

An Estimation Method for Relationship Strength in Weighted Social Network Graphs

Xiang Lin, Tao Shang, Jianwei Liu

School of Electronic and Information Engineering, Beihang University, Beijing, China
Email: shangtao@buaa.edu.cn

Received December 2013

Abstract

Previous works mainly focused on estimating direct relationship strength in social networks. If two users are not directly connected in a social network, there is no direct relationship. In order to estimate the relationship strength between two indirectly connected users as well as directly connected users, this paper proposes an estimation method for relationship strength in weighted social network graphs, which is based on the trust propagation strategy and the estimation of direct relationship strength. Our method considers the length of a relationship path, the number of relationship paths and the edge weights (direct relationship strength) along with a relationship path to estimate the strength of indirect relationship. Then it synthesizes the direct and indirect relationship strength to represent the strength of relationship between two users in social networks. Thus our method can fully estimate the relationship strength between any two users in a social network no matter whether they are directly connected or not.

Keywords

Social Networks; Relationship Strength; Estimation

1. Introduction

With the rapid development of internet services, social network (SN) becomes a most popular service where people interact with each other. Social network is a set of social actors and the relationships among them [1]. The relationship strength among users is different and can rapidly vary. Information on relationship strength between two users is useful in many areas such as link prediction, item recommendation, newsfeed, people search and so on. According to the estimation of relationship strength, the social network provider can improve the quality of some social network services.

Recently, several++ works have focused on computing the relationship strength in social networks. Interaction data has been used to predict the relationship strength [2,3], but this work only considered two levels of relationship strength, namely, weak and strong relationships. Viswanath *et al.* [4] analyzed the relationships between users based on the Facebook's "wall posts", which ignored other interaction data. Xiang *et al.* [5] proposed an unsupervised latent variable model for the estimation of relationship strength based on interaction activity and user's similarity. It uses a richer representation that can span the full spectrum from weak to strong ties. Srba *et al.* [6] proposed a method to calculate the relationship strength by means of the interaction data and

other “rate factors”. Yanagimoto *et al.* [7] proposed a relationship strength estimation method in social media. It estimates relationship strength between web pages in social bookmarking services using a tag vocabulary.

However, nearly all these methods focus on estimating the strength of direct relationship while ignoring indirect relationship between two users in social networks. For example, Alice and Bob have a common friend though they are not friends. How to estimate the relationship strength between those users is a key problem and is also very useful in social networks.

In this work, we mainly focus on indirect relationship strength as well as direct relationship strength. We combine the estimation method of relationship intensity strength to estimate the direct relationship strength between two users. Firstly, the indirect relationship strength is estimated according to trust propagation strategy. Secondly, the indirect relationship strength between two users is estimated according to the length of relationship paths, the number of relationship paths and the edge weights along with the relationship paths. Finally, the direct and indirect relationship strength is synthesized to represent the relationship strength between these two users. Our method can fully estimate the relationship strength between any two users in a social network no matter whether they are directly connected or not.

2. Related Work

2.1. Estimation of Direct Relationship Strength

Srba *et al.* [6] proposed a method to calculate the relationship intensity strength. This method denotes elementary interaction between two users (e.g., sending a message) or static common information (e.g., common hobby) as a “rate factor”. Every rate factor influences the strength of a relationship in a positive or negative way depending on the social aspect. And the rate factor from different sources (social networks) has a different importance, which is represented to be a weight. The final relationship strength is also influenced by the count of instances of the rate factor.

Partial relationship intensity depends on the weight, the count of instances of the rate factor and time:

$$I_f(k, j) = \frac{w_{kj} \sum_{i=1}^l f_i}{1 + \ln(1 + l_c)}, \quad (1)$$

where I_f is the partial relationship intensity for one rate factor, w_{kj} is the weight of the rate factor j for source k , l is the count of instances of the rate factor in the relationship of two traced users, l_c is the count of instances of the rate factor, and f_i is the function expressing time influence.

The final relationship intensity is calculated as the arithmetic average of the partial relationship intensity of all sources. This method can estimate the strength of direct relationship between two users, but it ignores the indirect relationship between the users who are not directly connected in the social network.

2.2. Trust Propagation Strategy

Until now, few of previous works have focused on indirect relationship, but there are a lot of works on indirect trust computation. Although relationship is a different concept from trust, the propagation of indirect relationship between two indirectly connected users is similar to the propagation of indirect trust.

Nasir *et al.* [8] introduced a shortest-path min-max strategy to compute the indirect trust between a source user s and a target user t . A shortest path is defined as the path with the minimum number of edges. Let $SP_{s,t}$ be the set of all shortest paths between s and t , the strongest shortest paths $\pi_{s,t}^{SST}$ are the shortest paths with maximum strength.

$$\pi_{s,t}^{SST} \in \left\{ \pi \in SP_{s,t} \mid str(\pi) = \max_{\rho \in SP_{s,t}} str(\rho) \right\} \quad (2)$$

Let $in_{s,t}^s$ be the reachable in-neighbors of t from s with the shortest distance from s among all the reachable in-neighbors of t . Then the most reliable in-neighbors of t for s in the shortest paths strategy $in_{s,t}^{smax}$ are the nodes belonging to $in_{s,t}^s$ having a maximum strength strongest shortest path from s .

$$in_{s,t}^{smax} \in \left\{ v \in in_{s,t}^s \mid str(\pi_{s,v}^{SST}) = \max_{x \in in_{s,t}^s} str(\pi_{s,x}^{SST}) \right\} \quad (3)$$

The final trust from s to t is calculated as

$$Trust(s,t) = \max_{v \in \Pi_{s,t}^{max}} \left(\min \left(str \left(\pi_{s,v}^{SST} \right), w(v,t) \right) \right). \quad (4)$$

This method can compute the trust value from a source user s to a target user t , but it is incomplete because it uses the minimum weight edge of a path to represent the path strength.

3. Estimation Method for Relationship Strength

We focus on estimating the relationship strength between two indirectly connected users as well as directly connected users.

3.1. Relationship Strength

We define relationship strength as the closeness between two users in social networks. Direct relationship strength (computed by users' interaction data, similarity and so on) describes the closeness between two directly connected users. If user A could find user B along with the edges in a social network graph, there exists a relationship path between them. Indirect relationship strength describes the closeness between two users who are not directly connected, and it is calculated by the relationship paths between them. In addition, two users might have direct and indirect relationship at the same time. So we should synthesize both cases to estimate users' relationship strength. The relationship types are illustrated in **Figure 1**.

Let $RS(v_i, v_j)$, $RS_d(v_i, v_j)$ and $RS_{id}(v_i, v_j)$ denote the synthetic relationship strength, direct relationship strength and indirect relationship strength between two individuals v_i and v_j . Thus

$$RS(v_i, v_j) = \alpha RS_d(v_i, v_j) + \beta RS_{id}(v_i, v_j), \quad (5)$$

where α and β denote the weight coefficient of direct relationship strength and indirect relationship strength. In addition, α and β satisfy the equation as follows:

$$\alpha + \beta = 1 \quad \alpha, \beta > 0. \quad (6)$$

Let $G = \{V, E, W\}$ denote a weighted social network graph. $V = \{v_1, v_2, \dots, v_n\}$ is the user set of the social network graph where v_i denotes the i th user and $|V| = n$. $E = \{e_{i,j}\}$ is the edge set of the social network graph where $e_{i,j}$ denotes the edge between v_i and v_j . And the number of edges is denoted as $|E| = m$. $W = \{w_{i,j}\}$ is the weight set of edges in the social network graph where $w_{i,j}$ denotes the weight of edge $e_{i,j}$, and the value of an edge weight varies continuously from 0 to 1. A simple social network graph is illustrated in **Figure 2**.

3.2. Estimation Method

We propose an estimation method for relationship strength in weighted social network graphs, which is based on the estimation of direct relationship strength and the trust propagation strategy, just as described in the following part:

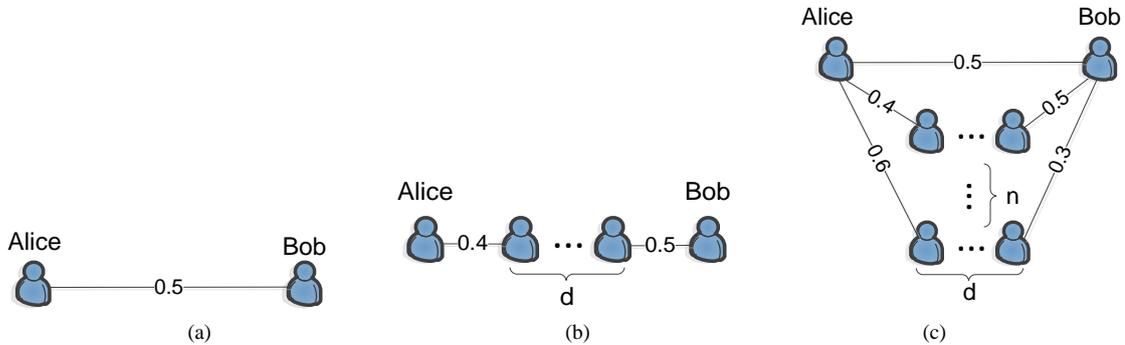


Figure 1. Relationship type. (a) Direct relationship; (b) Indirect relationship; (c) Synthetic relationship.

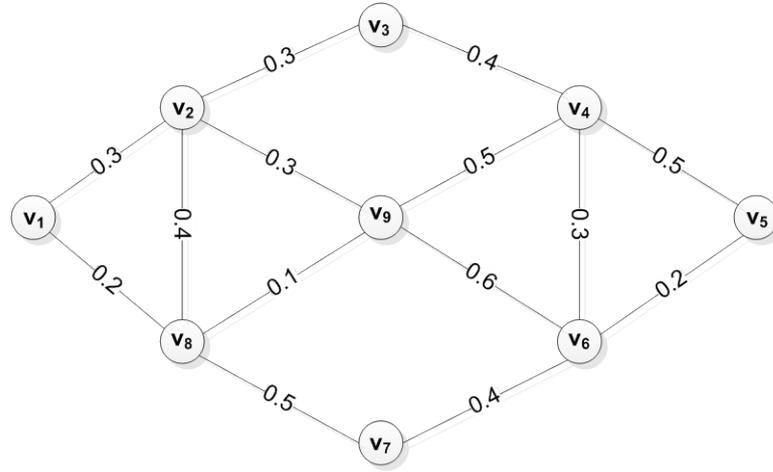


Figure 2. A simple social network graph.

1) The direct relationship strength is estimated by the estimation method of relationship intensity strength. And it is represented by the edge weight in the social network graph.

$$RS_d(v_i, v_j) = \frac{\sum_{k=1}^n \sum_{j=1}^m I_f(k, j)}{n} = w_{i,j} \quad (7)$$

2) The indirect relationship strength is decided by the length of a relationship path, the number of relationship paths and the edge weights of relationship paths. The number and the edge weights of relationship paths have a positive correlation with the indirect relationship strength, while the length of a relationship path has a negative effect.

Assuming that $P_{i,j}$ is a relationship path between v_i and v_j , the indirect relationship strength is described as follows:

$$RS_{id}(v_i, v_j) = e^{-\lambda d} \prod_{j=1}^d w_j, \quad (8)$$

where λ denotes the attenuation coefficient of the length of a relationship path, d denotes the length of the relationship strength, w_j denotes the j th edge weight of a relationship path.

Here $e^{-\lambda d}$ is an attenuation function and its value varies continuously from 0 to 1, which is illustrated in Figure 3. The function represents the weight coefficient of a relationship path among all the paths.

In general, there are more than one relationship path between two users. As the different length of a relationship path has different influence on the relationship strength, we should not treat all paths as equivalent influence. Therefore, we take the weighted average rather than the arithmetic average to compute indirect relationship strength of all paths.

Assuming that there are n relationship paths between v_i and v_j :

$$\{P_1, P_2, \dots, P_n\},$$

with weights

$$\{e^{-\lambda d_1}, e^{-\lambda d_2}, \dots, e^{-\lambda d_n}\},$$

thus the indirect relationship strength of all paths is described as follows:

$$RS_{id}(v_i, v_j) = \frac{\sum_{i=1}^n [e^{-\lambda d_i} P_i]}{\sum_{i=1}^n e^{-\lambda d_i}} = \frac{\sum_{i=1}^n \left[e^{-\lambda d_i} \prod_{j=1}^{d_i} w_j \right]}{\sum_{i=1}^n e^{-\lambda d_i}}, \quad (9)$$

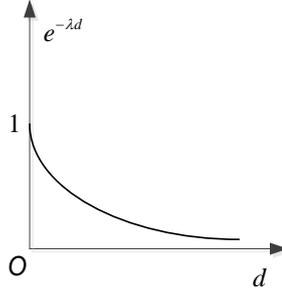


Figure 3. Function of $e^{-\lambda d}$.

where P_i denotes the i th relationship path of v_i and v_j , d_i denotes the length of P_i .

3) According to the direct and indirect relationship strength, the synthetic relationship strength is finally calculated as follows:

$$RS(v_i, v_j) = \alpha RS_d(v_i, v_j) + \beta RS_{id}(v_i, v_j) = \alpha \cdot w_{i,j} + \beta \cdot \frac{\sum_{i=1}^n \left[e^{-\lambda d_i} \prod_{j=1}^{d_i} w_j \right]}{\sum_{i=1}^n e^{-\lambda d_i}}. \quad (10)$$

3.3. Property

Theorem 1 The values of $RS(v_i, v_j)$, $RS_d(v_i, v_j)$ and $RS_{id}(v_i, v_j)$ are all in the interval of $[0, 1]$.

Proof

For $w_{i,j}$ is an edge weight between v_i and v_j and it is a continuous value in the interval of $[0, 1]$, we can get that the value of $RS_d(v_i, v_j)$ is in the interval of $[0, 1]$ because it equals to $w_{i,j}$. And the value of P_i is in the interval of $[0, 1]$ because it is achieved by multiplying all the edge weights along with the relationship path. The value of $RS_{id}(v_i, v_j)$ is also in the interval of $[0, 1]$ because it is estimated by weighted average of all relationship paths. For $\alpha + \beta = 1$ and $\alpha, \beta > 0$, the value of $RS(v_i, v_j)$ is in the interval of $[0, 1]$ because it is the summation of $\alpha RS_d(v_i, v_j)$ and $\beta RS_{id}(v_i, v_j)$.

Theorem 2 The continuous value of relationship strength from 0 to 1 can span the full spectrum from weak to strong relationship strength.

Proof

According to **Theorem 1**, we can know that the values of $RS(v_i, v_j)$, $RS_d(v_i, v_j)$ and $RS_{id}(v_i, v_j)$ representing the strength of these three relationship types are all in the continuous interval of $[0, 1]$. And if the value of relationship strength is close to 1, it represents stronger relationship strength. Meanwhile, if it is close to 0, it represents weaker relationship strength. Because there are countless values in the continuous interval of $[0, 1]$, the strength of direct relationship, indirect relationship and synthetic relationship represented by these values can span the full spectrum from weak to strong relationship strength.

4. Search Algorithm of Relationship Paths

We should search all the paths from a source user s to a target user t before estimating indirect relationship strength. However, there are too many paths if the social network is big enough and it results in high time overhead. Therefore, we use a modified breadth first search (BFS) algorithm to search the shortest paths from s to the neighbors of t . **Algorithm 1** shows the search algorithm. These shortest paths and the edges between t and its neighbors make up the relationship paths between s and t . The relationship paths between s and t are illustrated in **Figure 4**.

Let $SP_{s,t}$ be the set of all shortest paths between s and t . And $in_{s,t}^t$ is the set of all neighbors of t . The all relationship paths are denoted as $\pi_{s,t}^{SST}$.

$$\pi_{s,t}^{SST} \in \{SP_{s,v} \parallel e_{v,t}, v \in in_{s,t}^t\} \quad (11)$$

5. Experiments

In order to verify the effectiveness of the proposed method, we performed several experiments on a simple social network (**Figure 5**). First, we choose a source user s . Second, we estimate the relationship strength between the source user s and the other users with two methods, including the method of relationship intensity strength [6] and our estimation method of relationship strength.

We estimate the relationship strength between s and the other users based on the estimation method of relationship intensity strength, and then represent it with edge weight in the weighted social network graphs.

Before evaluating our method on estimating relationship strength, we set $\lambda = 0$, $\alpha = 0.6$ and $\beta = 0.4$. If two users have direct relationship strength besides indirect relationship strength, the direct relationship strength should be more important than indirect relationship strength. It agrees with common sense.

At first we use the edge weight, estimated by Equation (7), to represent the direct relationship strength between two users. Secondly, we search the all relationship paths from a source user s to the other users in the network. And then we estimate the indirect relationship strength using Equation (9). After that, we synthesize the direct relationship strength and indirect relationship strength by Equation (10). The comparison of these two methods is shown in **Table 1**.

Algorithm 1. Search algorithm of relationship paths.

Input: s : source node; t : target node

Output: $\pi_{s,t}^{SS}$

- 1) $\hat{in}_{s,t}^t \leftarrow$ find the neighbors of t
- 2) for each $v \in \hat{in}_{s,t}^t$
- 3) $SP_{s,t} \leftarrow BFS(s,v)$
- 4) $\pi_{s,t}^{SS} \leftarrow SP_{s,v} \parallel e_{v,t}$
- 5) return $\pi_{s,t}^{SS}$

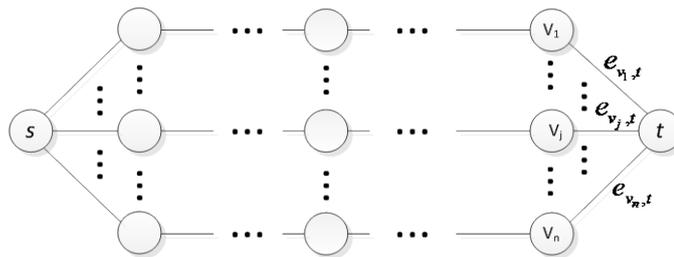


Figure 4. Relationship paths between s and t .

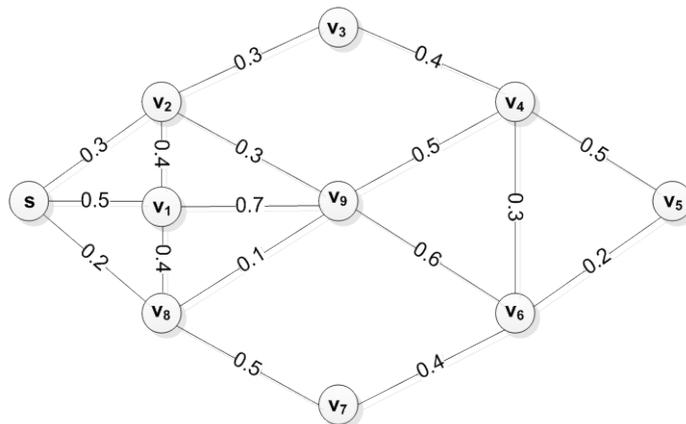


Figure 5. A simple weighted social network graph.

Table 1. Comparison of our method and reference [6].

Relationship strength Target user	Source user	s	
		Our Method	Method of [6]
v_1		0.328	0.500
v_2		0.207	0.300
v_3		0.018	None
v_4		0.013	None
v_5		0.010	None
v_6		0.014	None
v_7		0.021	None
v_8		0.144	0.200
v_9		0.040	None

As shown in **Table 1**, v_1 , v_2 and v_8 have the most strong relationship strength with the source user s in both methods. The value of relationship strength in our method is smaller than that of relationship intensity strength because our method considers the indirect relationship strength as well as direct relationship strength. The indirect relationship strength will decrease the synthetic relationship strength. We can also see that the method in that of relationship intensity strength cannot estimate the indirect relationship strength (e.g., relationship strength between s and v_3) though it can estimate the direct relationship strength. Since the indirect relationship is much more common than the direct relationship in a social network, this method has a limitation on estimating relationship strength from a source user to a target user. Therefore, our method is more comprehensive on the estimation of relationship strength in a social network.

6. Conclusion

In this paper, we proposed an estimation method for relationship strength in weighted social network graphs. This method focuses on estimating the indirect relationship strength as well as direct relationship strength. The indirect relationship strength of one relationship path is estimated by multiplying all the edge weights in the path. And the final indirect relationship strength is a weighted average of all relationship paths. Compared to other methods on estimating relationship strength, our method is much more comprehensive.

Acknowledgements

The authors would like to thank the National Basic Research Program of China (No. 2012CB315905), the National Natural Science Foundation of China (No. 61272501, 4132056, 61173154), and the Beijing Natural Science Foundation (No. 4132056) for valuable helps.

References

- [1] Hanneman, R.A. and Riddle, M. (2005) Introduction to Social Network Methods. <http://faculty.ucr.edu/~hanneman/nettext/>
- [2] Kahanda, I. and Neville, J. (2009) Using Transactional Information to Predict Link Strength in Online Social Networks. *Proceedings of the Third International ICWSM Conference*, 74-81.
- [3] Gilbert, E. and Karahalios, K. (2009) Predicting Tie Strength with Social Media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 211-220.
- [4] Viswanath, B., Mislove, A., Cha, M., Gummadi, K.P. (2009) On the Evolution of User Interaction in Facebook. *Proceedings of the 2nd ACM workshop on Online Social Networks*, 37-42.
- [5] Xiang, R., Neville, J. and Rogati, M. (2010) Modeling Relationship Strength in Online Social Networks. *Proceedings of the 19th International Conference on World Wide Web*, 981-990. <http://dx.doi.org/10.1145/1772690.1772790>

- [6] Srba, I. and Bielikova, M. (2010) Tracing Strength of Relationships in Social Networks. 2010 *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, Toronto, 31 August-3 September 2010, 13-16.
- [7] Yanagimoto, H. and Yoshioka, M. (2012) Relationship Strength Estimation for Social Media Using Folksonomy and Network Analysis. *WCCI 2012 IEEE World Congress on Computational Intelligence*, Brisbane, 10-15 June 2012, 1-8.
- [8] Nasir, S.U. and Kim, T.H. (2013) Fast Trust Computation in Online Social Networks. *IEICE Transactions on Communications*, **E96-B**, 2774-2783.