Estimating Network Structure from Anonymous Ego-centric Information

Takayasu Fushimi[1], Kazumi Saito[1], and Kazuhiro Kazama[2]

[1] Graduate School of Management and Information of Innovation, University of Shizuoka
52-1 Yada, Suruga-ku, Shizuoka 422-8526, Japan
takayasu.fushimi@gmail.com, k-saito@u-shizuoka-ken.ac.jp

[2] Faculty of Systems Engineering, Wakayama University
Sakaedani 930, Wakayama-city, Wakayama 640-8510, Japan
kazama@ingrid.org

Abstract. We address a problem of estimating the whole structure of an actual social network of people from only their two types of anonymous ego-centric information, personal attributes like sex and relational ones like the numbers of female and male friends, obtained as answers to questionnaires in a social survey. From these attribute values, we can obtain the degree of each node, which corresponds to the number of friends of each person, together with some macroscopic information about the network, like the ratio of links between female and male nodes to the total number of links, as the mixing matrices. However, we cannot directly know the actual connections between two nodes only from these observed mixing matrices. Thus, we propose a new method for estimating the whole structure of the hidden network by minimizing the Kullback-Leibler divergence between each pair of the observed and estimated mixing matrices, under the constraints with respect to the degree of each node. In our experiments using three types of networks, we show that the proposed method can produce much better estimation results, in comparison to a random baseline which is assigned arbitrary links under the degree constraints, especially for the cases of highly assortative, where each node has a tendency to connect to nodes with the same attribute values.

1 Introduction

In recent years, there have been many studies on complex networks which analyze their essential structures and/or phenomena happened over them (Leskovec et al. 2007; Wu et al. 2008). It has been widely known that many complex networks have common characteristics such as a small-world nature (Watts et al. 1998) and a scale-free nature, and many social networks have an assortative/homophily nature (Newman et al. 2003) which means a tendency that lots of linked pairs of nodes are likely to have same or similar properties. Furthermore, in order to investigate real world phenomena such as information diffusion and opinion formation over social networks, the Independent Cascade model, the

Y.S. Kim et al (Eds.): PKAW 2014, LNCS 8863, pp. 236–245, 2014.

Linear Threshold model (Kempe et al. 2003; Saito et al. 2009), and the voter model (Even-Dar et al. 2007) have been studied, under an implicit assumption that the whole structure of a target network is already given. In real world problems, however, it is not always easy to obtain a whole network structure because of the privacy issue, even for on-line social networks mainly due to a rate limitation of accessing data resources. In such cases, we can obtain partial information about the target network. By collecting some statistical or macroscopic information about a network, we need to estimate the whole structure as accurately as possible. Even for an estimated network, we can know some important nodes such as influential ones (Kimura et al. 2010) with a reasonable accuracy by using the above-mentioned models for information diffusion or opinion formation, if the precision of the estimated network is sufficient high.

In this paper, basically supposing that a social survey provides the source ego-centric information, we propose a method of estimating the whole structure from only anonymous ego-centric information. The ego-centric information consists of two types of information, personal attributes like sex and relational ones like the numbers of female and male friends. By collecting and aggregating these pieces of information, we can obtain some macroscopic information like the ratio of links between female and male nodes to the total number of links, as the mixing matrices. However, we cannot directly know the actual connections between two nodes only from these observed mixing matrices. Thus, we propose a new method for estimating the whole structure of the hidden network by minimizing the Kullback-Leibler divergence between each pair of the observed and estimated mixing matrices.

As existing studies on estimating network structure or predicting missing links, a variety of techniques have been proposed such as topological scores (Nowell et al. 2003), feature-based classification (Oyama et al. 2004), kernel-based methods (Kato et al. 2005), hierarchical property (Clauset et al. 2008), and matrix factorization (Menon et al. 2011). These methods attempt to learn link tendencies from observed linked pairs and predict the link existence or non-existence for the remaining node pairs. On the other hand, our method attempts to estimate all links from only ego-centric information, therefore these existing methods and our method differ in the problem settings.

In our experiments, in order to evaluate the precision of the estimated network obtained by our proposed method we utilize three networks each of whose link structure is completely known. For these networks, we assign artificial attributes to all the nodes, where the associated values of some attributes are determined by using the voter model, so that we can quantitatively evaluate the relation between assortativity of networks and estimation precisions.

This paper is organized as follows: after formalizing our problem framework in Section 2, we describe a detail of our proposed method in Section 3. Then, by using three networks, we compare the estimation precisions and representative network statistics with those of a random baseline in Section 4. Finally, we describe our conclusion in Section 5.

2 Problem Framework

Formally, we regard the N answerers of a social survey as nodes and define the node set as V. Let K be the number of categorical attributes we focus on. Let $S^{(k)}$ be the number of categories of the k-th attribute, and we assign each categorical attribute value to an integer ranging from 1 to $S^{(k)}$, i.e., $\{1, \ldots, S^{(k)}\}$. For each node u, we consider the K-dimensional attribute vector $\mathbf{f}_u = (f_u^{(1)}, \ldots, f_u^{(K)})$ whose k-th element is $f_u^{(k)} \in \{1, \ldots, S^{(k)}\}$.

Now, for each pair of node u and attribute k, we also consider the $S^{(k)}$-dimensional vector $\mathbf{g}_u^{(k)}$ whose t-th element $g_u^{(k)}(t)$ corresponds to the number of u's friends with the categorical value t for this attribute, where note that $t \in \{1, \ldots, S^{(k)}\}$ and $d_u = \sum_{t=1}^{S^{(k)}} g_u^{(k)}(t)$. Then, for the k-th attribute, we can define the following mixing matrix element $m_{s,t}^{(k)}$ with respect to a pair of categorical values s and t:

$$m_{s,t}^{(k)} = \frac{1}{L} \sum_{\{u : f_u^{(k)} = s\}} g_u^{(k)}(t) \qquad k = 1, \ldots, K$$

where L stands for the total number of links, i.e., $L \sum_{u \in V} d_u$. Thus, we can obtain the observed mixing matrix $\mathbf{M}^{(k)} = [m_{s,t}^{(k)}]$ for each attribute k. From each mixing matrix, we can calculate the assortative coefficient according to the Newman's method (Newman 2003), so that we can discriminate whether a certain attribute is assortative or not.

In order to derive our objective function, we construct an $N \times S^{(k)}$ projection matrix for the k-th attribute as follows:

$$\mathbf{W}^{(k)} = [w_{u,s}^{(k)}] = \begin{cases} 1 & \text{if} \quad f_u^{(k)} = s \\ 0 & \text{otherwise} \end{cases}$$

Now, we suppose that an adjacency matrix $\mathbf{A} = [a_{u,v}]$ is given where $a_{u,v}$ means the (u, v)-th element, i.e., $a_{u,v} \in \{0, 1\}$. Then, for a given projection matrix $\mathbf{W}^{(k)}$ of the k-th attribute, we can calculate an $S^{(k)} \times S^{(k)}$ estimated mixing matrix from an adjacency matrix $\mathbf{A}$, denoted by $\hat{\mathbf{M}}^{(k)}(\mathbf{A}) = [\hat{m}_{s,t}^{(k)}(\mathbf{A})]$, as follows:

$$\hat{m}_{s,t}^{(k)}(\mathbf{A}) = \frac{1}{L} \sum_{u \in V} \sum_{v \in V \setminus \{u\}} a_{u,v} w_{u,s}^{(k)} w_{v,t}^{(k)}.$$

Then we can define the optimal adjacency matrix $\hat{\mathbf{A}}$ which minimizes the KL divergences between the observed and estimated mixing matrices, $\mathbf{M}^{(k)}$ and $\hat{\mathbf{M}}^{(k)}(\mathbf{A})$, as follows:

$$\hat{\mathbf{A}} = \arg\min_{\mathbf{A}} \left\{ \sum_{k=1}^{K} \mathrm{KL}\left(\mathbf{M}^{(k)} || \hat{\mathbf{M}}^{(k)}(\mathbf{A}) \right) \right\}$$

under the condition that $\sum_{v \in V \setminus \{u\}} a_{u,v} = d_u$ for each node $u \in V$. Note that $\mathrm{KL}(P||Q)$ means the KL divergence between probabilistic distributions P and

Q. The calculation of the optimal adjacency matrix $\hat{\mathbf{A}}$ can be transformed as follows:

$$\begin{aligned}\hat{\mathbf{A}} &= \arg\min_{\mathbf{A}} \left\{ \sum_{k=1}^{K} \sum_{s=1}^{S^{(k)}} \sum_{t=1}^{S^{(k)}} m_{s,t}^{(k)} \log \frac{m_{s,t}^{(k)}}{\hat{m}_{s,t}^{(k)}(\mathbf{A})} \right\} \\ &= \arg\max_{\mathbf{A}} \left\{ \sum_{k=1}^{K} \sum_{s=1}^{S^{(k)}} \sum_{t=1}^{S^{(k)}} m_{s,t}^{(k)} \log \hat{m}_{s,t}^{(k)}(\mathbf{A}) \right\} \\ &= \arg\max_{\mathbf{A}} \left\{ \sum_{k=1}^{K} \sum_{s=1}^{S^{(k)}} \sum_{t=1}^{S^{(k)}} m_{s,t}^{(k)} \log \sum_{u \in V} \sum_{v \in V \setminus \{u\}} a_{u,v} w_{u,s}^{(k)} w_{v,t}^{(k)} \right\}\end{aligned}$$

Therefore, we can solve our problem by finding the optimal $\mathbf{A}$ which maximizes the following objective function:

$$J(\mathbf{A}) = \sum_{k=1}^{K} \sum_{s=1}^{S^{(k)}} \sum_{t=1}^{S^{(k)}} m_{s,t}^{(k)} \log \sum_{u \in V} \sum_{v \in V \setminus \{u\}} a_{u,v} w_{u,s}^{(k)} w_{v,t}^{(k)}, \tag{1}$$

under the condition that $\sum_{v \in V \setminus \{u\}} a_{u,v} = d_u$ for each node $u \in V$.

3 Solution Algorithm

In order to maximize the objective function Eq.(1), we employ the EM algorithm by defining the following posterior probability on the link between nodes u and v.

$$\bar{q}_{s,t,u,v}^{(k)} = \frac{\bar{a}_{u,v} w_{u,s}^{(k)} w_{v,t}^{(k)}}{\sum_{x \in V} \sum_{y \in V} \bar{a}_{x,y} w_{x,s}^{(k)} w_{y,t}^{(k)}}$$

where $\bar{a}_{u,v}$ means the current value of (u, v)-th element of the adjacency matrix. Then, we define the expectation value of log-likelihood of complete data on posterior probability referred to as Q function as follows:

$$Q\left(\mathbf{A}|\bar{\mathbf{A}}\right) = \sum_{k=1}^{K} \sum_{s=1}^{S^{(k)}} \sum_{t=1}^{S^{(k)}} m_{s,t}^{(k)} \sum_{u \in V} \sum_{v \in V \setminus \{u\}} \bar{q}_{s,t,u,v}^{(k)} \log a_{u,v}.$$

In our problem framework, due to the degree constraints, we add the following Lagrange multiplier terms.

$$\tilde{Q}\left(\mathbf{A}|\bar{\mathbf{A}}\right) = Q\left(\mathbf{A}|\bar{\mathbf{A}}\right) + \sum_{u \in V} \lambda_u \left(d_u - \sum_{v \in V \setminus \{u\}} a_{u,v} \right)$$

where λ_u means the Lagrange multiplier on node u.

Although the actual value of each adjacency matrix element is $a_{u,v} \in \{0, 1\}$, we relax our problem by allowing each value to be a real number $a_{u,v} \in [0, 1]$. and we refer to this matrix as a relaxed adjacency matrix.

The proposed method estimates the $\tilde{Q}$ function as the E step of the EM algorithm, and maximizes the $\tilde{Q}$ function by updating the relaxed adjacency matrix as the M step as follows:

$$a_{u,v} = \frac{d_u \sum_{k=1}^{K} \sum_{s=1}^{S^{(k)}} \sum_{t=1}^{S^{(k)}} m_{s,t}^{(k)} \bar{q}_{s,t,u,v}^{(k)}}{\sum_{k=1}^{K} \sum_{s=1}^{S^{(k)}} \sum_{t=1}^{S^{(k)}} m_{s,t}^{(k)} \sum_{v \in V \setminus \{u\}} \bar{q}_{s,t,u,v}^{(k)}}.$$

The proposed method repeats the E and M steps until convergence. Furthermore, since this relaxation problem is a convex optimization problem over a convex set, it is guaranteed that we can always obtain the unique global optimal solution from any initialized values. Let $\mathbf{A}_t$ be the relaxed adjacency matrix at t-th iteration of our algorithm. For a small value ε, we define the convergence criterion as $|(J(\mathbf{A}_t) - J(\mathbf{A}_{t-1}))/J(\mathbf{A}_{t-1})| < \varepsilon$. Finally, our proposed method selects the largest top-d_u values from $\{\hat{a}_{u,v} : v \in V\}$ for each u. and constructs an estimated link set $\hat{E}$. Then, the method outputs the estimated link set $\hat{E}$.

Consequently given the node set V and K categorical attributes, we propose a method to estimate all the links between actual nodes of their hidden social network by the following steps:

STEP1 For each node u, extract K categorical attribute values and then construct K-dimensional attribute vector $\mathbf{f}_u$. For each user u, extract the number of friends d_u and its breakdown by values of attribute k, and then construct observed mixing matrix $\mathbf{M}^{(k)}$;

STEP2 Estimate a relaxed adjacency matrix $\hat{\mathbf{A}} = [\hat{a}_{u,v}]$ by minimizing the KL divergence between each pair of the observed and estimated mixing matrices, and select the largest top-d_u values from $\{\hat{a}_{u,v} : v \in V\}$ for each u, then outputs an estimated link set $\hat{E}$;

4 Experimental Evaluation

Our goal is estimating the whole structure using only ego-centric information obtained from a social survey. However, in this paper, in order to quantitatively evaluate the precision of estimated network obtained by our proposed method, we utilize three networks each of whose link structure is completely known. Furthermore, it is naturally considered that the estimation precision depends on the numbers of attributes and categories, and the levels of assortative coefficients. Then, we evaluate the relations between estimation precisions and these conditions of utilizing attributes. Therefore, at the STEP1 of our method, as substitute for extracting K attributes from a social survey, we assign K artificial attributes by the method described in the subsection 4.2.

4.1 Network Data

We describe a detail of three networks used in our experiments.

First one is a synthetic network with a hierarchical property, just like employee relationships. In this network, we can assume two types of nodes, central and peripheral nodes. The central nodes are characterized by relatively high degree and low clustering coefficients, while the peripheral nodes by relatively low degree and high clustering coefficients. We generated this network according to Ravasz et al. (Ravasz et al. 2003). This network has 125 nodes and 410 links. Hereafter, this network is referred to as Hierarchical network.

Second one is a hyperlink network of a Japanese university Web site, where we obtained this network by crawling the Web site as of August 2010. This network has 600 nodes and 1,299 links. Hereafter, this network is referred to as Web network [2].

Last one is a co-author network of the international conference NIPS (Neural Information Processing Systems). We define an author as a node and add a link between two nodes who have at least one joint paper. There exist some researcher communities of similar research topics. This network has 1,036 nodes and 2,044 links. Hereafter, this network is referred to as Nips network.

4.2 Artificial Attributes

We describe the method for assigning attribute values to each node via the voter model. The voter model is one of the most basic stochastic process models, which simulates opinion formation processes over networks. Let $h_t(u)$ be the opinion value of node u at time t and $\Gamma(u)$ be the set of parent/adjacent nodes of u. For a given a network $G = (V, E)$, we assign categorical values of an attribute k for each node as follows:

VM1 Initialize $t \leftarrow 1$ and assign the initial opinion value as $h_0(u) \leftarrow s$ by selecting an atrribute value $s \in \{1, \dots, S^{(k)}\}$ uniformly at random for each node $u \in V$;

VM2 Update the opinion value as $h_t(u) \leftarrow h_{t-1}(v)$ by selecting a parent node $v \in \Gamma(u)$ uniformly at random for each node $u \in V$;

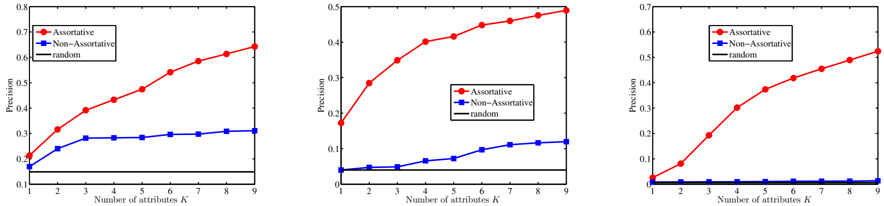
VM3 Terminate if the assortative coefficient $r > 0.8$, otherwise update $t \leftarrow t+1$ and go back to VM2;

In our experiments, this algorithm terminated within 10 iterations for all the trials.

4.3 Experimental Settings

In order to evaluate how assortativity of attributes on a network affects to the performance of our proposed method, we prepare two types of sets of K attributes, assortative ones and non-assortative ones. By repeating the trials of

[2] The site name and its address are "Faculty of Computer and Information Sciences, Hosei University" and `http://cis.k.hosei.ac.jp/`, respectively.



(a) Hierarchical network (b) Web network (c) Nips network

Fig. 1. Precision with respect to the number of attributes

the voter model, we obtain the assortative attribute for each pair of node u and attribute k, $f_u^{(k)} \leftarrow h_T(u)$, where T means the final step of our voter model process. On the other hand, by simply assigning an initial random attribute value, we obtain the non-assortative attribute for each pair of node u and attribute k, $f_u^{(k)} \leftarrow h_0(u)$. Hereafter, for a convenience, our proposed method is reffered to as the assortative method in the case that a series of assortative attributes are available, while it is reffered to as the non-assortative method in the case that only a series of non-assortative attributes are available.

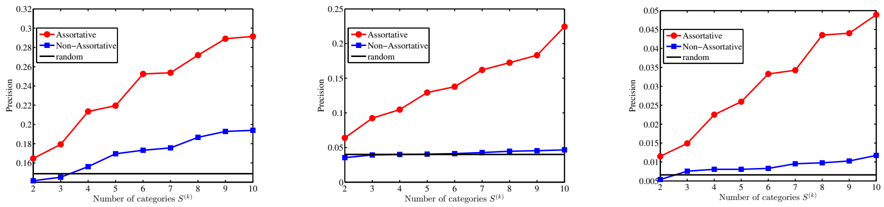
We set each initial value of the relaxed adjacency matrix element to $a_{u,v} = d_u/N$, and the convergence criterion of our algorithm is specified by $\varepsilon = 10^{-4}$. Then, we estimate the whole structure of each network by our proposed method using a set of attributes, and evaluate the performance as the average precision of 10 estimation results obtained by 10 sets of attributes generated independently.

4.4 Estimation Precisions with Respect to the Number of Attributes

In order to evaluate our proposed method, we employ a precision measure which has been widely used in the information retrival field. More specifically, we define the sets of the true and estimated links as E and $\hat{E}$, respectively, and calculate the precision of $\hat{E}$ to E as $F(E, \hat{E}) = |E \cap \hat{E}|/|E|$.

Fig. 1 shows the precisions of the assortative and non-assortative methods with respect to the number of attributes K ranging from 1 to 9, where the number of categories $S^{(k)}$ is set to 5 for each attribute. In these figures, we plot the estimation precisions obtained by the assortative method, the non-assortative method and random baseline. with a line with filled circles, a line with filled squares, and a simple line, respectively. Here, the random baseline is the performance obtained by arbitrary assigning links under the degree constraints.

From Fig. 1, we can observe the following characteristics. First, as expected, for each network, the performance of the assortative method streadly improved when the number of attributes increases in comparison to the random baseline. This fact suggests that the proposed method generally works well by using an enough number of attributes with high levels of assortataive coefficients. Here we should emphasize that a person has a tendency to connect to people with



(a) Hierarchical network (b) Web network (c) Nips network

Fig. 2. Precision with respect to the number of categories

the same attribute value in general. Second, unlike the assortative method, the performance of the non-assortative method did not improve sufficiently even for a case of large number of attributes. Especially, the performance improvement was almost nothing for the Nips network with the largest size. This fact indicates that our proposed method works well only for networks with a relatively small size in case that only a series of non-assirtative attributes are available.

4.5 Estimation Precisions with Respect to the Number of Categories

We evaluate relations between precisions and the number of categories of attributes used in estimation. Fig. 2 shows the precisions of the assortative and non-assortative methods with respect to the number of categories $S^{(k)}$ ranging from 2 to 10 and the number of attributes K is set to 1 for each result of estimation.

From Fig. 2, we can observe the following characteristics. First, similar to the previous experimental results as shown in Fig. 1, for each network, the performance of the assortative method streadly improved in comparison to the random baseline, when the number of categories increases. This fact also suggests that the proposed method generally works well by using even only one assortative attribute if it has large numbers of specific categories. Second, unlike the assortative method, the performance of the non-assortative method did not improve sufficiently even for a case of large number of categories, except for the small scale Hierarchical network. This fact also indicates that our proposed method works well only for networks with a relatively small size in case that only one non-assirtative attribute is available. Therefore, when combining our results observed in Fig. 1 and 2, by preparing an enough number of assortative attributes, each of which has a relatively large number of specified categories, we can expect that our proposed method is promising to estimating the whole structure of networks from only anonymous ego-centric information.

4.6 Estimation Performances in Terms of Network Statistics

We evaluate the structure of the obtained networks in terms of network statistics. To this end, we focus on the average clustering coefficient $\bar{C}$, the standard

Table 1. Statistics of true, proposed and random networks

(a) Hierarchical network

	$\bar{C}$	$C_{\mathrm{S.D.}}$	$\bar{D}$
true	0.837	0.164	2.128
proposed	0.623	0.211	2.826
random	0.063	0.087	2.499

(b) Web network

	$\bar{C}$	$C_{\mathrm{S.D.}}$	$\bar{D}$
true	0.540	0.413	4.219
proposed	0.340	0.390	14.433
random	0.004	0.031	4.650

(c) Nips network

	$\bar{C}$	$C_{\mathrm{S.D.}}$	$\bar{D}$
true	0.582	0.427	6.213
proposed	0.420	0.394	15.923
random	0.001	0.013	5.060

deviation of the clustering coefficient of each node $C_{S.D.}$, and the average shortest path length $\bar{D}$ as the representative network statistics. Note that the average of shortest path length is calculated by harmonic average to cope with disconnected networks. In this experiment, we used 9 assortative attributes each of which consists of 5 categories ($K = 9$, $S^{(k)} = 5$).

Table 1 compares the statistics of the true networks, the estimated networks by the assortative method, and random ones. From Table 1, we can see the following similar characteristic. First, the average values of clustering coefficient $\bar{C}$ of the estimated networks by the assortative method are significantly close to those of the true networks, in comparison to those of random networks. We can also see that the standard deviations of clustering coefficients $C_{\mathrm{S.D.}}$ of the estimated networks are also reasonably close to true ones. Existing assortative attributes means that the neighborhood nodes in the network are also likely to have same attribute values. Thus, our method can estimate the local relationships and deviation of clustering coefficients more accurately.

Moreover, in case of the relatively small networks like the Hierarchical network, we observe that the average shortest path length $\bar{D}$ of the estimated network is also close to that of the true network. In contrast, in each case of the relatively large networks such as the Nips and Web networks, we observe that the average shortest path length of the estimated network is not close to that of the true network. This is because our proposed method occasionally produced disconnected network in the case of large scale netowrks. On the other hand, the random network shows the small shortest path length because of random links like small-world model (Watts et al. 1998). From these results, we conjecture that our proposed method can estimate more accurately the local structure like communities.

5 Conclusion

We addressed the problem of estimating the whole structure of a network. In this paper, we proposed a method of estimating the whole structure from only anonymous ego-centric information by minimizing the Kullback-Leibler divergence between each pair of the observed and estimated mixing matrices.

In our experiments using several types of synthetic and real networks, we evaluated the estimation precisions with respect to the numbers of attributes and categories and assortativity of attributes. From our experimental results, our proposed method generally works well by using an enough number of attributes with high levels of assortataive coefficients, and the method can estimate more accurately the local structure like communities.

In future, we plan to evaluate the performance of our proposed method by using several types of attributes such as non-assortative but non-random ones, numerical ones like age, and so forth.

Acknowledgments. This work was partly supported by Asian Office of Aerospace Research and Development, Air Force Office of Scientific Research under Grant No. AOARD-13-4042, and JSPS Grant-in-Aid for Scientific Research(No. 25.10411).

References

Clauset, A., Moore, C., Newman, M.E.J.: Hierarchical structure and the prediction of missing links in networks. Nature 453, 98–101 (2008)

Even-Dar, E., Shapira, A.: A note on maximizing the spread of influence in social networks. In: Deng, X., Graham, F.C. (eds.) WINE 2007. LNCS, vol. 4858, pp. 281–286. Springer, Heidelberg (2007)

Kato, T., Tsuda, K., Asai, K.: Selective integration of multiple biological data for supervised network inference. Bioinformatics 21(10), 2488–2495 (2005)

Kempe, D., Kleinberg, J., Tardos, E.: Maximizing the spread of influence through a social network. In: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2003, pp. 137–146. ACM, New York (2003)

Kimura, M., Saito, K., Nakano, R., Motoda, H.: Extracting influential nodes on a social network for information diffusion. Data Min. Knowl. Discov. 20(1), 70–97 (2010)

Leskovec, J., Adamic, L.A., Huberman, B.A.: The dynamics of viral marketing. ACM Trans. Web 1(1), 5 (2007)

Menon, A.K., Elkan, C.: Link prediction via matrix factorization. In: Gunopulos, D., Hofmann, T., Malerba, D., Vazirgiannis, M. (eds.) ECML PKDD 2011, Part II. LNCS (LNAI), vol. 6912, pp. 437–452. Springer, Heidelberg (2011)

Newman, M.E.J.: Mixing patterns in networks. Physical Review E 67(2), 026126+ (2003)

Nowell, D.L., Kleinberg, J.: The link prediction problem for social networks. In: CIKM 2003: Proceedings of the Twelfth International Conference on Information and Knowledge Management, pp. 556–559. ACM, New York (2003)

Oyama, S., Manning, C.D.: Using feature Conjunctions across examples for learning pairwise classifiers. In: Boulicaut, J.-F., Esposito, F., Giannotti, F., Pedreschi, D. (eds.) ECML 2004. LNCS (LNAI), vol. 3201, pp. 322–333. Springer, Heidelberg (2004)

Ravasz, E., Barabási, A.L.: Hierarchical organization in complex networks. Physical Review E 67(2), 026112+ (2003)

Saito, K., Kimura, M., Ohara, K., Motoda, H.: Learning continuous-time information diffusion model for social behavioral data analysis. In: Zhou, Z.-H., Washio, T. (eds.) ACML 2009. LNCS (LNAI), vol. 5828, pp. 322–337. Springer, Heidelberg (2009)

Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature 393(6684), 440–442 (1998)

Wu, F., Huberman, B.A.: How public opinion forms. In: Papadimitriou, C., Zhang, S. (eds.) WINE 2008. LNCS, vol. 5385, pp. 334–341. Springer, Heidelberg (2008)