

## **Authors:**

**Thomas Plotkowiak, University of St. Gallen**

**Katarina Stanoevska-Slabeva, University of St. Gallen**

## **Title: Information Diffusion in Twitter Communities**

### **Extended Abstract**

Social scientists have long recognized the importance of social networks in the spread of information (Granovetter, 1973) and innovation (Rogers, 1995). Modern communications technologies, notably email and more recently social media, have only enhanced the role of networks in marketing (Domingos & Richardson, 2001), information dissemination (Gruhl, Guha, Liben-Nowell, & Tomkins, 2004; Wu & Huberman, 2004) or search (Adar, Zhang, Adamic, & Lukose, 2004a). The new emerging network data offers a rich source of evidence for studying the structure of social networks (Leskovec & Horvitz, 2008) and the dynamics of individual and group behavior (Lerman, 2007).

Yet in most studies, the structure of the underlying network was not directly visible but had to be inferred from the flow of information from one individual to another. The same held for the different forms of interactions of users or the strength of their connection with each other. Finally the actual diffusion of information was mostly also not visible but had also to be inferred. This posed a serious challenge to our efforts to understand how the structure of the network affects dynamics of information spread on it. In contrast to the historical hindrances the homogenous medium of Twitter allows us to study how individuals embed themselves in a social medium. With whom they interact in which form (@-replies, direct messages, etc.) allows us to observe the actual information diffusion in form of retweets.

Another shortcoming in existing information diffusion studies is that although a vast body of literature explains the diffusion of information in cohesive groups, there are only few studies that have analysed multiple groups and communities at the same time and studied the diffusion of information between bigger communities. The community has always either remained the outer cover of the data collection and has so created the framework for each study or it perished completely as a structural component when information diffusion has been studied for the whole population (Adar, Zhang, Adamic, & Lukose, 2004b; Leskovec, McGlohon, Faloutsos, Glance, & Hurst, 2007). Yet it is explicitly this "meso" - level of analysis, which describes different communities with their members and roles that allow for interesting insights into the mechanics information diffusion inside and between cohesive groups. For example an idea or innovation can arise in one community and then spread

throughout this community with the help of opinion leaders or be transported to another community with the help of brokers between those communities.

We argue that Twitter users embed themselves different topic groups (Byrne, 1971; Turner, Smith, Fisher, & Welser, 2005), where not the attributes of the actor and his profile decide about his interest but rather whom he follows and whom he interacts with. As a consequence of this fuzzy membership in a homogenous medium the individual is embedded in an ecosystem of topic communities, which compete for his attention. The study of such communities presents a very fruitful ground in order to analyse the information diffusion between such communities, which has not yet been studied before.

The third issue we want to address in this paper are the existing role concepts in literature that describe how certain roles in populations can significantly influence the diffusion of information (Bass, 2004; Gladwell, 2000; Katz & Lazarsfeld, 1955; Lazarsfeld, Berelson, & Gaudet, 1968; Merton, 1968; Schenk, 1993; Valente, 1993; Weinmann, 1994) or information (R. Burt, 1995). The current theory and methods for the identification of such roles, do often lack to address the issue if their potential influence might only be limited to their own community. In our described topic communities, it we therefore address the question how actors with a strong structural position are involved in diffusion processes and if their role inside the community plays an part in the diffusion of information inside and across different communities. Similar to opinion leaders we are asking the same question for structural brokers (Ronald S Burt, 2010) in online communities: If the population is fragmented like in our case into multiple topic communities, we want to ask if topic brokers that facilitate the information diffusion between certain communities, or if their influence is marginal.

## Method

To analyze how the structure of networks and the resulting communities influences the spread of information we analyzed 5 different twitter topic communities consisting of 500 members, their 1 Mio. Tweets and their interactions and retweets among each other. The 5 communities were sampled from the information science domain, where we collected communities around different programming languages.

In our study we developed our own method of collecting such communities, which are initially constructed by collecting suggested users from websites like *www.wefollow.com*, *www.twellow.com* and then substantially are being "grown" or extended by using the twitter list feature. By using lists that provide similar users to the "seed" users we are extending those communities to a specified size. By having information on how often certain members are listed for a given topic or keyword we can obtain communities that have a high density and a high amount of interaction. In this collection of communities we analyzed:

- **The hypothesis of homophily** (McPherson, Smith-Lovin, & Cook, 2001) **and information diffusion:** Can we verify that most retweets end up being transmitted to the same community? We therefore checked if the members that retweeted a given tweet were members of the same originating community as the original author or members one of the other communities, or could not be allocated at all.
- **The hypothesis of strong ties** (Granovetter, 1973) **and information diffusion:** Can we verify that the chance of a retweet is higher for members that hold strong ties between each other? Here the strength of a tie was defined by their interaction frequency i.e. how often member mention each other in @-replies.
- **The hypothesis of opinion leaders and brokerage roles** (R S Burt, 1999) **and information diffusion:** We verified if structural network roles of a **broker** and an **opinion leader** in a global (network of all members) and local (community wide) had an impact on the way retweets are transmitted through the network. In specific if opinion leaders were able to create more, faster or bigger retweets and if brokers were more successful than others to transfer retweets into other communities or create retweets outside of the community.

## Results

Our results show that although we could verify the hypotheses from literature their explanatory value is not as strong as we would tend to expect. The regression of actual observed diffusion of information with networks and networks measures, allowed us to obtain predictions of the information diffusion based on the presented structural properties of the networks. Our work provides an empirical contribution not only to the field of information diffusion but also towards the discussion of the validity of network measures in social media. The findings of our empirical analysis bear valuable information for seeding information on social networks or social news like twitter. They help to answer if forwarding information rather is done in a scattered unorganized way, or if cohesive groups, strong ties and structural roles have an influence on the dissemination of information on twitter.

## References

- Adar, E., Zhang, L., Adamic, L. A., & Lukose, R. M. (2004a). Implicit structure and the dynamics of blogspace. *Workshop on the Weblogging Ecosystem*, 13(1). Retrieved from [http://www.blogpulse.com/papers/Adar\\_blogworkshop2\\_ppt.pdf](http://www.blogpulse.com/papers/Adar_blogworkshop2_ppt.pdf).
- Adar, E., Zhang, L., Adamic, L. A., & Lukose, R. M. (2004b). Implicit structure and the dynamics of blogspace. *Workshop on the Weblogging Ecosystem*, 13(1). Retrieved from [http://www.blogpulse.com/papers/Adar\\_blogworkshop2\\_ppt.pdf](http://www.blogpulse.com/papers/Adar_blogworkshop2_ppt.pdf).
- Bass, F. M. (2004). A New Product Growth for Model Consumer Durables. *Management Science*, 50(12), pp. 1825-1832. INFORMS. Retrieved from <http://www.jstor.org/stable/30046153>.

- Burt, R S. (1999). *The Social Capital of Opinion Leaders*, 566, 37. Philadelphia: Published by A.L. Hummel for the American Academy of Political and Social Science.
- Burt, Ronald S. (2010). Structural holes in virtual worlds, 83.
- Burt, R. (1995). Structural holes: The social structure of competition. Retrieved March 2, 2011, from <http://books.google.com/books?hl=en&lr=&id=E6v0cVy8hVIC&oi=fnd&pg=PR7&dq=structural+holes+the+social+structure+of+competition&ots=omTPPeZ5IH&sig=8oK47dTOYqUmCvnKNOntZImJUgk>.
- Byrne, D. (1971). *The attraction paradigm* (p. xi, 474 p.). Academic Press. Retrieved from <http://www.amazon.com/Attraction-Personality-Psychopathology-Psycho-pathology-Monographs/dp/0121486508>.
- Domingos, P., & Richardson, M. (2001). Mining the network value of customers. *International Conference on Knowledge Discovery and Data Mining*, 57. Retrieved August 10, 2010, from <http://portal.acm.org/citation.cfm?id=502525>.
- Gladwell, M. (2000). *The Tipping Point*. *booksgooglecom* (Vol. 34, pp. 16-17). Little, Brown and Company. doi: 10.1016/S0024-6301(00)00101-1.
- Granovetter, M. S. (1973). The Strength of Weak Ties, 78(6), 1360-1380. The University of Chicago Press. Retrieved from <http://www.jstor.org/stable/2776392>.
- Gruhl, D., Guha, R., Liben-Nowell, D., & Tomkins, A. (2004). Information diffusion through blogspace. (S. I. Feldman, M. Uretsky, M. Najork, & C. E. Wills, Eds.) *Proceedings of the 13th conference on World Wide Web WWW 04*, 6(May 2004), 491. ACM Press. doi: 10.1145/988672.988739.
- Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence* (pp. 162-206). Free Press.
- Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1968). *The people's choice*. New York. Retrieved March 3, 2011, from <http://www.citeulike.org/user/cnlawrence/article/2317525>.
- Lerman, K. (2007). Social Information Processing in News Aggregation. *IEEE Internet Computing*, 11(6), 16-28. IEEE Computer Society. doi: 10.1109/MIC.2007.136.
- Leskovec, J., & Horvitz, E. (2008). Planetary-scale views on a large instant-messaging network. *Proceeding of the 17th international conference on World Wide Web WWW 08*, 393, 915. ACM Press. doi: 10.1145/1367497.1367620.
- Leskovec, J., McGlohon, M., Faloutsos, C., Glance, N., & Hurst, M. (2007). Cascading Behavior in Large Blog Graphs. *SIAM International Conference on Data Mining SDM 2007*, 1-21. Retrieved from <http://arxiv.org/abs/0704.2803>.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily. *Annual Review of Sociology*, 27, 415-444.
- Merton, R. K. (1968). Patterns of Influence: Local and Cosmopolitan Influentials. In R. K. Merton (Ed.), *Social Theory and Social Structure*. Free Press.
- Rogers, E. M. (1995). *Diffusion of innovations* (Fourth Edi.). New York, London, Toronto, Sydney, Tokyo, Singapore: The Free Press.

- Schenk, M. (1993). Die Ego-zentrierten Netzwerke von Meinungsbildnern ("Opinion Leaders"). *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 45, 254-269.
- Turner, T. C., Smith, M. A., Fisher, D., & Welser, H. T. (2005). Picturing Usenet: Mapping Computer-Mediated Collective Action. *Journal of Computer-Mediated Communication*, 10(4), 0. Blackwell Publishing Ltd. Retrieved from <http://jcmc.indiana.edu/vol10/issue4/turner.html>.
- Valente, T. (1993). Diffusion of Innovations and Policy Decision Making. *Journal of Communication*. Retrieved March 2, 2011, from <http://onlinelibrary.wiley.com/doi/10.1111/j.1460-2466.1993.tb01247.x/abstract>.
- Weinmann, G. (1994). *The influentials: people who influence people*. New York. Retrieved March 3, 2011, from [http://scholar.google.ch/scholar?hl=de&q=the+influentials,+weinmann&btnG=Suche&lr=&as\\_ylo=&as\\_vis=0#0](http://scholar.google.ch/scholar?hl=de&q=the+influentials,+weinmann&btnG=Suche&lr=&as_ylo=&as_vis=0#0).
- Wu, F., & Huberman, B. A. (2004). Social Structure and Opinion Formation. *Physics*, 1-25. Retrieved from <http://arxiv.org/abs/cond-mat/0407252>.