures.

A. Threat Model

We consider an adversary $\mathcal{A}$ that is interested to listen and profit from users communications, aiming to collect information about a user u, including the list of people u is connected to. In addition, other malicious parties, such as other users and third party applications, can collect communication information from the users and collude with the adversary. We assume that while $\mathcal{A}$ may profit in eavesdropping, he will not change the communication data as this would result in being detected. Also, we assume that FiM nodes will obey to the protocol, while not tampering with the communication. Thus, we model $\mathcal{A}$ to be *honest but curious*.

We assume that $\mathcal{A}$, aiming to de-anonymize a OSN $\mathcal{S}$, is able to collude or collect auxiliary information from a different OSN $\mathcal{S}'$ that is not anonymized. In addition, it is assumed that a subset of $\mathcal{S}'$ exists in $\mathcal{S}$. We note that $\mathcal{S}'$ could be the case of LinkedIn, where users share part of their profile for professional reasons. To de-anonymize the user, the adversary can use de-anonymization techniques, as in [19], based on the similarity of the graph components in the two colluding OSN.

It is assumed that the profile of users in $\mathcal{S}$ has been anonymized with basic methods: the profile itself does not contain any real information about the user. Also, all communication is performed using encryption. We note that this would be, for example, the outcome of the application of OSN privacy solutions as Scramble [1] or VPSN [4].

In this paper, we do not explicitly consider attacks as in [15], where authors assume an adversary can break into user's accounts. Even though, if FiM is properly implemented, i.e. the information a FiM node needs to act as such are stored and used only in trusted parties as application servers or user's devices, even running this attack the adversary would not get an easier task to run the de-anonymization.

B. De-Anonymization Attack

Based on our threat model, and consistently with the model presented in [19], the adversary $\mathcal{A}$ can break user's privacy if he can learn whether a certain u is in the network $\mathcal{S}$, or if he can match an anonimized node $u \in \mathcal{S}$ to an non-anonimized one $u' \in \mathcal{S}'$. To perform the attack, $\mathcal{A}$ is assumed to have access to a non-anonimyzed network $\mathcal{S}'$ containing a subset of nodes correlated with the anonymized network $\mathcal{S}$. To proceed with the attack, $\mathcal{A}$ inputs an initial mapping seed that can be done manually, and the non-anonymized dataset $\mathcal{S}'$. The initial mapping consists on the adversary knowledge of some nodes match, e.g. $\mathcal{A}$ is positively aware that user $u \in \mathcal{S}$ maps to $u' \in \mathcal{S}'$. According to the node similarity, based on homologous attributes such as connections, the algorithm outputs the result mapping matching user $u \in \mathcal{S}$ to user $u' \in \mathcal{S}'$.

IV. Friend in the Middle

In this section, we present our FiM approach and discuss its benefits in terms of privacy. Let u and v be arbitrary users of the OSN that aspire to preserve their privacy. In particular, u and v want to hide the fact that they have a "friendship" relation, as well as the fact that they might be exchanging messages. In general, a FiM node is a node on the OSN $\mathcal{S}$ that acts as a "mediator" between two users that want to protect their privacy, as u and v do. The FiM node is required to be consistently online. We use the following definitions to present our proposal.

*Definition 1 (Friend in the middle (*FiM*)):* A node $x \in \mathcal{S}$ is said to be a FiM node for the nodes $(u, v) \in \mathcal{S}$ (s.t. $x \neq u \wedge x \neq v$) if x is connected with both u and v, and it acts as an intermediary, by forwarding communication between u and v.

*Definition 2 (*FiM*-connected):* Two nodes $(u, v) \in \mathcal{S}$ s.t. $x \neq u \wedge x \neq v$, are said to be FiM-connected if they are connected via a FiM. Thus, the connection (u, v) is not present, while there are the links (u, x) and (x, v).

In practice, a FiM can be seen as a node that performs a service similar to a proxy. Applying a FiM node x to the edge (u, v) representing the connection between users u and v, both users will use x to forward all the communication between each other. An example of a network using FiM is depicted in Figure 1, where user u and user v are connected using FiM with id x.

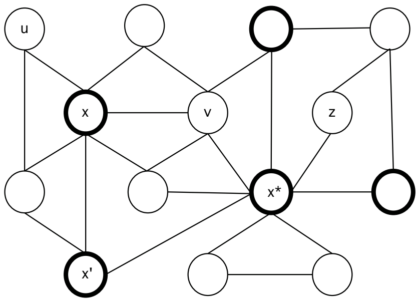


Fig. 1: An example of a FiM network, where non-FiM users (thin circles) are connected among them via FiM nodes (tick circles).

In the following, we first describe the different approaches for using FiM. Then, we discuss some challenges for a possible practical implementation of the FiM approach in real OSNs.

A. *Types of* FiM

Now we discuss the different FiMs approaches on our model and their benefits to privacy. We start to present the single FiM and then next we expound a multiple hop FiM approach.

1) Single FiM: The first approach is to use a single FiM x to connect two users (u, v). Thus, x acts as a simple mixer by only forwarding encrypted data from u to v, in a way similar to the one described in [2]. In this way, the OSN provider is kept oblivious of the source and destination nodes link (u, v), provided that the information about the role of x and its kept outside the domain of the OSN—in practice this might be the case where those additional information are kept in a (trusted) third party application server, or locally in the computer or a portable device of the user.

2) Multi FiM*s:* A more privacy friendly approach where x is not required to be trusted is the multi FiM. The users (u, v) can be separated by n FiM hops instead of a single one, described in Definition 3.

Definition 3 (n-hop FiM*):* A multiple or n-hop FiM is a network composed by a list of FiM nodes $X = \{x_0, ..., x_n\} \in \mathcal{S}$ s.t. a connection between two users in the network (u, v) is made through n FiM nodes.

For example, if $n = 3$ then (u, v) connection is as follows: $u \leftrightarrow x \leftrightarrow x' \leftrightarrow x^* \leftrightarrow v$, where $\leftrightarrow$ indicates two nodes being connected in the OSN, as represented in Figure 1. This way of applying FiM presents a similar functionality as Tor [8]. In fact, this allows a higher degree of anonymity for u and v communication with respect to the OSN. The entry FiM just learns that the communication started in u is forwarded to next node that is also a FiM, the middle just maps the entry and exit FiMs and the exit node just learns that a certain communication ended in node v.

B. *Implementation challenges*

To be effective, FiM nodes need to work outside of the domain of the adversary. For example, considering Facebook as the OSN, the FiM behaviour can be implemented for a node as a Facebook App, where information about the nodes served by the FiM are stored on an external application server outside the control of Facebook. Similarly, the computation that a FiM needs to do must be done on such external server. We notice that instead of an external server, a possible implementation might also be done in the user's device that access the OSN. For instance, this might be done via a browser extension running on the user device.

Another practical issue is how two nodes (u, v) become FiM-connected. An approach is that (u, v) both agree on the list of FiM nodes involved in the path. In practice, such agreement can be done by both users when they decide (e.g. offline) to have a FiM-connection on the OSN. While agreeing on the full list is the only option for the single FiM scenario, i.e. both (u, v) knowing the id and role of the FiM nodes, another approach is possible in the $Multi$ FiM scenario. In particular, user u can choose the full path and disclose to v only the id of the last FiM node. In both cases the intermediate FiM nodes need to be notified and instructed about their role. Finally, we note that v might set up and use to communicate with u a FiM-path which is different from the one established by u to communicate with v. While we leave this behaviour out of the evaluation reported in Section V, we expect this to increase users' link privacy, as the general graph becomes more interconnected.

V. Evaluation

In this section, we present the results of our evaluation via a thorough set of experiments on a real dataset. We start to present the dataset we considered. Then we discuss the experiments, followed by the analysis on the direct effect of FiM-connected nodes have on re-identified nodes. Finally we draw some observations on the obtained results.

A. Considered Dataset

We performed our evaluation analysis based on a social network datasets made available by SNAP[1]. The initial dataset $\mathcal{S}_{init}$ is composed by Slashdot connections from 2008, while the second dataset $\mathcal{S}_{aux}$ is from 2009. We outline the parameters of both datasets in Table I, where the node identification and its edges are exposed. We considered only the symmetric connections, thus the degree of a node is calculated by the sum of the its connections. Later, we populate the dataset $\mathcal{S}_{init}$ with different variances of FiMs to produce test sets to execute attacks using the de-anonymization algorithm [19] and evaluate it using the dataset $\mathcal{S}_{aux}$ as reference input for the auxiliary information.

TABLE I: Datasets Characteristics

Network	Slashdot (2008)	Slashdot (2009)
Nodes	77359	82166
Edges	905468	948464
Av. Degree	90	88

B. Experiments

We investigated the impact that the number of nodes using FiM has on the ability of the de-anonymization algorithm to achieve large-scale re-identification. To measure the accuracy of our method, we started by constructing several different graphs $\mathcal{S}_{\mathsf{FiM}}$ from the initial dataset $\mathcal{S}_{init}$ by populating with different FiM approaches. This affects directly the topology of the graph, by increasing the average degree, specially on the nodes acting as FiM. After, we run several tests by applying the de-anonymization attack from [19] to evaluate our methods. Such algorithm is required an initial mapping seed and an auxiliary dataset $\mathcal{S}' = \mathcal{S}_{aux}$, where certain users are also present. We used the nodes with highest degree to compose the initial seed for the mapping, and a subset of 10k nodes of the de-anonymyzed version of the initial dataset $\mathcal{S}'$ as the auxiliary data, over $\mathcal{S} = \mathcal{S}_{\mathsf{FiM}}$. We believe, that using the initial version of the network as auxiliary data represents a stronger security assumption as gives the exact nodes of the network.

In order to demonstrate and evaluate our model we divided our evaluation tests in test sets, were we used a total number of 50 nodes in the network to act as a FiM. Later we extended the total number of FiMs to 150. Prior to apply our model, we have also run the algorithm on the dataset $\mathcal{S}_{init}$ with no FiM. The outcome of this experiment was a re-indetification of 98% of the total network. Our test sets are described below.

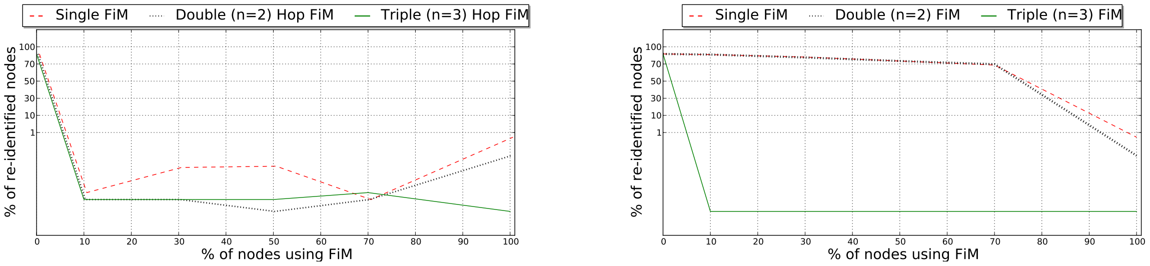
1) FiM *applied to all nodes:* For our first experiment we populated the dataset with a FiM for all connections (edges). As it is revealed from the data just a very few nodes were re-identified by the de-anonymization algorithm. During our tests the algorithm identified a total of 0.76% of the nodes. However, 0.75% represents false positives match.

2) FiM *applied per node connections:* Our second experiment consisted of applying FiM to a percentage of connections per node on the initial dataset $\mathcal{S}_{init}$. Hence, just a percentage of each nodes connections on $\mathcal{S}_{\mathsf{FiM}}$ have a FiM acting as intermediary. The results are illustrated in Figure 2.a. It is visible from the figure that the results present a decrease trend, once the number of FiMs per connection is applied. As the outcome presents very low values we used a logarithmic scale to show the variations when applying different FiM approaches. Figure 3.a displays the amount of false positives matched nodes compared to the correctly matched. When applied a single FiM the de-anonymization algorithm matches a maximum of 0.14% of nodes, however, only 0.01% represents a correct mapping. When the hops increase, it also increases the distance from u to v. This has negative influence to the matching algorithm. For $n = 2$ the de-anonymization algorithm outcome performs an average of 0.02% matches where half represent false positives. As so, when $n = 3$ the algorithm can correctly re-identify only 0.01% of the nodes, which represents a negligible value. Thus, our results show that when populating the graph with extra nodes and edges to perform a real connection (u, v) disturbs the de-anonymization classifier.

3) FiM *applied per network:* Subsequently, we created a network $\mathcal{S}_{\mathsf{FiM}}$ where we chose a percentage of the total connections of the initial dataset $\mathcal{S}_{init}$ to apply different FiM approaches. This changes the graph topology to the extend of some percentage of connections. An overview of the outcome of our tests is shown in Figure 2.b. The results obtained from applying a single FiM present a different behaviour from the approaches previously described. However, the trend presents a decreasing reaction when increasing the number of edges using FiM. Figure 3.b illustrates the percentage of the total matched nodes, and the ration of positive and false positive matches. By increasing the number of hops it increases anonymity towards the de-anonymization algorithm. Still, this only occurs for high percentage of FiMsin the network or when $n > 2$. Thus, when applied $n = 3$ hop FiM the de-anonymization classifier performance decreases compared to the previous case, re-identitfying only 0.01% of the nodes.

4) Increasing the number of FiM*:* We measured the impact on modifying the number of nodes acting as FiM used in the network. To distinguish if this presents an important parameter, we extended $\mathcal{S}_{init}$ to use a total of 150 FiMs, instead of only 50 FiMs as before. Then, we repeated the experiment from Section V-B2. Figure 4 illustrates the growth of nodes acting as FiM on the network improves link anonymity.

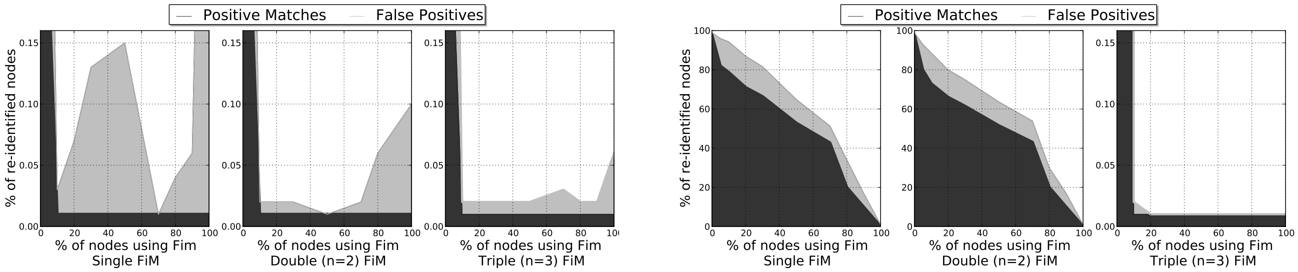
[1] Standford Network Analysis Project: http://snap.stanford.edu/



(a)

(b)

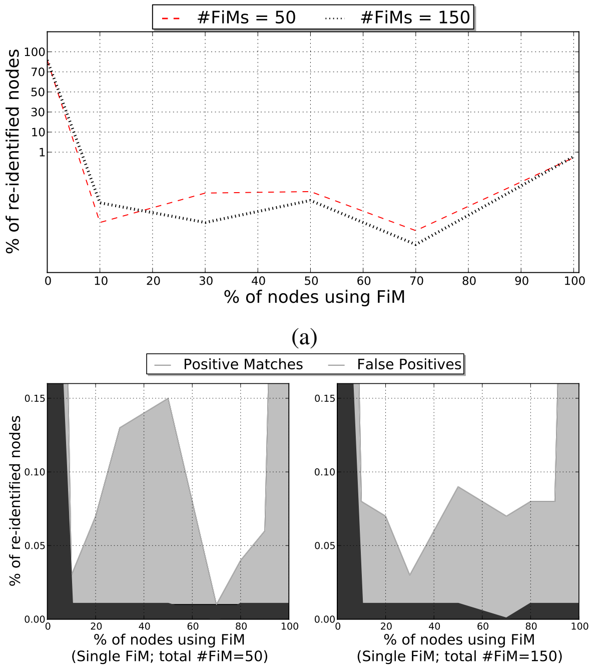
Fig. 2: Re-identified nodes varying % of: (a) FiM per node; (b) FiM per network connections.



(a)

(b)

Fig. 3: Positive and False Positive re-identified nodes varying % of (a) FiM per node; (b) FiM per network connections.



(a)

(b)

Fig. 4: Differences when varying the total number of FiMs on the network with respect to: (a) % of re-identified nodes; (b) Positive and False Positives re-identified nodes when applied single FiM per node

C. *Re-identified nodes with* FiM

Consequently, we analyzed the ratio of the correctly re-identified nodes that have FiM present on their connections. For the experience where FiM nodes were applied to a percentage of each node's connections (Section V-B2) we noticed that from the correct matched only one node was using FiMs. However, this node symbolizes the node with an extreme higher degree compared with the average of the network. This, represents an unique property which is directly captured by the de-anonymization algorithm. Regarding the experiment discussed on Section V-B3, where just a percentage of the full network connections are FiM-connected, we notice that none of the nodes correctly re-identified were using FiM for the single FiM and $n = 3$ hop FiM approach. However, for a $n = 2$ hop FiM the 10% of correctly re-identified nodes were FiM-connected. Figure 5 outlines that variation when applied FiM to $\{10, 30, 50, 70\%\}$ connections on the network.

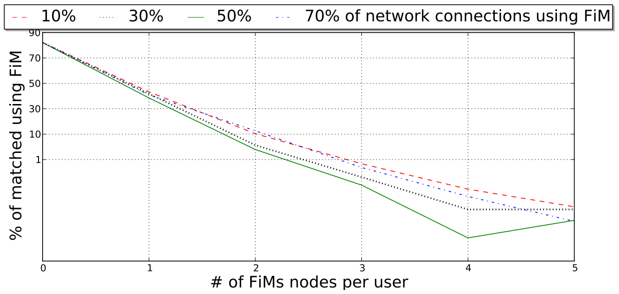


Fig. 5: Percentage of correctly re-identified nodes containing a number of FiM-connected nodes.

D. *Result Analysis*

The outcome of our experiences support our initial predictions, using FiM nodes link privacy is improved. The result of the tests demonstrate a more efficient node anonymization when applied a FiM per connection. In addition, it is revealed that by using nodes from the graph to act as FiM nodes (and being those FiM nodes a percentage of the node's connections), the anonymization presents similar effect. The algorithm outputs false positive matches which represent the map of nodes with similar characteristics. Still, a node can use the same FiM to connect to several others nodes and decrease or keep its degree value constant. It is also possible to use FiM with a percentage of its connections.

Regarding the correctly matched nodes, those are generally not FiM-connected as described in Section V-C and illustrated by Figure 5. In addition, some of those nodes present extreme unique characteristics, like a higher degree value than the average from the network. This finding reveals that the node degree value represents an important attribute for de-anonymization techniques, and when it is much higher than average it becomes an unique property.

VI. Conclusion and Future work

In this work, we considered the problem of re-identification of nodes in anonymized OSNs graphs. We addressed this problem by proposing a new approach to make anonymized OSN graphs significantly more resilient against the general de-anonymization model presented in [19]. Our approach uses FiM nodes to interconnect nodes in the OSN graph, avoiding direct connections. We performed a thorough set of experiments on several datasets produced from extensions to a real OSN graph with different FiM approaches. The evaluations confirm the viability of our solution. While without our solution the de-anonymization attack [19] was able to re-identify 98% of the nodes, with our solution the numbers of correctly re-identified nodes decreases to 0.01%—considering the simplest version of the proposal (i.e. single FiM) and with only 10% of the links having FiM nodes.

We note that while we aim to protect connections, our model can be extended to address communication privacy. At the moment it does not provide protection against traffic analysis, meaning that the provider could infer who is communicating. Protection against this kind of attacks is left as a subject of future work, by giving more capabilities to each FiM. Such capabilities include introducing timing when forwarding the message, dummy traffic or message size constant. This can be seen that a FiM would have similar behaviour as tools like Tor [8].

As future work, we plan to validate our results on different social networks (such as Facebook and LinkedIn). Also, we want to implement our FiM approach through a Facebook application, to allow nodes to act as FiM and a third party server to allow those nodes to be online. Also, we want to assess the complexity for the OSN manager to identify nodes acting as FiM nodes, along with an evaluation of the practicality of using short lived FiMs (similar to the short live proxies in [9]).

References

[1] Filipe Beato, Markulf Kohlweiss, and Karel Wouters. Scramble! your social network data. In Simone Fischer-Hübner and Nicholas Hopper, editors, *PETS*, volume 6794 of *Lecture Notes in Computer Science*, pages 211–225. Springer, 2011.

[2] David Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2), February 1981.

[3] Sebastian Claubeta, Dogan Kesdogan, and Tobias Kölsch. Privacy enhancing identity management: protection against re-identification and profiling. In Vijay Atluri, Pierangela Samarati, and Atsuhiro Goto, editors, *Digital Identity Management*, pages 84–93. ACM, 2005.

[4] Mauro Conti, Arbnor Hasani, and Bruno Crispo. Virtual private social networks. In *Proceedings of the first ACM conference on Data and application security and privacy*, CODASPY '11, pages 39–50, New York, NY, USA, 2011. ACM.

[5] Mauro Conti, Radha Poovendran, and Marco Secchiero. Fakebook: Detecting fake profiles in on line social networks. In *Proceedings of the First IEEE/ACM International Workshop on Cybersecurity of Online Social Network*, CSOSN '12, 2012.

[6] Emiliano De Cristofaro, Claudio Soriente, Gene Tsudik, and Andrew Williams. Hummingbird: Privacy at the time of twitter. In *IEEE Symposium on Security and Privacy*, pages 285–299. IEEE Computer Society, 2012.

[7] Leucio Antonio Cutillo, Refik Molva, and Melek Önen. Safebook: A distributed privacy preserving online social network. In *WOWMOM*, pages 1–3. IEEE, 2011.

[8] Roger Dingledine, Nick Mathewson, and Paul Syverson. Tor: the second-generation onion router. In *Proceedings of the 13th conference on USENIX Security Symposium - Volume 13*, SSYM'04, pages 21–21, Berkeley, CA, USA, 2004. USENIX Association.

[9] David Fifield, Nate Hardison, Jonathan Ellithorpe, Emily Stark, Dan Boneh, Roger Dingledine, and Phil Porras. Evading censorship with browser-based proxies. In Simone Fischer-Hübner and Matthew Wright, editors, *Privacy Enhancing Technologies*, volume 7384 of *Lecture Notes in Computer Science*, pages 239–258. Springer, 2012.

[10] Kahina Gani, Hakim Hacid, and Ryan Skraba. Towards multiple identity detection in social networks. In *WWW (Companion Volume)*, pages 503–504. ACM, 2012.

[11] Gábor Gy. Gulyás and Sándor Imre. Analysis of identity separation against a passive clique-based de-anonymization attack. *Infocommunications Journal*, III(4), December 2011.

[12] Michael Hay, Gerome Miklau, David Jensen, Donald F. Towsley, and Chao Li. Resisting structural re-identification in anonymized social networks. *VLDB J.*, 19(6):797–823, 2010.

[13] Carter Jernigan and Behram F. T. Mistree. Gaydar: Facebook friendships expose sexual orientation. *First Monday*, 14(10), 2009.

[14] Lei Jin, Hassan Takabi, and James B.D. Joshi. Towards active detection of identity clone attacks on online social networks. In *Proceedings of the first ACM conference on Data and application security and privacy*, CODASPY '11, pages 27–38, New York, NY, USA, 2011. ACM.

[15] Aleksandra Korolova, Rajeev Motwani, Shubha U. Nabar, and Ying Xu. Link privacy in social networks. In *Proceedings of the 17th ACM conference on Information and knowledge management*, CIKM '08, pages 289–298, New York, NY, USA, 2008. ACM.

[16] Wanying Luo, Qi Xie, and Urs Hengartner. Facecloak: An architecture for user privacy on social networking sites. In *CSE'09*, pages 26–33, Washington, DC, USA, 2009. IEEE Computer Society.

[17] Shah Mahmood and Yvo Desmedt. Your facebook deactivated friend or a cloaked spy. In *PerCom Workshops*, pages 367–373, 2012.

[18] Arvind Narayanan, Elaine Shi, and Benjamin I. P. Rubinstein. Link prediction by de-anonymization: How we won the kaggle social network challenge. In *IJCNN*, pages 1825–1834. IEEE, 2011.

[19] Arvind Narayanan and Vitaly Shmatikov. De-anonymizing social networks. *Security and Privacy, IEEE Symposium on*, 0:173–187, 2009.

[20] Mudhakar Srivatsa and Mike Hicks. Deanonymizing mobility traces: using social network as a side-channel. In Ting Yu, George Danezis, and Virgil D. Gligor, editors, *ACM Conference on Computer and Communications Security*, pages 628–637. ACM, 2012.

[21] Elena Zheleva and Lise Getoor. Preserving the privacy of sensitive relationships in graph data. In *Proceedings of the 1st ACM SIGKDD international conference on Privacy, security, and trust in KDD*, PinKDD'07, pages 153–171, Berlin, Heidelberg, 2008. Springer-Verlag.

[22] Bin Zhou and Jian Pei. The *k*-anonymity and *l*-diversity approaches for privacy preservation in social networks against neighborhood attacks. *Knowl. Inf. Syst.*, 28(1):47–77, 2011.