

Measuring tie strength in multidimensional networks

Luca Pappalardo, Giulio Rossetti, and Dino Pedreschi

Department of Computer Science, University of Pisa, Italy
`{lpappalardo,rossetti,pedre}@di.unipi.it`

Discussion Paper
original paper is [10]

Abstract. Online social networks have allowed us to build massive networks of *weak* ties: acquaintances and nonintimate ties we use all the time to spread information and thoughts. Conversely, strong ties are people we really trust, persons most like us and whose social circles tightly overlap with our own. Unfortunately, social media do not incorporate tie strength in the creation and management of relationships, and treat all users the same: friend or stranger, with little or nothing in between. In the current work, we address the challenging issue of detecting on online social networks the strong and intimate ties from the huge mass of such mere social contacts. In order to do so, we propose a novel multidimensional definition of tie strength which exploits the existence of multiple online social links between two individuals. We test our definition on a multidimensional network constructed over users in Foursquare, Twitter and Facebook, analyzing the structural role of strong e weak links, and the correlations with the most common similarity measures.

Keywords: Social network analysis, Multidimensional networks, Tie strength

1 Introduction

In the last decade, social networking has completely redefined the way we conceive our social relationships, creating the sensation of having broken the constraints of time and geography that limited people's social world. Since in these environments establishing new friendships is immediate and effortless, it is reasonable to think that online social networks (OSNs) removed the social boundaries of our modern, technological era. However, what OSNs have allowed us to do is to build massive networks of *weak* ties: acquaintances and nonintimate ties we use all the time to reach out to persons, business requests, speaking engagements, or ideas and advice. Despite such enormous quantity of acquaintances, some works revealed that people still have the same circle of intimacy as ever [5, 2]; and that the formation of friendships is strongly influenced by the geographic distance, breaking down the illusion of living in a "global village" [4].

Network	Nodes	Edges	Weighted
Foursquare	5783	42691	No
GeoFoursquare	4901	17987	Yes
Facebook	2081	5618	No
Twitter	3745	31638	No
Complete Network	7500	97934	-

Table 1. Statistics of the four dimensions and the multidimensional network.

People tend to interact intensely with a small subset of individuals, carrying out a social grooming in order to maintain and nurture strong, intense ties. Strong ties are people we really trust, the ones whose social circles tightly overlap with our own and, often, they are also the people most like us. Presumably, these are the persons we contact for emotional and economic support [13, 3] and often join together to lead organizations through times of crisis [6]. Unfortunately, social media do not incorporate tie strength in the creation and management of relationships, and treat all users the same: friend or stranger, with little or nothing in between.

In the current work, we address the following issue: how to define a tie strength measure that is capable to discriminate between intimate ties and mere online social contacts?

Actually, it does not exist a formal, unique and shared definition of tie strength, and literature has often provided very personal interpretations of Granovetter’s intuition: “the strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding) and the reciprocal service which characterize the tie” [12]. The most frequently used measurements of tie strength in OSNs are based on the number of conversations between users [5]. However, in our opinion these common approaches suffer two major shortcomings. Firstly, the number and intensity of conversations strongly depends from user to user, making it difficult to understand which of these conversations are dedicated to intimate relationships. Secondly, they do not take into account that strong ties must be powered by a form of social grooming, that is mainly based on geographical proximity and face-to-face contacts.

In order to overcome such shortcomings, we propose a new definition for tie strength, which exploits the existence of multiple online social links between two individuals. Indeed, while weak ties often rely on a few commonly available media, strong ties tend to diversify communicating through many different channels. Moreover, the patterns of homophily tend to get stronger as more types of relationships exist between two people, indicating that homophily on each type of relation cumulates to generate greater homophily for multidimensional than monodimensional ties [8]. To model this behavior, we introduce a strength function and test its meaningfulness on a 4-dimensional social network.

2 Related work

The concept of tie strength was introduced by Granovetter in his seminal paper “The Strength of Weak Ties” [12]. He proposed four main factors shaping tie

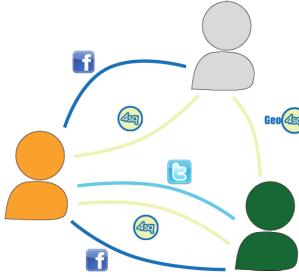


Fig. 1. A schematization of our 4-dimensional social network

strength: amount of time, intimacy, intensity and reciprocal services. In [7], authors used survey data from three metropolitan areas to discover the predictors of tie strength. Onnela et al. [9] utilized the duration of calls as a measure for tie strength, discovering that social networks are robust to the removal of the strong ties but fall apart after a phase transition if the weak ties are removed.

Multidimensional network analysis is a relatively recent field. Authors in [14] analyzed degree distributions on various dimensions, highlighting the need for analytical tools for the multidimensional study of hubs. A framework for the analysis of multidimensional networks is introduced in [1], defining a large set of measures capturing the interplay of the dimensions both at the global and at the local level. In [11] the link prediction problem is addressed in the context of multidimensional networks.

3 Multidimensional tie strength

On the vast online world, two individuals can interact and share interests through several social networking platforms. They can be coworkers on LinkedIn, friends on Facebook or Google+, followers/following on Twitter, they can frequent the same venues on Foursquare, or all of these things together.

To express this kind of information, as done by the authors of [1], we choose as model the one offered by multidimensional networks.

Definition 1 (Multidimensional Network) A multidimensional network is a network in which two nodes can be connected, at the same time, by multiple edges that belong to different dimensions.

We model such structure with an edge-labeled undirected multigraph denoted by a tuple $G = (V, E, L)$ where: V is a set of nodes; L is a set of labels; E is a set of labeled edges, i.e. a set of triples (u, v, d) where $u, v \in V$ are nodes and $d \in L$ is a label. Henceforth, we use the term *dimension* to indicate *label*.

Since strong ties tend to diversify communicating through many different channels, it makes sense to define a tie strength measure that exploits the multidimensional nature of online interactions. In order to do this, we extend traditional approaches adding three other features.

The first one takes into account the intensity of interaction and the similarity of the nodes in a single dimension:

Definition 2 (Node interaction and similarity)

$$h_d(u, v) = w_d(u, v) \frac{|\Gamma_d(u) \cap \Gamma_d(v)|}{\min(|\Gamma_d(u)|, |\Gamma_d(v)|)} \quad (1)$$

where w_d is a weight function representing the intensity of the interaction between the nodes in the dimension d , and Γ_d is the set of neighbors of a node w.r.t dimension d . In order to capture whether they belong to the same circle of friendships, and whether such circle is prominent for one of them, the intensity of interactions is influenced by the percentage of common neighbors with respect to the more selective node (the one with less friends). The second feature regards the relevance of a dimension for the connectivity of a user: the removal of the links belonging to a dimension should not affect significantly the capacity to reach his real strong connections.

Definition 3 (Connection Redundancy)

$$\varphi_d(u, v) = (1 - DR(u, d))(1 - DR(v, d)) \quad (2)$$

The dimension relevance DR [1] is the fraction of neighbors that become directly unreachable from a node if all the edges in a specified dimension were removed. We give an higher score to the edges that appear in several dimensions, so we are interested in the complement of those values. If the two nodes are linked in more than one dimension, the score is raised until a maximum of 2.

We merge these aspects taking into account the multidimensionality of a tie: a greater number of connections on different dimensions is reflected in a greater chance of having a strong tie [8]:

Definition 4 (Multidimensional Tie Strength) Let $u, v \in V$ be two nodes and L the set of dimension of a multidimensional network $G = (V, E, L)$. The strength function $\text{str} : V \times V \rightarrow \mathbb{R}$ between two users u, v is defined as:

$$\text{str}(u, v) = \sum_{d \in D} h_d(u, v)(1 + \varphi_d(u, v)) \quad (3)$$

The measure proposed, given its formulation, could be used to estimate the strength of ties even in monodimensional networks: in that scenario the φ_d function assume a value equal to zero and the overall sum became the value of h_d . This scoring function is our final measure of tie strength.

4 Experiments

We built a multidimensional network $G = (V, E, L)$ by collecting friendships existing between the same 7500 individuals¹ in three online social networks

¹ All considered users are geographically located in the city of Osaka (Japan).

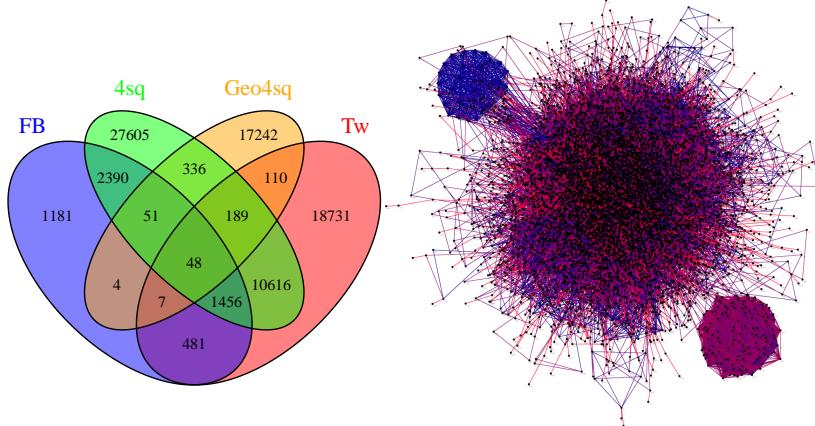


Fig. 2. (Left) Venn diagram showing how the edges of the four dimensions overlap in the multidimensional network; (Right) Global visualization of the network N : colors of edges gradient from blue to red indicate the tie strengths, from strong to weak respectively.

(Foursquare, Twitter and Facebook). Moreover, we inferred a co-occurrence network linking two users if they made a Foursquare checkin in the same venue within a time interval of 15 minutes, during a time span of one month. The number of co-occurrences between two individuals was taken as the weight for the corresponding edge. Figure 1 presents a schematic example of our 4-dimensional network, whereas Table 1 summarizes some characteristics of the multidimensional network and of its dimensions.

To analyze the structural role of strong and weak links, we calculated the strength measure on G and, using the scores obtained, inferred a weighted network $N = (V, E_N)$, collapsing all the edge between two nodes into one. Figure 2 shows a global visualization of N , from which three main clusters clearly emerge, with the one on the left representing people communicating in many different social networking platforms. Our measure seems to be consistent with the “*strength of weak ties*” hypothesis [12], with strong tie connecting local communities, and weak ones acting as bridge between them (Figure 5). To test more rigorously this aspect, we studied the resilience of N and the individual networks to the removal of either strong and weak links. Since weak ties act as bridges between different communities, we expect that their removal made the network structure fall apart quickly [9]. Indeed, the deletion of strong ties do not infect considerably the connectivity of the networks, with the 70% of the nodes still reachable in N removing almost all the strong arcs (Figure 3). Conversely, the removal of weak ties rapidly “*destroys*” the networks, splitting them into several small connected components (Figure 3). Our definition is therefore capable to discriminate between intimate circles and the edges acting as bridges between them.

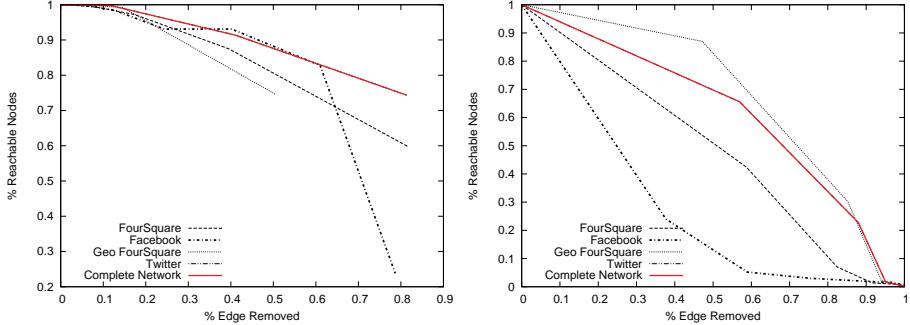


Fig. 3. The stability of the networks to strong link removal. (*Left*) Curves correspond to removing first the high-strength links, moving toward the weaker ones; (*Right*) Curves correspond to removing first the low-strength links, moving toward the stronger ones.

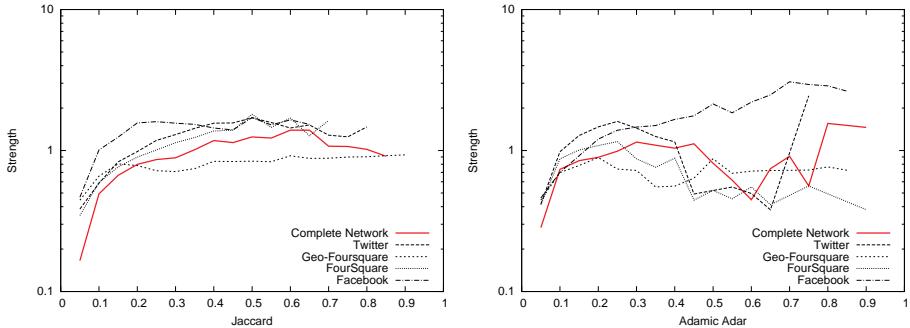


Fig. 4. (*Left*) Relation between Jaccard coefficient and strength values; (*Right*) Relation between Adamic-Adar coefficient and strength values.

Figure 2 shows a Venn diagram representing the number of ties belonging to each possible intersection of the dimensions. It clearly shows that there are only 48 bonds appertaining to all the 4 dimensions. Such links represent a sort of “super strong” ties, i.e. those having a high probability of being real and intimate friendships.

With the purpose of investigate if the proposed measure assigns a strength value correctly, we analyzed how its score correlate with three well-known network measure: Jaccard, Adamic-Adar and Edge Betweenness.

4.1 Strength vs. Jaccard

Comparing the values assigned by our measure with the corresponding Jaccard coefficient, we want to verify the existence of a correlation between the strength of a tie and the similarity of the individuals involved. We plot the tie strength against the Jaccard coefficient, both for the network N and the single dimensions. As shown in Figure 4, weak ties tend to have a small Jaccard coefficient, whereas

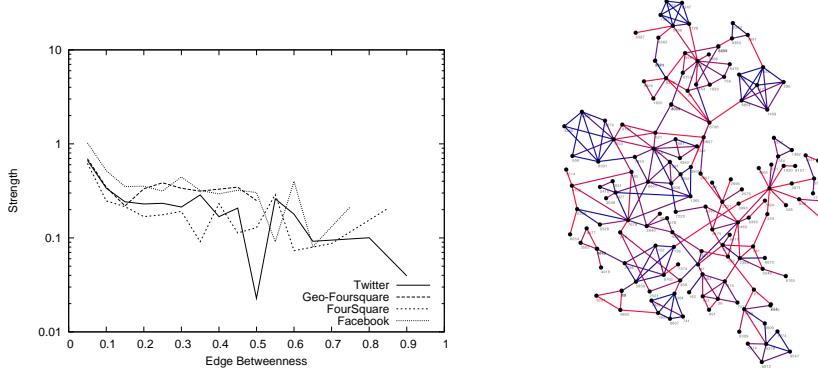


Fig. 5. (Left) Relation between Edge Betweenness coefficient and strength values; (Right) A portion of the network N : colors of the edges in a color gradient from blue to red indicate the strength of ties, from strong to weak respectively.

those with higher strength seems more similar. However, there are cases in which an high similarity does not reflect in higher strength. This is because the Jaccard coefficient is defined as the ratio between the common neighbors and all the friends, whereas our measure takes into account the prominence of the circle of friendships (equation 1).

4.2 Strength vs. Adamic Adar

As done with the Jaccard coefficient, we compare our measure with Adamic-Adar. This measure considers how the mutual neighbors of two nodes are selective in establishing connections: the more selective the friendships are, the more likely the two individuals belong to the same friendship community. As we can see in Figure 4, it seems that the strength increases together with the Adamic score in Facebook, Twitter and the network N . It does not happen with Foursquare, presumably because of the peculiar typology of the service that it offers. Anyway, the trend shown by the figure suggests the following conclusion: two nodes belonging to selective circles of friendships have a greater chance to establish a strong bond.

4.3 Strength vs. Edge Betweenness

The edge betweenness is a measure of edge's centrality, equal to the number of shortest paths that pass through that edge. An edge with an high betweenness is likely a bridge between two different communities and, by definition, a weak link. We compare our strength function against this score computed over the single dimensions only. The computation of this measure on the network N is meaningless because, in such network, an edge could establish paths that are not real. As expected, Figure 5 shows that when the edge betweenness increases, the value of strength seems to decrease.

5 Conclusion

In this work we have introduced a measure of tie strength for multidimensional networks. Supported by a validation on a 4-dimensional social network, we observed that the strength is strictly related to the number of interactions among the individuals involved. Moreover, we assessed the hypothesis that strength is also related to the number of different contexts in which those connections take place. In the future, we plan to investigate how the information provided by the tie strength can be exploited to tackle well-known problems such as link prediction and community discovery.

Acknowledgements. This research work has been partially supported by European FP7 project DATASIM (<http://www.datasim-fp7.eu/>).

References

1. Michele Berlingario, Michele Coscia, Fosca Giannotti, Anna Monreale, and Dino Pedreschi. Foundations of multidimensional network analysis. In *ASONAM*, pages 485–489. IEEE Computer Society, 2011.
2. R.I.M. Dunbar. The social brain hypothesis. *brain*, 9:10, 1998.
3. James H. Fowler and Nicholas A. Christakis. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the framingham heart study. *BMJ*, 337(dec04.2):a2338+, December 2008.
4. Jacob Goldenberg and Moshe Levy. Distance is not dead: Social interaction and geographical distance in the internet era. *CoRR*, abs/0906.3202, 2009.
5. Bruno Gonçalves, Nicola Perra, and Alessandro Vespignani. Modeling users' activity on twitter networks: Validation of dunbar's number. *PLoS ONE*, 6(8):e22656, 08 2011.
6. David Krackhardt and Robert N Stern. Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly*, 51(2):123–140, 1988.
7. Peter Marsden. Measuring tie strength. *Social Forces*, 63:482–501, 1984.
8. Miller McPherson, Lynn Smith-Lovin, and James M. Cook. Birds of a feather: Homophily in social networks. *Annu. Rev. Sociol.*, 27(1):415–444, 2001.
9. J. P. Onnela, J. Saramaki, J. Hyvonen, G. Szabo, D. Lazer, K. Kaski, J. Kertesz, and A. L. Barabasi. Structure and tie strengths in mobile communication networks, 2006.
10. Luca Pappalardo, Giulio Rossetti, and Dino Pedreschi. "how well do we know each other?" detecting tie strength in multidimensional social networks. In *ASONAM*, pages 1040–1045, 2012.
11. Giulio Rossetti, Michele Berlingario, and Fosca Giannotti. Scalable link prediction on multidimensional networks. In Myra Spiliopoulou, Haixun Wang, Diane J. Cook, Jian Pei, Wei Wang 0010, Osmar R. Zaane, and Xindong Wu, editors, *ICDM Workshops*, pages 979–986. IEEE, 2011.
12. Granovetter Mark S. The strength of weak ties. 1973.
13. Catherine Schaefer, JamesC. Coyne, and RichardS. Lazarus. The health-related functions of social support. *Journal of Behavioral Medicine*, 4:381–406, 1981.
14. Michael Szell, Renaud Lambiotte, and Stefan Thurner. Trade , conflict and sentiments : Multi-relational organization of large-scale social networks. *arXiv*, physics.so, 2010.