Deanonymizing Users in Social Networking Services: an Ego-Network Analysis Approach

Vladislav Chesnokov
Information Security Department
Bauman Moscow State Technical University
Moscow, Russia
v.o.chesnokov@yandex.ru

Peter Klyucharev
Information Security Department
Bauman Moscow State Technical University
Moscow, Russia
pk.iu8@yandex.ru

***Abstract* — One of the main problems in social networking services monitoring systems is the incompleteness of analyzed data. Anonymous users may participate in information warfare tampering public opinion. A new method for user profiling in social networking services is proposed. It is based on analysis of user ego-network communities. The core idea of the method is that each user has a unique set of attributes and user attributes are strongly related to his ego-network communities. User profile attributes can be obtained as the union of attributes relevant to each community. The method was compared with majority voting and two community detection based approaches. Experiments on four datasets from Facebook, Twitter, LiveJournal and VKontakte social networking services showed that the proposed method outperforms others and some user attributes can be determined with high precision and recall. The method is tolerant to node attributes partial absence. User attributes determined by the proposed method combined with additional sources of information may lead to user deanonimization.**

Keywords — deanonimization; user profiling; ego-network; social graph; user profile attributes; community detection; information warfare; social networking service

I. Introduction

Social networking service (SNS) monitoring is one of the key instruments for public opinion analysis, which is used in marketing, politics, science, etc. A good report must contain statistics about its target audience attributes. However, some users in SNS may be anonymous, and some people may create profiles just to tamper the public opinion regarding some topic. In recent years there were many examples how SNS can be used to manipulate its users and influence "the real world" [1, 2, 3]. Revealing and deanonimization such profiles are the key problems of information warfare.

SNS allow users to create profiles filled with personal data. Most users publish a huge amount of personal attributes such as age, location, education, favorites, photos, etc. However, not everyone does that. There are several reasons to have empty or scarcely filled profile: privacy issue, laziness, paranoia and others. Some users may think that fake names and no profile photo can assure their anonymity. Meanwhile they can still use SNS to its full extent, communicating with their friends and family, listing social links in their profile, reposting publications they interested in and performing other actions revealing their identity.

Depending on quality and amount of data SNS users deanonimization methods can be divided into four categories:

- ones that use only publicly available data: user profiles, photos, publications, etc;
- technical methods (e.g. browser fingerprints);
- cross-SNS analysis;
- social engineering.

Of course, these methods can be combined. Methods from the first category are based on user attributes inferring and their further analysis. It is quite obvious that a set of attributes like age, sex, location, education, workplace, etc is unique for each person, so having them one can obtain user identity. Using additional data sources like graduates lists or other user databases one can even get their full name.

II. Problem Definition

Consider unweighted undirected graph *G'(V', E')* with *D(G')* = 2 which has $u \in V'$ such that

$$\forall v \in V', v \neq u \; \exists \{u, v\} \in E', \quad (1)$$

where *D(G')* is the diameter of graph *G'*. Then it is called an ego-network for node *u*, while *u* is an ego or center. Denote

$$V = V' \setminus \{u\}$$

$$E = E' \setminus \{\{u, v\} | v \in V\}' \quad (2)$$

. (3)

Each vertex from *V'* has a set of attributes (sometimes called features) from set *F*. I.e. there exists a map $f\colon V \to 2^F$. In practice, only a part of vertex attribute set is observable, so denote it by *f'* such that

$$\forall v : f'(v) \subseteq f(v).$$

$p \subseteq F$ (4)

This work was supported by RFBR (grant № 16-29-09517).

The task of profile inferring is to obtain such user profile of vertex u that

$$\delta(p, p^*) \to \max \tag{5}$$

where $p^* = f(u)$ and δ – some similarity measure of two sets, e.g. F-measure.

III. RELATED WORK

There are quite lot methods which can be used to infer user profile attributes. Most of them are based on machine learning. Others can use attribute inferring, majority voting, user preferences or community detection.

The most basic method is the majority voting: the more neighbors (or “friends”) of vertex has some attribute, the higher the probability that this vertex has this attribute. This method was used to determine the geolocation of Twitter users in [4]. It has quite high precision (86%) but low recall (20%), because only small fraction of users provides their location. More complex methods include attribute transfer. For example, in [5] user geolocation was determined by median value of its neighbors location. The error of the method was less than 10 kilometers for half of the users from Twitter dataset. In [6] user preferences were transferred through their social ties and in [7] Facebook user geolocation was determined in the same way.

Methods based on user preferences use more complex information about a user, which is harder to obtain. In [8] profile attributes were determined by music preferences, and in [9] – by user reactions (“likes”). PGPI method proposed in [10] allows determining some attributes of user profile with accuracy higher than 90%. More than that, it needs only limited amount of information (number of “facts”) and uses users attributes, group membership, publications view count and likes count.

The quality of user profiles got by machine learning based methods highly depends on quality of training data. In [11] a comparative analysis of several models was conducted. The task was to determine age and sex of 7 million telecom provider clients by calls and SMS data. The best model showed 0.85 for sex and 0.72 for age values of F-measure with 90% data as training. One of the common approaches is to use user attributes with features extracted from their publications [12-15]. Accuracy higher than 89% for age and sex detection can be reached with this method [12]. Machine learning can be used to combine several approaches. For example, in [16] features like distance between user favorite publications and profiles were used to determine his geolocation. Of course one can build an ensemble of classifiers to determine user attributes.

The proposed method is based on community detection approach. In [17] a greedy community detection algorithm regarding each attribute based on conductivity maximization is proposed. Given that value of some attribute is known for at least 20% of users, the accuracy for this attribute for other users is 80% on Facebook dataset for students and professors of two universities [17]. In [18] it was concluded that applying community detection algorithm to full social graph is redundant and may lead to low results. More efficient methods are based on ego-network community detection. For example, in [19] a generative model was used to obtain user circles (local communities). In [18] some attributes of user profile were determined by ego-network communities with 70% accuracy on LinkedIn dataset. One of the main ideas of the method is to determine a type of links between ego-user and others with respect to vertex attributes. It was shown that this method outperforms methods from [17] and [19] by accuracy of user profiling and community detection. However, this method requires knowledge about attributes nature and supposes that communities may not intersect.

IV. THE PROPOSED METHOD

It is $\mathscr{C} = \{C_i\}$ well known that social ties of a person are not random. Users connected to him can be separated into several groups formed by some attribute: classmates, colleagues, relatives, close friends, etc. Consider ego-network communities. The result of the community detection algorithm proposed in [20] is a cover with corresponding label sets $A_i \subseteq F$ describing them. Suppose that the central vertex belongs to all its ego-network communities. So, most likely, it would have all attributes of them. The proposed method is of determining user profile is simply a union of attributes sets corresponding to these communities:

$$p = \bigcup_i A_i \tag{6}$$

The main problem is to obtain this correspondence. The algorithm from [20, 21] can be simplified as we don’t need the community membership, just labels. First, associate with each vertex v an empty set of key attributes K_v. Next, start an iterative process of updating sets of key attributes. At each iteration all nodes are visited. For each vertex v with neighbors set $\mathcal{N}_v$ and each attribute a sets

$$\mathcal{N}'_{v,a} = \{w | w \in \mathcal{N}_v \wedge a \in f'(w)\} \tag{7}$$

and

$$Q_{v,a} = \{w | w \in \mathcal{N}_v \wedge a \in K_w\} \tag{8}$$

are computed. If the sum is $|\mathcal{N}'_{v,a}| + |Q_{v,a}|$ greater than the threshold $\max(2, \alpha_a|\mathcal{N}_v|)$, where α_a is a fraction defining qualified majority for attribute a, then attribute a is added to the set of key attributes K_v. The process stops when there were no changes at the last iteration. The original community detection algorithm has four more steps to determine community membership from key attributes sets and filter some attributes. However without great loss of accuracy only described steps are required in the simplified algorithm. So, the profile can be determined as

$$p = \bigcup_{v \in V} K_v \tag{9}$$

Of course, not all attributes can be determined reliably using this method, e.g. sex or last name. So, the attributes should be filtered before applying the method. Some attributes can be grouped if their nature is known – e.g. education. Unfortunately, attribute filtering is quite hard task nowadays, and may depend on SNS type. Some solutions for that may include statistical analysis of obtained profiles or machine learning. Most frequently, a data scientist may filter them by hand on in semi-automatic way.

V. Method Evaluation

The method was evaluated on four ego-networks datasets from Facebook[1], Twitter[2], VKontakte[3] and LiveJournal[4]. The first two datasets are from SNAP dataset [22]. Two other datasets were crawler by author. All datasets characteristics are shown in Table 1. Each graph in all datasets has a profile (a set of binary attributes) for each vertex, including ego.

Precision, recall and *F*-measure were used to measure the quality of obtained profiles. The proposed method was also compared with some others: simple majority; majority by communities (obtained by one of the five community detection algorithms: modularity maximization [23], Infomap [24], AGM-fit [25], BigCLAM [26], CESNA [27] or ground-truth communities); method based on attribute weights in communities obtained by CESNA. For each method several threshold values were used, and for each graph threshold with the best value of *F*-measure was used. For each SNS and profile detection method quality measures were averaged.

TABLE I. Datasets

SNS	number of graphs	avg. number of vertices	avg. number of edges	avg. number of attributes
Facebook	10	417	17 017	228
Twitter	973	138	2 350	665
VKontakte	2 000	164	1 561	1 665
LiveJournal	700	149	1 335	2 368

TABLE II. Method evaluation results (no attribute filtering). Average values of *F*-measure

	Facebook	Twitter	VKontakte	LiveJournal
simple majority	0.462	0.155	0.058	0.210
ground-truth communities	0.515	0.190	n/a	n/a
Infomap communities	0.491	0.183	0.170	0.099
modularity maximization communities	0.532	0.190	0.154	0.152
AGM-fit communities	0.516	0.189	0.187	0.209
BigCLAM communities	0.570	0.234	0.260	0.264
CESNA communities	0.535	0.222	0.218	0.236
CESNA attribute weights	0.312	0.187	0.225	0.300
proposed method	0.600	0.274	0.323	0.369

[1] https://facebook.com
[2] https://twitter.com
[3] https://vk.com
[4] https://livejournal.com

TABLE III. Method evaluation results (with attribute filtering). Average values of *F*-measure

	Facebook	Twitter	VKontakte	LiveJournal
simple majority	0.769	0.642	0.424	0.678
ground-truth communities	0.759	0.756	n/a	n/a
Infomap communities	0.726	0.641	0.557	0.504
modularity maximization communities	0.793	0.675	0.578	0.625
AGM-fit communities	0.780	0.669	0.624	0.668
BigCLAM communities	0.860	0.720	0.728	0.750
CESNA communities	0.842	0.719	0.683	0.727
CESNA attribute weights	0.466	0.752	0.767	0.874
proposed method	0.841	0.866	0.860	0.908

The results for all methods without attributes filtering are shown in Table I. The values for all methods are quite low. So, for each datasets attributes were filtered. They were grouped by similarity and description. For Facebook dataset attributes corresponding to education and hometown were kept (612 total). In Twitter dataset attributes are hashtags and replies, so only most popular hashtags were kept (86 total). For VKontakte dataset, education and workplace attributes were used (61 361 total). For LiveJournal dataset attributes corresponding to education and hometown were kept (47 106 total).

The results for datasets with filtered attributes are shown in Table III. Average values for *F*-measure are significantly higher for all methods. Perhaps they can be improved even further with another attribute filtering. First of all, the result for Twitter, VKontakte and LiveJournal datasets are similar for all methods, but different for Facebook datasets. It can be connected with small size of Facebook dataset or its nature (it was created by volunteers [22], while others were crawled from random profiles).

Simple majority method showed quite low results for all datasets except Facebook. Precision and recall tests demonstrated that this method lacks precision. One of the best results was given by ground-truth communities. Of course, this method cannot be applied in practice, but high results for it confirm hypothesis that community structure is strongly connected with profiles attributes, and improvement of community detection algorithms may lead to increase of precision and recall of profile attributes detection methods.

Methods based on communities obtained by graph clustering algorithms (modularity maximization and Infomap) are not suitable for attributes detection. One of the probable reasons for that is that communities do not intersect and their union must contain all graph vertices. Similar results for community detection evaluation were obtained in [20], so the better community detection, the more accurate user profiling. Community detection algorithms based on affiliation model outperform graph clustering algorithms, and the similar results can be declared for user profiling based on communities obtained by these algorithms.

Second best method is user profiling by attributes weights in communities given by CESNA. However it showed the worst result on Facebook dataset.

The proposed method ties for the best result with CESNA communities method for Facebook dataset, and the top results for Twitter, VKontakte and LiveJournal datasets outperforming closest competitors by 14.5%, 12% and 4% respectively.

Considering some chosen groups of attributes (see table IV) the results given by the proposed method by *F*-measure, precision and recall are close to the maximum, 1.

TABLE IV. *F*-MEASURE, PRECISION AND RECALL OF THE PROPOSED METHOD FOR CHOSEN ATTRIBUTE GROUPS

Dataset	Attribute group	F-measure	Precision	Recall
Facebook	education	0.902	0.967	0.875
Facebook	hometown	1.000	1.000	1.000
VKontakte	middle school	0.919	0.910	0.973
VKontakte	faculty	0.994	0.995	0.996
VKontakte	work place	0.994	0.997	0.994
LiveJournal	education	0.975	0.991	0.970
LiveJournal	hometown	0.984	0.977	1.000

VI. TOLERANCE TO ATTRIBUTE ABSENCE

Real datasets often lack some information about nodes attributes. So, the methods' tolerance to node attributes partial absence was tested. Note that the experiments were performed on crawled datasets, so they have already have been imperfect.

For each graph in each dataset 10, 20 … 90% random attributes values were removed. In other words, each dataset had nine copies of it with less and less data.

The precision of most methods with exception of the proposed method and CESNA weights method sufficiently drops on Facebook dataset after deleting 60% of node attributes, but recall grows. On Twitter dataset precision of all methods grows but recall drops because the obtained profiles have less "extra" attributes. After removing up to 80% of attribute values the proposed method and CESNA weights method significantly outperform others. For VKontakte and LiveJournal datasets results are similar to Twitter dataset.

Summing up, the experiments showed two obvious leaders which preserve high values of precision and recall outperforming other methods: the proposed method and CESNA weights method. Both of them are highly tolerant to attribute partial absence. So, methods based on joint analysis of ego-network structure and nodes attributes are better suited for user profiling than others.

VII. CONCLUSION

A user profile detection method was proposed. It is based on community detection algorithm. The core idea of the method is that community structure of user ego-network is strongly connected with attributes of his profile.

The experiments showed that the proposed method outperforms other attribute transfer based methods by *F*-measure, precision and recall. Some attributes can be determined with precision and recall close to one. The method preserves its performance if node attributes are partially absent.

This method can be used to determine user identity and deanonymize users in social networking services. Method application could improve public opinion analysis and help to detect and counteract users who tries to tamper it.

References

[1] J. Ratkiewicz et al.Detecting and Tracking Political Abuse in Social Media. Proc. 5th International AAAI Conference on Weblogs and Social Media (ICWSM). — North America, 2011. DOI: 10.1.1.646.5073

[2] Metaxas P. T., Mustafaraj E. Social Media and the Elections. Science. — 2012. — Vol. 338, No. 6106. — P. 472–473. DOI: 10.1126/science.1230456.

[3] Khondker H. H. Role of the New Media in the Arab Spring. Globalizations. — 2011. — Vol. 8, no. 5. — P. 675–679. DOI: 10.1080/14747731.2011.621287.

[4] Inferring the Location of Twitter Messages Based on User Relationships. / C. A. J. Davis [et al.]. T. GIS. — 2011. — Vol. 15, no. 6. — P. 735–751. DOI: 10.1111/j.1467-9671.2011.01297.x .

[5] Jurgens D. That's What Friends Are For: Inferring Location in Online Social Media Platforms Based on Social Relationships.. Proceedings of the AAAI International Conference on Weblogs and Social Media. — 2013.

[6] Pennacchiotti M., Popescu A.M. Democrats, Republicans and Starbucks Afficionados: User Classification in Twitter. Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. — San Diego, California, USA: ACM, 2011. — P. 430–438. DOI:10.1145/2020408.2020477.

[7] Backstrom L., Sun E., Marlow C. Find Me if You Can: Improving Geographical Prediction with Social and Spatial Proximity. Proceedings of the 19th International Conference on World Wide Web. — Raleigh, North Carolina, USA: ACM, 2010. — P. 61–70.DOI: 10.1145/1772690.1772698.

[8] Chaabane A., Acs G., Kaafar M. You Are What You Like! Information Leakage Through Users' Interests. Proc. Annual Network and Distributed System Security Symposium. — San Diego, United States, 2012.

[9] Kosinski M., Stillwell D., Graepel T. Private traits and attributes are predictable from digital records of human behavior. Proceedings of the National Academy of Sciences. — 2013. — Vol. 110, no. 15. — P. 5802–5805.DOI: 10.1073/pnas.1218772110.

[10] Dougnon R. Y., Fournier-Viger P., Nkambou R. Inferring User Profiles in Online Social Networks Using a Partial Social Graph. Advances in Artificial Intelligence: 28th Canadian Conference on Artificial Intelligence, Canadian AI 2015, Halifax, Nova Scotia, Canada, June 25, 2015, Proceedings. — Cham: Springer International Publishing, 2015. — P. 84–99.DOI: 10.1007/978-3-319-18356-5_8.

[11] Inferring User Demographics and Social Strategies in Mobile Social Networks / Y. Dong [et al.]. Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. — New York, New York, USA: ACM, 2014. — P. 15–24.DOI: 10.1145/2623330.2623703.

[12] Analiz socialnikh setei: metody i prilozaniya / A.V. Korshunov et al.. Trudi Instatuta sistemnogo programmirovaniya RAN. -2014. – v.26. – issue 1. – P. 357-394.

[13] Cheng Z., Caverlee J., Lee K. You Are Where You Tweet: A Contentbased Approach to Geolocating Twitter Users. Proceedings of the 19th ACM International Conference on Information and Knowledge Management. — Toronto, ON, Canada: ACM, 2010. — P. 759–768. DOI: 10.1145/1871437.1871535.

[14] Discriminating gender on Twitter / J. D. Burger [et al.]. Proceedings of the Conference on Empirical Methods in Natural Language Processing. — Association for Computational Linguistics. 2011. — P. 1301–1309.

[15] Ikawa Y., Enoki M., Tatsubori M. Location inference using microblog messages.. WWW (Companion Volume). — ACM. - 2012. — P. 687–690. DOI: 10.1145/2187980.2188181.

[16] Towards Social User Profiling: Unified and Discriminative Influence Model for Inferring Home Locations / R. Li [et al.]. Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. — Beijing, China: ACM, 2012. — P. 1023–1031. DOI: 10.1145/2339530.2339692.

[17] You Are Who You Know: Inferring User Profiles in Online Social Networks / A. Mislove [et al.]. Proceedings of the Third ACM International Conference on Web Search and Data Mining. — New York, New York, USA: ACM, 2010. — P. 251–260. DOI: 10.1145/1718487.1718519.

[18] Li R., Wang C., Chang K. C.C. User Profiling in an Ego Network: Coprofiling Attributes and Relationships. Proceedings of the 23rd International Conference on World Wide Web. — Seoul, Korea: ACM, 2014. — P. 819–830.DOI: 10.1145/2566486.2568045.

[19] McAuley J., Leskovec J. Discovering Social Circles in Ego Networks. ACM Trans. Knowl. Discov. Data. — 2014. — Vol. 8, no. 1. — P. 1–28.DOI: 10.1145/2556612.

[20] Chesnokov V. Overlapping Community Detection in Social Networks with Node Attributes by Neighborhood Influence. Models, Algorithms, and Technologies for Network Analysis: NET 2016, Nizhny Novgorod, Russia, May 2016 / ed. by V. A. Kalyagin [et al.]. — Cham: Springer International Publishing, 2017. — P. 187–203. DOI: 10.1007/978-3-319-56829-4_14.

[21] Chesnokov V. Application of the community allocation algorithm in the information confrontation in the social networks. Voprosy kiberbezopasnosti *[Cybersecurity issues]*, 2017, N 1(19), pp**.** 37-44. DOI: 10.21681/2311-3456-2017-1-37-44.

[22] Leskovec J., Krevl A. SNAP Datasets: Stanford Large Network Dataset Collection. — June 2014. — http://snap.stanford.edu/data.

[23] Clauset A., Newman M. E. J., Moore C. Finding community structure in very large networks. Phys. Rev. E. — 2004. — Vol. 70, issue 6. — P. 1–6. DOI: 10.1103/PhysRevE.70.066111

[24] Rosvall M., Bergstrom C. T. Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences. — 2008. — Vol. 105, no. 4. — P. 1118–1123. DOI: 10.1073/pnas.0706851105

[25] Yang J., Leskovec J. Community Affiliation Graph Model for Overlapping Network Community Detection. 12th IEEE International Conference on Data Mining, ICDM 2012, Brussels, Belgium, December 1013, 2012. — 2012. — P. 1170–1175. DOI: 10.1109/ICDM.2012.139

[26] Yang J., Leskovec J. Overlapping Community Detection at Scale: A Nonnegative Matrix Factorization Approach. Proceedings of the Sixth ACM International Conference on Web Search and Data Mining. — Rome, Italy: ACM, 2013. — P. 587–596. DOI: 10.1145/2433396.2433471

[27] Yang J., McAuley J. J., Leskovec J. Community Detection in Networks with Node Attributes. 2013 IEEE 13th International Conference on Data Mining. – 2013. – P. 1151-1156. DOI: 10.1109/ICDM.2013.167