Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 129 (2018) 52–58

Procedia Computer Science

www.elsevier.com/locate/procedia

2017 International Conference on Identification, Information and Knowledge in the Internet of Things

Deeply Understanding Structure-based Social Network De-anonymization

Wenqian Tian[a], Jian Mao[a,*], Jingbo Jiang[a], Zhaoyuan He[a], Zhihong Zhou[b], Jianwei Liu[a]

[a]*School of Electronic and Information Engineering, BeiHang University, Beijing 100183, China*
[b]*Shanghai Key Laboratory of Integrated Administration Technologies for Information Security, Shanghai 200240, China*

Abstract

Anonymization techniques are widely adopted to protect users' privacy during social data publishing and sharing. In this paper, we conduct a comprehensive analysis on the typical structure-based social network de-anonymization algorithms. We aim to understand the de-anonymization approaches and disclose the impacts on their application performance caused by different factors. We design the analysis framework and define three experiment environments to evaluate a few factors' impacts on the target algorithms. Based on our analysis architecture, we simulate two typical de-anonymization algorithms and evaluate their performance under different pre-configured environments.

Copyright © 2018 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the 2017 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI2017).

Keywords: De-anonymization; Anonymization; Social Privacy; Social Network

1. Introduction

Nowadays, social network services have been developed rapidly as a fast-growing business. Social network websites/applications (e.g., Facebook, Youtube, Twitter, Reddit, etc.) are getting more and more popular. Social network platforms become huge social data resources, which have great commercial value and significant sociological impacts. Considering the commercial benefit and the social impact of these social network information, the social network service providers may release their social data to third parties for academic or commercial purposes. However, it will consequently introduce the risk of leaking users' sensitive information (e.g., identity, location, personal interests, etc.) [17, 30, 31, 32]. To protect users' privacy, the most straightforward solution is to anonymize data by removing users' identities. However, recent research demonstrates that this naïve solution is vulnerable to *auxiliary information-based de-anonymization* [17, 18, 24]. As a widely adopted approach in relational data anonymization, *k-anonymity* is introduced to protect social network data privacy against different attacks [33, 13, 34, 3]. In addition,

* Jian Mao. Tel.: +86-10-8231-7212; Fax: +86-10-8231-9474.
E-mail address: maojian@buaa.edu.cn

1877-0509 Copyright © 2018 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the 2017 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI2017).
10.1016/j.procs.2018.03.045

Aggregation/Class/Cluster based anonymization [5, 2, 25], differential-privacy-based mechanism [23, 22, 21, 27] and random-walk-based methods [15] are also proposed to preserve users' private information.

De-anonymization (DA) techniques are actively studied to identify vulnerabilities in current social-network data-publishing mechanisms [18]. Typically, these approaches can be classified into two categories, *Seed-based De-anonymization* [18, 19, 24, 28, 8, 12] and *Seed-free De-anonymization* [20]. In a seed-based de-anonymization approach, seed nodes are selected first between two networks and then propagation is conducted iteratively to the whole network. Nilizadeh et al. propose community-level-based de-anonymization [19] that can be used to improve other existing seed-based de-anonymization mechanisms [24, 28, 8, 12, 20]. Besides, some semantic-based de-anonymization methods are developed to break link privacy [11] or infer private attributes [26]. Ji et al. give a survey on the graph data anonymization and de-anonymization approaches [10].

There are many factors that influence the performances (e.g., accuracy, scalability) of existing DA algorithms, for example, the methods (anonymization) sanitizing the original data, parameter configuration, the size of the testing graph, link direction, density distribution of the graph nodes, etc. However, most of the existing de-anonymization approaches aim at one specific anonymization approach. In some cases, they did not provide specification of the methods sanitizing the raw data. Meanwhile, some of them also do not the evaluation on the de-anonymization accuracy under different parameter configuration, such as in [7].

In this paper, we conduct a comprehensive analysis on the typical structure-based social network de-anonymization algorithms to understand the structure-based de-anonymization approaches and disclose the impacts on their application performance caused by the factors mentioned above. We design the analyzing framework and define three experiment environments (anonymized algorithms, topology of graph and key parameters) to evaluate the factors' impacts on the target algorithms. Based on our analyzing architecture, we simulate two typical De-anonymization algorithms (N-DA [18] and Ji-DA [6]), and evaluate their performance under different pre-configured environments.

We summarize the contributions made by this paper as follows.We make a comprehensive analysis of different dimensions' influences on the accuracy of De-anonymization algorithms. We design the analyzing architecture and define three experiment environments. We simulate two typical structure-based De-anonymization algorithms and evaluate their performance under different pre-configured environments. Based on our evaluation, we conclude the parameter impacts on the testing approaches and the influences introduced by the topology properties of testing datasets.

2. Design and Implementation

In this section, we analyze the typical de-anonymization mechanisms in different experiment configurations. Our analysis framework consists of three parts, *Anonymization*, *De-Anonymization*, *Parameter Configuration*, as well as two datasets, *Original Data*, *Anonymized Data*. To evaluate the influence of the data processing methods (specifically, anonymization algorithms) and the key parameters on the efficiency of de-anonymization approaches, the overall process includes the following steps: First, we collect the original social data from the online social platforms and public social databases; Then, we create anonymized social data by using typical anonymized algorithms (selected by the configuration module); Finally, we conduct experiments with different parameter configurations and evaluate the efficiency (accuracy) of the target de-anonymization algorithms.

2.1. Dataset Preparation

We use a subset of the *Twitter* social network as our evaluation input. The dataset [14] was created in 2010, which consists of 90,907 users and 443,399 "follow" relationships. It is a directed graph. We divide the twitter-network into several parts to achieve better evaluation results. We use a *center-spread* method to obtain several smaller subsets for experiments. And we process the raw/original data in different ways with regard to directed and undirected graphs. The dataset splitting method is listed as follows.

Step 1. When undirected subsets are required, we transfer the original directed network into an undirected graph by using the approach mentioned in the work [4], keeping the edge that only exists bilaterally. If directed subsets are required, we directly go to the next step.

Step 2. Select m max-degree nodes $\{v_1, v_2, \cdots, v_m\}$ in the original graph G(twitter-network), and put them in the top-degree set denoted as $T\,set$.

Table 1: Subset samples of anonymized methods

Original Graph	Node	Edge	Anonymized Graph	Edge Overlap
graph_1 – 2	196	431	*graph_1 – 2_add – del/kda/union*	0.82
Digraph_1 – 2	379	1739	*Digraph_1 – 2_add – del/kda/union*	0.61

Step 3. Let each $v_i \in T set$ be the center. Select the neighbors (both in_edge neighbor and out_edge neighbor for directed graph) of v_i and add them into $T set$.

Step 4. Repeat the Step 3 for n times and obtain a subset $graph\_m - n$ for undirected graph and $Digraph\_m - n$ for directed graph.

We choose different m and n to obtain different size of datasets and to see whether the size of graph affect the accuracy of de-anonymization algorithms. The subsets obtained by following Step 1 through Step 4 are used as auxiliary graphs to re-identify anonymized graphs.

2.2. Selected Anonymization Algorithms

In this work, we use the following anonymized methods to prepare sanitized social dataset in our experiments.

Naïve Add/Del Edges Method. For naïve *edge-edit* anonymized method, we choose the Add/Del Edges method [29] as one of our sanitized approaches, which protects node and link privacy of graph data by adding or deleting edges randomly through the whole graph. We use this method to anonymize all the datasets with different sizes. When using it we set the fraction of edges that we want to edit. For instance, if we set the edition fraction is 0.1, actually the overlap of edges we get will be lower than 0.9. After a subset $graph\_m - n$ ($Digraph\_m - n$) being anonymized by Add/Del Edges Method, it is denoted as $graph\_m - n\_add - del$ ($Digraph\_m - n\_add - del$).

***k*-Degree Anonymization Method**. k-anonymity-based solutions are also a typical choice for preserving social data privacy. We select a representative variant of k-anonymity based method, that is k-degree anonymization [13], as a candidate algorithm for social data sanitization/pre-processing. It works as that for every node there exist at least $k-1$ nodes with same degree in the graph. In the algorithm, we set the different k to get different anonymized graph with different overlaps. A subset $graph\_m - n$ ($Digraph\_m - n$) being anonymized by k-Degree Anonymization Method, we denote it as $graph\_m - n\_kda$ ($Digraph\_m - n\_kda$) in this paper.

Union-Split Method. The cluster-based methods [5, 25, 2] are similar to k-anonymity based methods. The aim is to make nodes in a cluster indistinguishable on structure. There are several approaches to implement it, such as t-means [25], union-split [25], and so on. We use union-split method to anonymize graphs, denoted as $graph\_m{-}n\_union$ ($Digraph\_m - n\_union$).

We use three anonymization algorithms to obtain anonymized datasets by using secGraph [7] tool. The processed subset samples are listed in Table 1. The evaluation results are illustrated in the Section 3.

2.3. Target De-Anonymization Algorithms

We select two representative graph de-anonymization algorithms to analyze in detail and simulate to test their de-anonymizability, respectively.

2.3.1. Narayanan et al. De-Anonymization (N-DA)

The N-DA algorithm [18] takes two directed graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ where weights of edges are set to 1, and outputs a re-identified match $\mu : V_1 \rightarrow V_2$ between two graphs. It consists of seed production and propagation steps. In seed-based de-anonymization, as the seed production is not our first priority, we mainly focus on the propagation stage. In our implementation, seed production can be implemented in the same way as existing solutions [1, 18, 16, 24]. So in the seed production stage, we assume we have selected k seed mappings, denoted by $M_s = \{(s_1, s_1^{'}), (s_2, s_2^{'}), \cdots, (s_k, s_k^{'})\}$, where $s_i \in V_1$, $s_i^{'} \in V_2$ and $s_i^{'} = \mu(s_i)$. In the propagation stage, for each iteration, it picks an unmapped node $v \in V_1$ and calculates a *score* for each $(v, v^{'}), v^{'} \in V_2$. The mappings between $(v, v^{'})$ with a score over a threshold will remain. When switching the two input graphs, if $v^{'}$ maps back to v, then the mapping

between v and $v^{'}$ will be added to the output mapping list. The propagation does not converge until no more mappings can be added to the final list. The *score* above equals to the number of common nodes of v and $v^{'}$ that have been mapped. In this algorithm, θ is an important parameter that influence the output accuracy greatly. It is the difference of the max score and the second max score divided by the standard deviation of the mapping set, denoted as *eccentricity* in [18]. In this paper, we analyze the parameter in different angles.

2.3.2. Ji et al. De-Anonymization (Ji-DA)

Ji-DA algorithm [9] takes two undirected graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ and seed mappings M_s as inputs. The output is the mapping between these two graphs. For each iteration, it starts from the neighbors of the already-mapped nodes M, calculating a *unified similarity score* $s(v, v^{'})$ between every pair of $\{v|v \in G_1 \& v \notin M\}$ and $\{v^{'}|v^{'} \in G_2 \& v^{'} \notin M\}$ to construct a weighted bipartite graph B based on $s(v, v^{'})$. It uses the Hungarian algorithm to obtain a *maximum weighted bipartite matching* $M^{'}$ of B. A *threshold* and a *TOP-K* strategy are used to remove some improper mappings. Finally, the remained mappings are added into the mapping M. The whole algorithm contains many parameters that may be hard to control. In our experiment, we especially pay attention to the impacts on the similarity score s caused by three weighing parameter c_S, c_D and c_I, where c_S, c_D and c_I represents the weights of structural similarity, relative distance similarity and inheritance similarity, correspondingly.

2.4. Critical Factors and Experiment Design

We use the mechanisms in [18, 9] as our target algorithms. For the selected algorithms [18, 9] and other latest de-anonymization attacks, the number of seeds and noise proportion are two general pre-conditions for researchers evaluating the efficiency of their approaches. In this work, we demonstrate that the key parameters and the anonymization algorithms used to process the raw social data are two other important issues that significantly influence the efficiency of the de-anonymization approaches.

Algorithm parameter configuration. Most de-anonymization algorithms have one or more parameters. For the algorithm in [18], we analyze the influence of accuracy with regard to the eccentricity θ in different angles. We choose some directed subsets we described above and use different anonymization methods. For the algorithm in [9], we analyze the effect of accuracy with respect to three weighing factors c_S, c_D and c_I to observe the connections among them.

Topology properties of the social data. The graph structure is bound to influence the accuracy. In previous section, we have created many subsets of graph data by using a *center-spread* method. There are two ways of spreading. One is *depth-spread* when we increase n keeping the constant m and the other is *width-spread* when we increase m keeping the constant n. So we will obtain two types of subsets. In this paper, we use both types to analyze the de-anonymization methods.

- **Depth-Spread.** In our experiments, we choose the maximum degree node and expand to its neighbors in deep-level. The selected subsets are listed in Table 2[1].
- **Width-Spread.** In this paper, we choose several max-degree node in width-level and expand to its neighbors in two-level. The subsets are shown in Table 3[2] .

3. Evaluation

In this section we evaluate different de-anonymization approaches described in Section 2 using three different anonymization methods from different angles we mentioned above. In the following part, we will discuss each de-

[1] As the complexity of algorithms increase when the m and n increasing so here we choose some small m and n here to complete high efficiency and different n values to obtain different depth size.

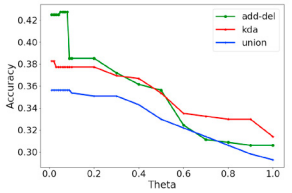
[2] Here we choose some small m and n here to complete high efficiency and different m values to obtain different width size.

Table 2: Sample subsets selected according to Depth-spread

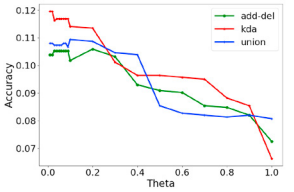
Raw Network	Nodes	Edges	Av.Deg
*graph*_1 – 2	196	431	4.398
*graph*_1 – 3	416	785	3.774
*graph*_1 – 4	1062	1680	3.164
*graph*_1 – 5	3599	5905	3.281
*Digraph*_1 – 2	379	1739	9.177
*Digraph*_1 – 3	1463	3639	4.975
*Digraph*_1 – 4	7036	15041	4.275

Table 3: Sample subsets selected according to Width-spread

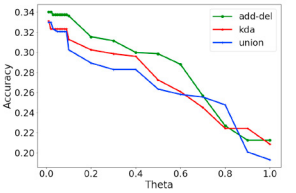
Raw Network	Nodes	Edges	Av.Deg
*graph*_1 – 2	196	431	4.398
*graph*_2 – 2	601	1029	3.424
*graph*_3 – 2	753	1204	3.293
*graph*_4 – 2	981	1578	3.217
*Digraph*_1 – 2	379	1739	9.177
*Digraph*_2 – 2	767	3032	7.906
*Digraph*_3 – 2	1096	7702	14.055



(a) θ's accuracy influence in *Digraph*_1 – 2



(b) θ's accuracy influence in *Digraph*_1 – 3



(c) θ's accuracy influence in *Digraph*_2 – 2

Fig. 1: Experiment results of θ's accuracy influence on N-DA

anonymization algorithm respectively. All the datasets of the following experiments are based on the subsets we configured in Section 2.

3.1. Experiment Analysis of N-DA

In this experiment, we evaluate the influence of the essential parameter *eccentricity*, θ, on the accuracy of N-DA algorithm under different anonymization methods and graph topologies. We directly set 50 top-degree nodes as seeds, as we described previously that the seed identification stage is not our primary purpose. When we analyze the effects of different anonymization methods regarding to one subset, we set the edge overlap between anonymized graph and auxiliary graph to be same. For simplicity the node overlaps in this paper are set to be 1, i.e., the edge overlap between *Digraph*_1 – 2 and *Digraph*_1 – 2_*add* – *del*, *Digraph*_1 – 2_*kda*, *Digraph*_1 – 2_*union* respectively is the same. We use three groups of graphs used for N-DA. *Digraph*_1 – 2 and its corresponding three anonymized graphs of three anonymized methods and edge overlap of each pair of datasets is about 61%. And similarly edge overlap of each pair of *Digraph*_1 – 3 and its corresponding three anonymized graphs is about 58% and edge overlap of each pair of *Digraph*_2 – 2 and its corresponding three anonymized graphs is about 63%. The results for different preferences are displayed in Figure 1.

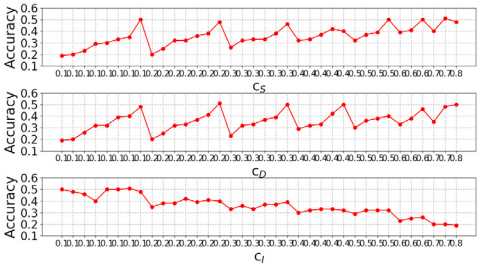
Observation. The experiment results demonstrate that regardless of the dataset topology or anonymization algorithms, the accuracy goes down with the parameter θ getting higher. So, a lower value of the parameter θ will contribute to a higher re-identification accuracy. And as the depth-spread dataset got much lower accuracy then the width-spread dataset, so width-spread graphs seem to be more vulnerable to N-DA.

3.2. Experiment Analysis of Ji-DA

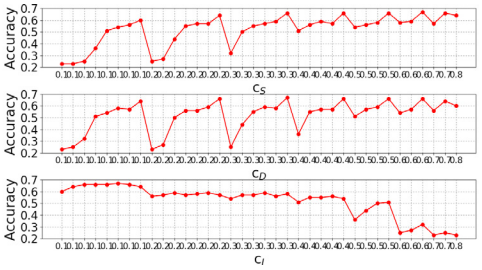
In this subsection, we evaluate the influence of three key parameters, different anonymization, and graph topology on accuracy of Ji-DA. We choose the subset *graph*_1 – 4 with *graph*_1 – 4_*add* – *del* and *graph*_1 – 4_*kda* (in which the edge overlap is 82%) as anonymized graphs to analyze the three weighing parameters c_S, c_D and c_I.

To set different preferences, we use full arrangement of three parameters that each value of the parameter varies from the interval [0.1, 0.8] whose increasing step is 0.1. Besides, the sum the of three parameters c_S, c_D, c_I is 1. For example, $c_S = 0.1$, $c_D = 0.1$, $c_I = 0.8$; $c_S = 0.2$, $c_D = 0.5$, $c_I = 0.3$. There are 36 groups of parameter settings of these three parameters. The other parameters are configured as: $C = 0.9, \theta = 0.9, \delta = 1$, and $\epsilon = 0.5$.

The results of three key parameters regarding to subset *graph*_1 – 4 with *graph*_1 – 4_*add* – *del* are displayed in Figure 2 (a). The results of three key parameters regarding to subset *graph*_1 – 4 with *graph* 1 – 4 *kda* are shown



(a) Effect w.r.t. $graph\_1-4$ with $graph\_1-4\_add-del$



(b) Effect w.r.t $graph\_1-4$ with $graph\_1-4\_kda$

Fig. 2: Effect of weighing parameters on accuracy.

in Figure 2 (b). As we do not optimize the other parameter such as θ and the overlap between two graphs is small, the accuracy may be a little bit low. However, our purpose here is not to maximum the accuracy but to analyze the influence tendency of the three weighing parameters.

Observation. The experiment results show that the accuracy goes down when c_I gets bigger, while c_S and c_D seems to present period tendency. In [9], c_S represents the *Structural Similarity* of a node, which considers the node's global information and c_D represents the *Relative Distance Similarity* of a node, which also considers the nodes' global information partly. So, the two parameters have some common meaning. However, the c_I represents the *Inheritance Similarity* of a node, which considers the node's nearby neighborhood that have mapped so it is different from the previous two parameters. Accordingly, we consider that the parameter c_I may influence the re-identification accuracy of this algorithm greatly. A small value of c_I contributes to a high accuracy.

4. Conclusion

In this paper, we conduct a comprehensive analysis on the typical structure-based social network de-anonymization algorithms to achieve a deep understanding on the de-anonymization approaches and disclose the impacts on their application performance caused by the different factors. We design the analyzing framework and define three experiment environments to evaluate the impacts by the factors on the target algorithms. Based on our framework, we simulate two typical de-anonymization algorithms and evaluate their performance under several pre-configured environments.

Acknowledgements

This work was supported in part by the National Key R&D Program of China (No. 2017YFB0802400), the National Natural Science Foundation of China (No. 61402029), and the Funding Project of Shanghai Key Laboratory of Integrated Administration Technologies for Information Security (No. AGK201708).

References

[1] Backstrom, L., Dwork, C., Kleinberg, J., 2007. Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography, in: Proceedings of the 16th international conference on World Wide Web, ACM. pp. 181–190.

[2] Bhagat, S., Cormode, G., Krishnamurthy, B., Srivastava, D., 2009. Class-based graph anonymization for social network data. Proceedings of the VLDB Endowment 2, 766–777.

[3] Cheng, J., Fu, W.C., Liu, J., 2010. K-isomorphism: Privacy preserving network publication against structural attacks, in: Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data, ACM. pp. 459–470.

[4] Gong, N.Z., Talwalkar, A., Mackey, L., Huang, L., Shin, E.C.R., Stefanov, E., Song, D., et al., 2011. Jointly predicting links and inferring attributes using a social-attribute network (san). Computer Science 1112.3265.

[5] Hay, M., Miklau, G., Jensen, D., Towsley, D., Weis, P., 2008. Resisting structural re-identification in anonymized social networks. Proceedings of the VLDB Endowment 1, 102–114.

[6] Ji, S., Li, W., Gong, N.Z., Mittal, P., Beyah, R., 2016a. Seed-based de-anonymizability quantification of social networks. IEEE Transactions on Information Forensics and Security 11, 1398–1411.

[7] Ji, S., Li, W., Mittal, P., Hu, X., Beyah, R.A., 2015. Secgraph: A uniform and open-source evaluation system for graph data anonymization and de-anonymization., in: Proceedings of the 24th USENIX Security Symposium, USENIX. pp. 303–318.
[8] Ji, S., Li, W., Srivatsa, M., He, J.S., Beyah, R., 2014. Structure based data de-anonymization of social networks and mobility traces, in: Proceedings of the International Conference on Information Security (ISC), Springer. pp. 237–254.
[9] Ji, S., Li, W., Srivatsa, M., He, J.S., Beyah, R., 2016b. General graph data de-anonymization: From mobility traces to social networks. ACM Transactions on Information and System Security (TISSEC) 18, 12.
[10] Ji, S., Mittal, P., Beyah, R., 2016c. Graph data anonymization, de-anonymization attacks, and de-anonymizability quantification: A survey. IEEE Communications Surveys & Tutorials 19, 1305–1326.
[11] Korolova, A., Motwani, R., U. Nabar, S., Xu, Y., 2008. Link privacy in social networks, in: Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM), ACM. pp. 239–298.
[12] Korula, N., Lattanzi, S., 2014. An efficient reconciliation algorithm for social networks, VLDB Endowment. pp. 377–388.
[13] Liu, K., Terzi, E., 2008. Towards identity anonymization on graphs, in: Proceedings of the 2008 ACM SIGMOD international conference on Management of data, ACM. pp. 93–106.
[14] Lou, T., Tang, J., Hopcroft, J., Fang, Z., Ding, X., 2013. Learning to predict reciprocity and triadic closure in social networks. ACM Transactions on Knowledge Discovery from Data (TKDD) 7, 5.
[15] Mittal, P., Papamanthou, C., Song, D., 2012. Preserving link privacy in social network based systems. Computer Science 1208.6189, 1–16.
[16] Narayanan, A., Shi, E., Rubinstein, B.I., 2011. Link prediction by de-anonymization: How we won the kaggle social network challenge, in: Neural Networks (IJCNN), The 2011 International Joint Conference on, IEEE. pp. 1825–1834.
[17] Narayanan, A., Shmatikov, V., 2008. Robust de-anonymization of large sparse datasets, in: Proceedings of 2008 IEEE Symposium on Security and Privacy, IEEE. pp. 111–125.
[18] Narayanan, A., Shmatikov, V., 2009. De-anonymizing social networks, in: Proceedings of 2009 IEEE Symposium onSecurity and Privacy, IEEE. pp. 173–187.
[19] Nilizadeh, S., Kapadia, A., Ahn, Y.Y., 2014. Community-enhanced de-anonymization of online social networks, in: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, ACM. pp. 537–548.
[20] Pedarsani, P., Figueiredo, D.R., Grossglauser, M., 2013. A bayesian method for matching two similar graphs without seeds, in: Proceedings of 51st Annual Allerton Conference on Communication, Control, and Computing, IEEE. pp. 1598–1607.
[21] Proserpio, D., Goldberg, S., Mcsherry, F., 2012a. Calibrating data to sensitivity in private data analysis: a platform for differentially-private analysis of weighted datasets. Proceedings of the VLDB Endowment 7, 637–648.
[22] Proserpio, D., Goldberg, S., Mcsherry, F., 2012b. A workflow for differentially-private graph synthesis, in: Proceedings of the 2012 ACM workshop on Workshop on Online Social Networks, ACM. pp. 13–18.
[23] Sala, A., Zhao, X., Wilson, C., Zheng, H., Zhao, B.Y., 2011. Sharing graphs using differentially private graph models, in: Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference (IMC), ACM. pp. 81–98.
[24] Srivatsa, M., Hicks, M., 2012. Deanonymizing mobility traces: Using social network as a side-channel, in: Proceedings of the 2012 ACM conference on Computer and communications security, ACM. pp. 628–637.
[25] Thompson, B., Yao, D., 2009. The union-split algorithm and cluster-based anonymization of social networks, in: Proceedings of the 4th International Symposium on Information, Computer, and Communications Security (ASIACCS), ACM. pp. 218–227.
[26] Wondracek, G., Holz, T., Kirda, E., Kruegel, C., 2010. A practical attack to de-anonymize social network users, in: Proceedings of the 2010 IEEE Symposium on Security and Privacy, IEEE. pp. 223–238.
[27] Xiao, Q., Chen, R., Tan, K.L., 2014. Differentially private network data release via structural inference, in: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), ACM. pp. 911–920.
[28] Yartseva, L., Grossglauser, M., 2013. On the performance of percolation graph matching, in: Proceedings of the 1st ACM Conference on Online Social Networks, ACM. pp. 119–130.
[29] Ying, X., Wu, X., 2008. Randomizing social networks: a spectrum preserving approach, in: Proceedings of the 2008 SIAM International Conference on Data Mining, SIAM. pp. 739–750.
[30] Zheng, X., Cai, Z., Li, J., Gao, H., 2017a. Location-privacy-aware review publication mechanism for local business service systems, in: IEEE INFOCOM 2017 - IEEE Conference on Computer Communications, IEEE. pp. 1–9.
[31] Zheng, X., Cai, Z., Yu, J., Wang, C., Li, Y., 2017b. Follow But No Track: Privacy Preserved Profile Publishing in Cyber-Physical Social Systems. IEEE Internet of Things Journal 4, 1868–1878.
[32] Zheng, X., Luo, G., Cai, Z., 2018. A Fair Mechanism for Private Data Publication in Online Social Networks. IEEE Transactions on Network Science and Engineering , 1–1.
[33] Zhou, B., Pei, J., 2008. Preserving privacy in social networks against neighborhood attack, in: Proceedings of the 24th IEEE International Conference on Data Engineering (ICDE), IEEE. pp. 506–515.
[34] Zou, L., Chen, L., Özsu, M.T., 2009. K-automorphism: A general framework for privacy preserving network publication. Proceedings of the VLDB Endowment 2, 946–957.