Decomposition of Social Ties with Blau's Exchange Theory

Luca Maria Aiello
Yahoo Labs Barcelona

Rossano Schifanella
University of Torino

Bogdan State
Stanford University

ABSTRACT

The explosion of data from online social media has encouraged the often uncritical adoption of the notion of *social tie* as the atomic interaction quantum of any social network structure. Social ties are usually treated as *a priori* entities, immediately available to the researcher (e.g., friending on Facebook) and have been interpreted as indicative of one social process or another (e.g., status exchange or trust), often with little systematic justification regarding the relation between observed data and theoretical concept. Even though previous research has explored several aspects of social links including their intensity and polarity, there is still much to investigate about the *nature* of the social interactions implied by social ties.

To breach this gap in computational social science, we study social ties under the light of Peter Blau's Exchange Theory [2], conceiving every social dyad as a repeated set of exchanges of different types of non-material *resources* transacted in an interpersonal situation, such as knowledge, social support or manifestation of approval. Being able to describe a conversation in terms of these resources would provide a new abstraction level that could facilitate the interpretation of the meaning of social connections. To operationalize this notion, we mine online *conversations*, namely dyadic exchange of textual messages, and we define a method to cluster messages by the type of resource they convey, rather than by their topical aspect. Our algorithm is based on the intuition that in a dyad, social interactions conveying a resource tend to be reciprocated with the same resource type. As an illustration, if two individuals exchange knowledge now, their next exchange will be most likely to also involve knowledge, rather than affection.

By applying our method on two online datasets different by scope and type of interaction (Flickr and aNobii) we observe the spontaneous emergence of three types of resources exchanged: *status*, *knowledge* and *social support*. By finding significant relations between such resources and classic social network analysis issues (tie strength, assortativity, dyadic interaction over time) we show how the network of interactions induced by the extracted domains can be used as a starting point for more nuanced analysis of online social data that may one day incorporate the normative grammar of social interaction.

1. METHODOLOGY

Our method has the following input/output:
Input: a population of users U and a set of messages M where each message $m^t_{u,v} \in M$ is a textual communication between source $u \in U$ and destination $v \in U$ at time t.
Output: a probabilistic clustering of messages in M with probability of a message m to be assigned to cluster D being ≥ 0.

The novel aspect of the method is that messages are grouped according to the type of social exchange those messages convey instead of their topic. The algorithm is composed by four phases:

Preprocessing. All messages are passed through a common IR preprocessing pipeline which includes stopword removal, stemming, and extraction of 2-3grams. The vector of stemmed grams representing the messages are stacked in a *term-document matrix* $\Gamma_{m \times n}$ where m is the total number of terms and n is the total number of messages.

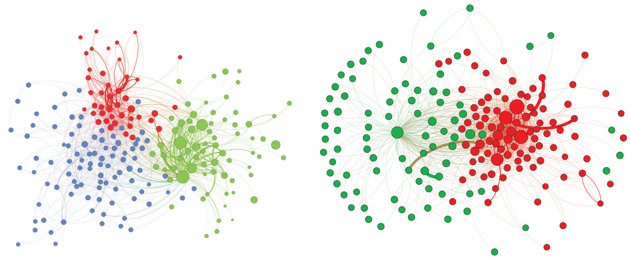

Figure 1: Conversation Graphs (Left: aNobii, right: Flickr). Nodes represent buckets of topically-coherent messages and edges the transitions between them in dyadic conversations. Colors encode messages exchanging *Social Support* (red), *Status* (green), and *Knowledge* (blue). The node size accounts for the number of messages in the bucket.

Message Bucketing. Messages conveying the same semantics are grouped together through a low rank *Non-negative Matrix Factorization (NMF)* of $\Gamma_{m \times n}$, yielding a probabilistic assignments of messages to topical *buckets*.

Creation of Conversation Graph. A weighted directed *Conversation Graph* (*CG*) is built, where nodes are buckets and edges represent transitions between them based on the conversational flow. A weighted arc (i, j) captures the likelihood that a message in bucket i is followed by reply message in bucket j. The Conversation Graph shapes the transition between classes of coherent messages during social interactions.

Extraction of message clusters. We assume that a message that conveys a certain type of resource will most likely get a reply that conveys the same resource type. Under this interpretation, highly-clustered parts of the *CG* aggregate buckets whose messages carry a homogeneous resource type. This scenario is consistent with the most common definition of graph *community*. We therefore use network a community detection algorithm (*Spinglass*, in our experiments) to spot dense areas in the *CG*. According to this intuition, each community is supposed to include (with a certain probability) buckets whose messages convey the same type of resource.

2. EXPERIMENTAL RESULTS

We test our framework on conversation datasets from two social media: aNobii, a website for book lovers where users can post direct messages on each other's "walls" (62,235 users, 545,656 conversations), and Flickr, the popular image sharing website where people can post comments on each other's photos (95,397 users 100,000 conversations).

Running our algorithm on both datasets yields clusters over the *CG*s (Figure 1). A manual inspection of the aNobii messages done by a sociologist suggests that the three communities found corre-

	DoI	Tie Share	Structural sim			Intensity		Sentiment		Kinship
			σ_n	σ_g	σ_i	$\langle conv_{len} \rangle$	$\langle msg_{len} \rangle$	Intim.	Emo.	
aNobii	*Status*	0.48	0.045	0.062	0.041	2.13	16.32	0.026	0.033	n/a
	Support	0.33	0.064	0.077	0.054	3.03	18.81	0.040	0.040	n/a
	Knowledge	0.19	0.068	0.075	0.059	2.48	23.27	0.038	0.036	n/a
Flickr	*Status*	0.51	0.028	0.024	0.0011	8.83	6.26	0.370	0.393	0.049
	Support	0.49	0.040	0.024	0.0013	12.70	7.35	0.410	0.440	0.057

Table 1: Strength of ties connecting pairs of users, in terms of: i) Jaccard similarity σ between their neighbors (n), the groups they are subscribed (g) and their items (i), books for aNobii and favorited photos for Flickr; ii) length of the conversation in terms of number of messages exchanged; iii) ratio of words belonging to the intimacy and emotions categories in the LIWC categories; iv) ratio of dyads reciprocally declaring a "family" or "friend" relation (Flickr only). The average portion of messages in a dyad carrying a certain resource is also reported as *tie share*.

spond to as many fundamental processes of social exchange:

Status exchange. According to the Power-Dependence Theory [3], heterogeneity of resource endowments in a dyadic relationship leads to *power imbalances*. *Status giving* is a way in which a low-power actor may attempt to reduce their power inequality. In practical terms, status giving is often instantiated in messages displaying appreciation, esteem, or admiration sent to social partners with higher power.

Social Support. The minute exchanges between individuals that form the essential structure of social interactions, a basic process of friendship through which one partner provides emotional valuation to another through good wishes, colloquial chat, jokes and laughter.

Knowledge exchange. The act of sharing one's knowledge or opinions with others.

In Flickr analogous domains emerge, with the exception of the knowledge exchange one. To check the accuracy of our approach, we resort to three independent editors who marked 1,000 messages with zero or more labels corresponding to the three above mentioned resource types (Fleiss' Kappa agreement 0.70). Our algorithm matches the ground truth in about 80% of the cases.

The possibility of extracting the resources transacted on social ties allows social analysts to isolate the different processes of social exchange and to directly check sociological theories specific to them. We color each edge of the message-exchange graph (nodes are users, edges their conversations) with the resources exchanged in that dyad. We then extract the subgraphs with edges of homogeneous color and study them in isolation, across several aspects.

Coverage and reciprocation. The proportions of graph edges of each type is balanced in Flickr (about 66% for status and 64% for support), while more skewed on status exchange (75%) in aNobii. High reciprocity (computed as the ratio of reciprocated messages between two endpoints) is found for all resource types but for status exchange especially (0.861), likely a reflection of social norms imposing the ritualized reciprocation of status exchange.

Tie strength. To understand whether different resource exchanges are characterized by different tie strengths, we adopt the framework presented by Gilbert and Karahalios [4] based on Granovetter's definition of tie strength. We measure the strength based on three main families of metrics: *structural similarity* (sharing of common acquaintances and features), *intensity* (duration of their interaction), and *sentiment* (amount of words expressing intimacy and emotion). In addition, Flickr data allows us to investigate also the *kinship* dimension, namely whether the endpoints declared to be friends or family members. Detailed results are presented in Table 1. In short, weaker ties tend to convey status giving and stronger ties (longer conversations, higher similarity) either social support or knowledge.

Inequality and assortativity. Resources can be considered as *goods* generated by the social actors and exchanged between them. The indegree of a node on a resource-specific subgraph is a proxy of the amount of resource owned. The *social inequality*, measured as the Gini index of the indegree, is high for all the resources, but for status especially (0.72). This supports the intuition that the status, more than other goods, tends to flow unidirectionally from low- to high-status individuals. This is confirmed also by the status network being the only one exhibiting degree in-in disassortativity.

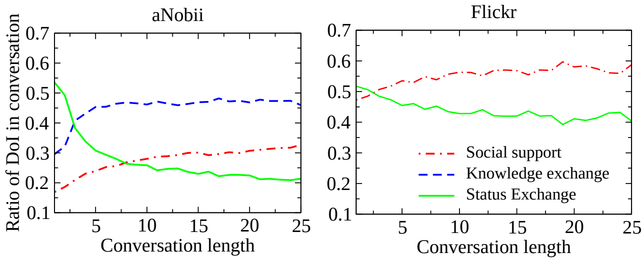

Figure 2: Average proportion of messages belonging to each DoI for pairs of users with fixed conversation length.

Tie evolution. We compute across all the users the average ratio of resources exchanged in conversations with different lengths (Figure 2). It thus appears that status exchange serves to set the foundation for the future relationship, being present in the first stages of a conversation and fading to the interactional background after the tie-formation stage, leaving space to exchange of knowledge and support. In both datasets the status giving curve starts losing its predominance exactly after 3 messages exchanged.

3. CONCLUSIONS

The characterization of messages in terms of their type of social exchange opens a plethora of opportunities for applications ranging from analytics to user/link profiling and summarization of social relationships (e.g., Alice and Bob exchange 30% of knowledge, 20% of status and 50% of social support). This naturally leads one to the idea of understanding social ties as *strings of interactions*. With this understanding, we can use insights from theoretical Computer Science to establish the computational properties of social rituals. The ultimate goal of such analysis is the unpacking of "culture" as a formal, computational concept. If we see social ties as interactional sequences, then we may understand the resource types we discover as the "grammar of society" [1] – in other words, the bits of "source-code" that prescribe how individuals are to act in a certain situation. We hope our work provides yet another step towards a truly computational understanding of human societies.

4. REFERENCES

[1] C. Bicchieri. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press, 2006.

[2] P. Blau. *Exchange and Power in Social Life*. Transaction Publishers, 1964.

[3] K. S. Cook and R. M. Emerson. Power, Equity and Commitment in Exchange Networks. *American Sociological Review*, 43(5), 1978.

[4] E. Gilbert and K. Karahalios. Predicting Tie Strength with Social Media. In *CHI*, 2009.