International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-3, Issue-8)

Research Article | **August 2014**

Obviation of Secluded Information on Social Networks from Inference Assails

G. Sateesh[1]
Asst. Prof., Department of CSE,
SVCET, Chittoor, India

Bhanu Prakash Ande[2]
M.Tech, II Year, Department of CSE
SVCET, Chittoor, India

Abstract-

In today's social networks, such as facebook, twitter and linkedIn are utilized by so many people. By using these networks people are publishing their details about themselves and they can communicate with their friends. Some of the information posted in these networks is meant to be private. But that private information can be seen by some other people this is the problem. So in this we explore how to overcome this problem by launching inference attacks using released social network data to predict confidential private information about individuals. We then arrange some possible sanitization techniques that could be used in various situations. Then, we explore the effectiveness of these techniques by studying Social network analysis, data mining, social network privacy and implementing them on a data set. We show that we can decrease the effectiveness of both local and relational classification algorithms by using the sanitization methods we described.

Keywords— online social network study, data mining, social network security study

I. INTRODUCTION

Online social network applications that allow their users to connect by means of various link types. As part of their offerings, these networks allow people to list details about themselves that are relevant to the nature of the network. For instance, Facebook is a general-use social network, so individual users list their favorite activities, books, and movies. Conversely, LinkedIn is a professional network; because of this, users specify details which are related to their professional life (i.e., reference letters, previous employment, and so on.) Because these sites gather extensive personal information, social network application providers have a rare opportunity: direct use of this information could be useful to advertisers for direct marketing.

However, in practice, privacy concerns can prevent these efforts. This conflict between the desired use of data and individual privacy presents an opportunity for privacy-preserving social network data mining—that is, the discovery of information and relationships from social network data without violating privacy. Privacy concerns of individuals in a social network can be classified into two categories: privacy after data release, and private information leakage. Private information leakage, conversely, is related to details about an individual that are not explicitly stated, but, rather, are inferred through other details released and/or relationships to individuals who may express that detail.

A trivial example of this type of information leakage is a scenario where a user, say John, does not enter his political affiliation because of privacy concerns. We model an attack scenario as follows: Suppose Facebook wishes to release data to electronic arts for their use in advertising games to interested people. We explore how the online social network data could be used to predict some individual private detail that a user is not willing to disclose and explore the effect of possible data sanitization approaches on preventing such private information leakage, while allowing the recipient of the sanitized data to do inference on nonprivate details. This problem of private information leakage could be an important issue in some cases. Recently, both ABC News and the Boston Globe published reports indicating that it is possible to determine a user's sexual orientation by obtaining a relatively small subgraph from Facebook that includes only the user's gender, the gender they are interested in, and their friends in that subgraph. Predicting an individual's sexual orientation or some other personal detail may seem like inconsequential, but in some cases, it may create negative repercussions. For example, using the disclosed social network, predicting an individual's likelihood of getting Alzheimer disease for health insurance and employment purposes could be problematic.

1.1 Our Contributions

To the best of our knowledge, this is the first paper that discusses the problem of sanitizing a social network to prevent inference of social network data and then examines the effectiveness of those approaches on a real-world data set. In

order to protect privacy, we sanitize both details and the underlying link structure of the graph. That is, we delete some information from a user's profile and remove some links between friends. We also examine the effects of generalizing detail values to more generic values. We then study the effect these methods have on combating possible inference attacks and how they may be used to guide sanitization. We further show that this sanitization still allows the use of other data in the system for further tasks. In addition, we discuss the notion of "perfect privacy" in social networks and give a formal privacy definition that is applicable to inference attacks discussed in this paper.

II. LITERATURE SURVEY

In this paper, we touch on many areas of research that have been heavily studied. The area of privacy inside a social network encompasses a large breadth, based on how privacy is defined. In, Zheleva and Getoor propose several methods of social graph anonymization, focusing mainly on the idea that by anonymizing both the nodes in the group and the link structure, that one thereby anonymizes the graph as a whole. However, their methods all focus on anonymity in the structure itself. For example, through the use of k anonymity or t-closeness, depending on the quasi-identifiers which are chosen, much of the uniqueness in the data may be lost. Through our method of anonymity preservation, we maintain the full uniqueness in each node, which allows more information in the data postrelease. In, Backstrom et al. consider an attack against an anonymized network. In their model, the network consists of only nodes and edges. Detail values are not included. The goal of the attacker is simply to identify people. Further, their problem is very different than the one considered in this paper because they ignore details and do not consider the effect of the existence of details on privacy. Hay et al. and Liu and Terzi consider several ways of anonymizing social networks. However, our work focuses on inferring details from nodes in the network, not individually identifying individuals. Other papers have tried to infer private information inside social networks.

In, Gross et al. examine specific usage instances at Carnegie Mellon. They also note potential attacks, such as node reidentification or stalking, that easily accessible data on Facebook could assist with. They further note that while privacy controls may exist on the user's end of the social networking site, many individuals do not take advantage of this tool. This inding coincides very well with the amount of data that we were able to crawl using a very simple crawler on a Facebook network. We extend on their work by experimentally examining the accuracy of some types of the demographic reidentification that they propose before and after sanitization. The Facebook platform's data has been considered in some other research as well. In, Jones and Soltren crawl Facebook's data and analyze usage trends among Facebook users, employing both profile postings and survey information. However, their paper focuses mostly on faults inside the Facebook platform. They do not discuss attempting to learn unrevealed details of Facebook users, and do no analysis of the details of Facebook users. Their crawl consisted of around 70,000 Facebook accounts. The area of link-based classification is well studied. However, their comparisons do not consider In , Tasker et al. present an alternative classification method where they build on Markov networks. However, none of these papers consider ways to combat their classification methods. In, Menon and Elkan use dyadic data methods to predict class labels. We show later that while we do not examine the effects of this type of analysis, the choice of technique is arbitrary for anonymization and utility.

In, Zheleva and Getoor attempt to predict the private attributes of users in four real-world data sets: Facebook, Flickr, Dogster, and BibSonomy. They do not attempt to actually anonymize or sanitize any graph data. Instead, their focus is on how specific types of data, namely, that of declared and inferred group membership, may be used as a way to boost local and relational classification accuracy. Theydefined method of group-based (as opposed to details-based or link-based) classification is an inherent part of our details-based classification, as we treat the group membership data as another detail, as we do favorite books or movies. In fact, Zheleva and Getoor work provides a substantial motivation for the need of the solution proposed in our work. In, Talukder et al. propose a method of measuring the amount of information that a user reveals to the outside world and which automatically determines which information (on a per-user basis) should be removed to increase the privacy of an individual. Finally, in, we do preliminary work on the effectiveness of our Details, Links, and Average classifiers and examine their effectiveness after removing some details from the graph.

III. LEARNING METHODS ON SOCIAL NETWORKS

3.1 Social Network Description

Definition 1. A social network is represented as a graph, G ={V, E,D}, where V is the set of nodes in the graph, where each node ni represents a unique user of the social network. E represents the set of edges in the graph, which are the links defined in the social network. For any friendship link Fi;j between user ni and user nj, we assume that both Fi;j ∈ E and Fj;i2 E. D is the set of details from the social network.

Definition 2. A detail type is a string defined over an alphabet _ that represents a specific category name within the social network details set. The set of all detail types is represented by H. A detail value is a string defined over an alphabet _ that represents a user's input for a detail type. A detail is a (detail type, detail value) pair, represented uniquely by an identifier Jk. Dj i is the jth (detail type, detail value) pair specified by the user ni. Di is the set of all Dj i for a node ni. D is the set of Di for all i. It is important to note that for any detail type, the expected response can either be single or multivalued, and that a user has the option of listing no detail values for any given detail. For example, consider Facebook's "home town" and "activities" detail type. A user can only have one home town, but can list multiple activities (for instance, soccer, reading, video games).

We further define a set of private details $\mathcal{I}$, where any detail is private if for any $h_m \in \mathcal{H}$, $h_m \in \mathcal{I}$. Consider the following illustrative example:

$$\mathcal{H} = \{\text{favorite books, favorite movies, political affiliation, religion, activities}\}, \quad (1)$$

$$\mathcal{I} = \{\text{political affiliation, religion}\}, \quad (2)$$

$$n_1 = (\text{Jane Doe}), \quad (3)$$

$$n_2 = (\text{John Smith}), \quad (4)$$

$$D_2 = \{\text{John Smith's Details}\}, \quad (5)$$

$$D_2^4 = (activities,\ fishing), \quad (6)$$

$$\mathcal{N}_1 = \{n_2\}, \quad (7)$$

$$F_{1,2}, F_{2,1} \in \mathcal{E}. \quad (8)$$

That is, from among four possible detail types (1), we define two detail types to be private, a person's political affiliation and their religion (2). Then, say we have two people, named Jane Doe and John Smith, respectively, (3) and (4). John Smith has specified that one of the activities he enjoys is fishing (6), which is also recorded as the fourth possible (detail type, detail value) pair. Also, John and Jane are friends (7).

Obviously, the detail types of I are varied based on an individual's choice. Generally, however, we consider a user's I to be any details that they do not specify. For experimentation, however, we choose to use

$$\mathcal{I} = \{\text{political affiliation, sexual orientation.}\}$$

We use these detail types as our C in all classification methods. Further, for political affiliation, we consider only C_{lib} and C_{cons} as possible class values—that is, "liberal" and"conservative." For sexual orientation, we consider $C_{heterosexual}$ and $C_{homosexual}$ as the possible class values.To evaluate the effect that changing a person's details has on their privacy, we needed to first create a learning method that could predict a person's private details (for the sake of example, we assume that political affiliation is unspecified for some subset of our population). Since our goal is to understand the feasibility of possible inference attacks and the effectiveness of various sanitization techni-ques combating against those attacks, we initially used a simple naı¨ve Bayes classifier

3.2.Naı¨ve Bayes Classification

Determining an individual's political affiliation is an exercise in graph classification.

Given a node n_i with m details and p potential classifica-tion labels, C_1; . . . ; C_p, C_x^i, the probability of n_i being in class C_x, is given by the equation

$$\underset{1\le x\le p}{\operatorname{argmax}} \left[P\left(C_x^i \middle| D_i^1, \ldots, D_i^m\right)\right],$$

where arg max1 <= x<= p represents the possible class label that maximizes the previous equation. However, this is difficult to calculate, since P ðC$_x^i$Þ for any given value of x is unknown. By applying Bayes' theorem, we have the equation

$$\underset{1\le x\le p}{\operatorname{argmax}} \left[\frac{P\left(C_x^i\right) \times P\left(D_i^1, \ldots, D_i^m | C_x^i\right)}{P\left(D_i^1, \ldots, D_i^m\right)}\right].$$

Further, by assuming that all details are independent, we are left with the simplified equation

$$\underset{1\le x\le p}{\operatorname{argmax}} \left[P\left(C_x^i\right) \times P\left(D_i^1 | C_x^i\right) \times \cdots \times P\left(D_i^m | C_x^i\right)\right], \quad (9)$$

3.3. Naı¨ve Bayes on Friendship Links

Consider the problem of determining the class detail value of person n_i given their friendship links using a naı¨ve Bayes model. That is, of calculating P ðC$_x^i$jN $_i$Þ. Because there are relatively few people in the training set that have a friendship link to n_i, the calculations for P ðC$_x^i$jF$_{i;j}$Þ become extremely inaccurate. Instead, we choose to decompose this relationship. Rather than having a link from person n_i to n_j, we instead consider the probability of having a link from n_i to someone with n_j's details. Thus,

$$P(C_x^i|F_{i,j}) \approx P(C_x^i|L_1, L_2, \ldots, L_m) \approx \frac{P(C_x^i) \times P(L_1|C_x^i) \times \cdots \times P(L_m|C_x^i)}{P(L_1, \ldots, L_m)}, \quad (10)$$

where L_n represents a link to someone with detail J_n.

3.4. Weighing Friendships

There is one last step to calculating P ðC$_x^i$jN $_i$Þ. In the specific case of social networks, two friends can be anything from acquaintances to close friends or family members. While there are many ways to weigh friendship links, the method we used is very easy to calculate and is based on the assumption that the more public details two people share, the more private details they are likely to share. This gives the following formula for $W_{i;j}$, which represents the weight of a friendship link from n_i to node n_j:

$$W_{i,j} = \frac{\left|\left(D_i^1, \ldots, D_i^n\right) \cap \left(D_j^1, \ldots, D_j^m\right)\right|}{|D_i|}. \quad (11)$$

Equation (11) calculates the total number of details n_i and n_j share divided by the number of details of n_i.

Note that the weight of a friendship link is not the same for both people on each side of a friendship link. In other words, $W_{j;i}/W_{i;i}$. The final formula for person i becomes the following, where Z represents a normalization factor and calculated by $P(C_x^i|F_{i,j})$

$$\rho(C_x^i, \mathcal{N}_i) = \frac{1}{Z} \sum_{n_j \in \mathcal{N}_i} \left[P(C_x^i|F_{i,j}) \times W_{i,j}\right]. \quad (12)$$

The value $\rho(C_x^i, \mathcal{N}_i)$ is used as our approximation to $P(C_x^i|\mathcal{N}_i)$

3.5. Network Classification

Collective inference is a method of classifying social network data using a combination of node details and connecting links in the social graph. Each of these classifiers consists of three components: a local classifier, a relational classifier, and a collective inference algorithm.

3.5.1 Local Classifiers

Local classifiers are a type of learning method that are applied in the initial step of collective inference. Typically, it is a classification technique that examines details of a node and constructs a classification scheme based on the details that it finds there. For instance, the naı¨ve Bayes classifier we discussed previously is a standard example of Bayes classification. This classifier builds a model based on the details of nodes in the training set. It then applies this model to nodes in the testing set to classify them.

3.5.2 Collective Inference Methods

Unfortunately, there are issues with each of the methods described above. Local classifiers consider only the details of the node it is classifying. Conversely, relational classifiers consider only the link structure of a node. Specifically, a major problem with relational classifiers is that while we may cleverly divide fully labeled test sets so that we ensure every node is connected to at least one node in the training set, real-world data may not satisfy this strict requirement. If this requirement is not met, then relational classification will be unable to classify nodes which have no neighbors in the training set.

IV. HIDING PRIVATE INFORMATION

Existing privacy definitions such as k-anonymity , l-diversity , and so on are defined for relational data only. They provide syntactic guarantees and do not try to protect against inference attacks directly. For example, k-anonym-ity tries to make sure that an individual cannot be identified from the data but does not consider inference attacks that can be launched to infer private information.

Recently developed differential privacy definition provides interesting theoretical guarantees. Basically, it guarantees that the result of a differential private algorithm are very similar with or without the data of any single user. In other words, differentially privacy guarantees that the change in one record, does not change the result too much. On the other hand, this definition does not protect against the building of an accurate data mining model that can predict sensitive information. Actually many differentially private data mining algorithms have been developed that has similar accuracy to nondifferentially private versions. Since our goal is to release rich social network data set while preventing sensitive detail disclosure through data mining techniques, differential privacy defini-tion is not directly applicable in our scenario.

4.1 Formal Privacy Definition

Problem 1. Given a graph, G, from a social network, where I is a subset of H, and */I/>1*, is it possible to minimize the classification accuracy on I when using some set of classifiers C while preserving the utility of H - I?

Definition 3. Background knowledge, K, is some data that is not necessarily directly related to the social network, but that can be obtained through various means by an attacker.

Examples of background knowledge in terms of a social network such as Facebook include voter registration, election results, phone book listings, and so on.

4.2 Manipulating Details

Clearly, details can be manipulated in three ways: adding details to nodes, modifying existing details and removing details from nodes. However, we can broadly classify these three methods into two categories: perturbation and anonymization. Adding and modifying details can both be considered methods of perturbation—that is, introducing various types of "noise" into D to decrease classification accuracies. Removing nodes, however, can be considered an anonymization method.

Definition 7. DVD is a process by which an attribute is divided into a series of representative tags. These tags do not necessarily reassemble into a unique match to the original attribute.

We provide a general outline of the generalization process in Algorithm 1. At each step, we generalize each detail type by one level [Lines 3-5] by determining which attributes can be further generalized without complete removal and keep a list of the accuracy of this generalization. At the end of each round, we "permanently" store the individual detail type that provides the greatest privacy savings [Line 4].

Algorithm 1. Generalize($\Delta, \mathcal{G}$)

1: $\mathcal{G}' \leftarrow \mathcal{G}$
2: **while** Classify($\mathcal{G}$) - Classify($\mathcal{G}'$) $\leq \Delta$ **do**
3: $S \leftarrow$ all details that can be further generalized
4: $s \leftarrow$ getHighestInfoGainAttrib(S)
5: Gen($s, \mathcal{G}'$)
6: **end while**
7: return $\mathcal{G}'$

TABLE 2
General Information about the Data

Diameter of the largest component	16
Number of nodes in the graph	167,390
Number of friendship links in the graph	3,342,009
Total number of listed details in the graph	4,493,436
Total number of unique details in the graph	110,407
Number of components in the graph	18

V. EXPERIMENTS

5.1. Data Gathering

We wrote a program to crawl the Facebook network to gather data for our experiments. Written in Java 1.6, the crawler loaded a profile, parsed the details out of the HTML, and stored the details inside a MySQL database. Then, the crawler loaded all friends of the current profile and stored the friends inside the database both as friend-ship links and as possible profiles to later crawl.

Because of the sheer size of Facebook's social network, the crawler was limited to only crawling profiles inside the Dallas/Forth Worth (DFW) network. This means that if two people share a common friend that is outside the DFW network, this is not reflected inside the database. Also, some people have enabled privacy restrictions on their profile which prevented the crawler from seeing their profile details. The total time for the crawl was seven days.

5.2. Experimental Setup

In our experiments, we implemented four algorithms to predict the political affiliation of each user. The first algorithm is called "Details Only." This algorithm uses (9) to predict political affiliation and ignores friendship links. The second algorithm is called "Links Only." This algorithm uses (12) to predict political affiliation using friendship links and does not consider the details of a person. The third algorithm is called "Average." The Average algorithm predicts a node's class value based on the following equation:

TABLE 3
Baseline Probabilities

Probability of being Liberal	0.45
Probability of being Conservative	0.55
Probability of being Heterosexual	0.95
Probability of being Homosexual	0.05

$$P_A(C_a^i) = 0.5 * P_D(C_a^i) + 0.5 * P_L(C_a^i),$$

where P_D and P_L are the numerical probabilities assigned by the Details Only and Links Only algorithms, respec-tively. The final algorithm is a traditional naı¨ve Bayes classifier, which we used as a basis of comparison for our proposed algorithms.

5.3 Local Classification Results

Tables 4 and 5 list the most liberal or conservative details. For example, the most liberal detail, as shown in Table 4, is being a member of the group "legalize same sex marriage." Tables 6 and 7 show the most liberal and conservative details for each detail type.

TABLE 5
A Sample of the Most Conservative Detail Values

Detail Name	Detail Value	Likelihood
group member	george w bush is my homeboy	45.88831329
group member	bears for bush	30.86484689
group member	kerry is a fairy	28.50250433
favorite movies	end of the spear	14.53703765

TABLE 6
Most Conservative Detail Value for Each Detail

Detail Name	Detail Value	Likelihood
activities	college republicans	5.846955271
favorite books	redeeming love	6.348153362
favorite movies	end of the spear	14.53703765
favorite music	delirious	18.85227471
favorite tv shows	fox news	7.753312932
grad school	sw seminary	2.749648395
group member	george w bush is my homeboy	45.88831329
interests	hunting and fishing	7.614995442
relationship status	married	1.667495517
religious views	christian	2.441063037
sex	male	1.087798286

"likelihood" value for a given detail value, J_k, is calculated as

$$\frac{P(J_k|C_i)}{P(C_i)} \times 100.$$

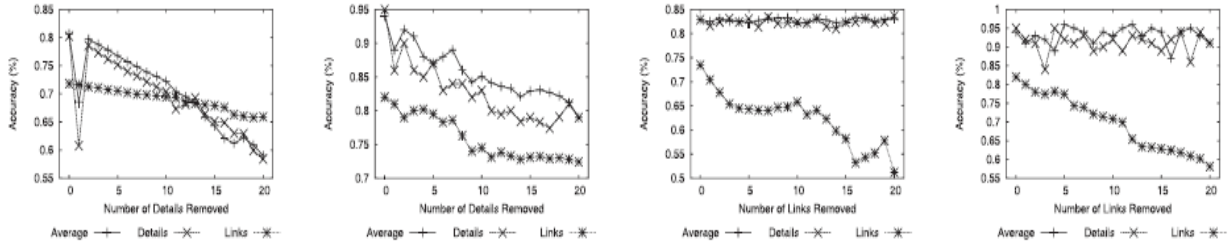
In Table 8, we show the details that most indicate the "homosexual" classification. In contrast to political affiliation, there are no single details which are very highly correlated with that classification. For example, the three details we have selected here are more highly indicative of being "Liberal" than of being "homosexual." Conversely, we see in Table 9 that there are a few categories that are very highly representative of the "heterosexual" classification.

5.3.1 Detail Removal

As can be seen from the results, our methods are generally successful at reducing the accuracy of classification tasks. Fig. 1 shows that removing the details most highly connected with a class is accurate across the details and average classifiers. Counter-intuitively, perhaps, is that the accuracy of our links classifier is also decreased as we remove details. However, as discussed in Section 4.4, the details of two nodes are compared to find a similarity. As we remove details from the network, the set of "similar" nodes to any given node will also change. This can account for the decrease in accuracy of the links classifier.

5.3.2 Link Removal

As seen in Figs. 1c and 1d, when we remove links, we have a generally more stable downward trend, with only a few exceptions in the "political affiliation" experiments.



(a) Political Affiliation (b) Sexual Orientation (c) Political Affiliation (d) Sexual Orientation

Fig. 1 Local classification accuracy

However, we show that by applying our technique, we routinely restrict classification accuracy to some arbitrary value below 95 percent. As we mentioned in Section 4.1, this means that graph is effectively ð_; C; G; KÞ-private because an attacker would be forced to use only K to determine classification labels.

5.4 Generalization Experiments

Each detail falls into one of several categories: religion, political affiliation, activities, books, music, quotations, shows/movies, and groups. Due to the lack of a reliable subject authority, that is, a source who could definitively categorize a given quotation without additional human input, quotations were discarded from all experiments. To generate the DGH for each activity, book, and show/movie, we used Google directories. To generate the DVD for Music, we used the Last.fm tagging system. To generate the hierarchy for Groups, we used the classification criteria from the Facebook page of that group.

5.5 Effect of Sanitization on Other Attack Techniques

We further test the removal of details as an anonymization technique by using a variety of different classification algorithms to test the effectiveness of our method. For each number of details removed, we began by removing the indicated number of details in accordance with the method as described in Section 4. We then performed tenfold cross validation on this set 100 times, and conduct this for 0-20 details removed. The results of these tests are shown in Figs. 6a and 6b. As can be seen from these figures, our technique is effective at reducing the classification of networks for those details which we have classified as sensitive.

While the specific accuracy reduction is varied by the number of details removed and by the specific algorithm used for classification, we see that we do in fact reduce the accuracy across a broad range of classifiers. We see that linear regression is affected the least, with approximately a 10 percent reduction in accuracy. Also that decision trees are affected the most, with a roughly 35 percent reduction in classification accuracy.

VI. CONCLUSION

We addressed various issues related to private information leakage in social networks. We show that using both friendship links and details together gives better predict- ability than details alone. In addition, we explored the effect of removing details and links in preventing sensitive information leakage. In the process, we discovered situations in which collective inferencing does not improve on using a simple local classification method to identify nodes. When we combine the results from the collective inference implications with the individual results, we begin to see that removing details and friendship links together is the best way to reduce classifier accuracy. This is probably infeasible in maintaining the use of social networks. However, we also show that by removing only details, we greatly reduce the accuracy of local classifiers, which give us the maximum accuracy that we were able to achieve through any combination of classifiers. We also assumed full use of the graph information when deciding which details to hide. Useful research could be done on how individuals with limited access to the network could pick which details to hide. Similarly, future work could be conducted in identifying key nodes of the graph structure to see if removing or altering these nodes can decrease information leakage.

REFERENCES

[1] Facebook Beacon, 2009.
[2] M. Hey, G. Miklau, D. Jensen, p. Weis, and S. Srivastava, “Anonymizing Social Networks”, Technical report 07-19, Univesity of Massachusetts Amherst,2007.
[3] K. Liu and E. Terzi, “Towards Identity Anonymization on Graphs”, Proc. ACM SIGMOD 2008.
[4] J. He, W. Chu, and V. Liu, “Inferring Privacy Information from Social Networks”, Proc. Inteligence and Security Informatics”,2006.
[5] L. Backstrom, C. Dwork, and J. Kleinberg, “Wherefore Art Thour3579x?: Anonymized Social Networks, Hidden Patterns, and Structural Steganography,” Proc. 16th Int’l Conf. World Wide Web (WWW ’07), pp. 181-190, 2007.
[6] H. Jones and J.H. Soltren, “Facebook: Threats to Privacy,” technical report, Massachusetts Inst. of Technology, 2005.
[7] A. Menon and C. Elkan, “Predicting Labels for Dyadic Data,” Data Mining and Knowledge Discovery, vol. 21, pp. 27-343, 2010.
[8] C. Dwork, “Differential Privacy,” Automata, Languages and Programming, M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener,eds., vol. 4052, pp. 1-12, Springer, 2006.
[9] A. Friedman and A. Schuster, “Data Mining with Differential Privacy,” Proc. 16th ACM SIGKDD Int’l Conf. nowledge Discovery and Data Mining, pp. 493-502, 2010.

