Social Networks 34 (2012) 470–480

Contents lists available at SciVerse ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Influentials, novelty, and social contagion
The viral power of average friends, close communities, and old news☆

Nicholas Harrigan[a,*], Palakorn Achananuparp[b], Ee-Peng Lim[b]

[a] School of Social Sciences, Singapore Management University, 90 Stamford Road, Level 4, Singapore 178903, Singapore
[b] School of Information Systems, Singapore Management University, 80 Stamford Road, Singapore 178902, Singapore

ARTICLE INFO

Keywords:
Social contagion
Subgraphs
Network motifs
Influentials hypothesis
Community structures
Twitter

ABSTRACT

What is the effect of (1) popular individuals, and (2) community structures on the retransmission of socially contagious behavior? We examine a community of Twitter users over a five month period, operationalizing social contagion as 'retweeting', and social structure as the count of subgraphs (small patterns of ties and nodes) between users in the follower/following network.

We find that popular individuals act as 'inefficient hubs' for social contagion: they have limited attention, are overloaded with inputs, and therefore display limited responsiveness to viral messages. We argue this contradicts the 'law of the few' and 'influentials hypothesis'.

We find that community structures, particularly reciprocal ties and certain triadic structures, substantially increase social contagion. This contradicts the theory that communities display lower internal contagion because of the inherent redundancy and lack of novelty of messages within a community. Instead, we speculate that the reasons community structures show increased social contagion are, first, that members of communities have higher similarity (reflecting shared interests and characteristics, increasing the relevance of messages), and second, that communities amplify the social bonding effect of retransmitted messages.

1. Introduction

We know that many forms of social behavior and ideas can be thought of as contagious: participation in strikes, riots, voting or migration; the spread of innovations, fashions, fads, rumors and job advertisements; the forwarding of emails, blog links, status updates or 'wire' news stories; changes in long term behavior such as education, smoking, diet and crime; the solving of collective action problems, such as climate change; and the failure of infrastructure and social institutions, from power grids to judicial and currency systems.

It is also self-evidently true that the social structure of a community will heavily affect the pattern of the spread of a contagious idea: ideas can only pass between individuals who have a relationship, and so the pattern of these relationships will affect the spread of ideas.

But what is the exact relation between social structures – particularly local social structures – and the successful spread of contagious ideas and behaviors? We examine a community of Twitter users over a five month period, studying 'retweeting' behavior on the follower/following network of users. Twitter messages are modeled in a similar way to a disease on a network, with 'infections' (tweets) being attributes that are passed along the follower/following network, and retweeting another user's message being a sign of infection.

We focus on the effect of two types of social structures: (1) popular individuals (such as users with a large number of followers) and (2) community structures (such as mutual and triadic structures between users). We operationalize the concepts of popular individuals and community structures by counting subgraphs – small patterns of ties and nodes also called graph statistics or network motifs – in the networks of senders and receivers of contagious messages. Subgraphs measuring popular individuals included the number of followers, number following, and number of messages sent (tweets).[1] Subgraphs measuring community structures included the mutual dyad (reciprocity), and a range of triadic structures, including the 3-cycle and the transitive triad.

☆ This research was supported by Singapore Management University Internal Grant C242/MSS9S014, and by a post-doctoral position at Nuffield College, University of Oxford, under the supervision of Peter Hedström. This research was supported by the Singapore National Research Foundation under its International Research Centre@ Singapore Funding Initiative and administered by the IDM Programme Office.

* Corresponding author. Tel.: +65 6828 0842; fax: +65 6828 0423.
E-mail addresses: nharrigan@smu.edu.sg, nickharrigan@gmail.com (N. Harrigan), palakorna@smu.edu.sg (P. Achananuparp), eplim@smu.edu.sg (E.-P. Lim).

0378-8733/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.socnet.2012.02.005

[1] Note that the word 'popularity' is used in this paper to refer to both nodes with high in-degree (the traditional meaning) and nodes with high out-degree (traditionally called 'activity').

We test a range of competing theories of the effect of social structure on contagious ideas. For popular individuals, we test the competing theories of efficient hubs (such as the 'influentials hypothesis' and the 'law of the few' Katz and Lazarfeld, 1955; Merton, 1968; Gladwell, 2000; Rosen, 2000; Watts and Dodds, 2007; Watts, 2007) and inefficient hubs (Barabasi, 2002). For community structures, we test two schools of thought: firstly, those that emphasize the negative effect of community structures on contagion, with a particular focus on theories of the negative effect of message redundancy (and lack of novelty) on contagion (Granovetter, 1973, 1983; Burt, 1992, 2005; Cha et al., 2010; Hill et al., 2006; Leskovec et al., 2007; Wu and Huberman, 2007); and secondly, those that emphasize the positive effect of community structures on contagion, with a focus on theories of the increased user similarity and increased social bonding effect within a community.

This paper begins with a literature review and an outline of our hypotheses (Section 2), followed by an overview of our dataset (Section 3) and methods (Section 4), particularly those methods which involve the counting of subgraphs. We then present our results (Section 5) and a discussion (Section 6) of our findings in light of existing research.

2. Literature review and hypotheses

So what do previous studies argue are the major effects of (1) popularity, and (2) community structures, on the spread of contagious ideas and behaviors? We organize this literature review by major competing theories (explanations), grouping previous studies with the theories they support.

2.1. Popularity

2.1.1. Efficient hubs

A large number of authors emphasize the positive 'influential' role played by highly popular individuals – those with a large number of friends or followers, or a high communication volume – in the spread of socially contagious phenomena. These individuals are presented as analogous to communications hubs that become more efficient and influential as they gain more network partners.

This theory is exemplified by the 'law of the few' and the roles of 'connector' and 'maven' in Gladwell's famous book *The Tipping Point* (Gladwell, 2000). This 'influential's hypotheses', as Watts calls it (Watts and Dodds, 2007; Watts, 2007), has a long tradition in sociological and marketing theory: highly connected individuals have been called variously "opinion leaders" (Katz and Lazarfeld, 1955), "influentials" (Merton, 1968), or "hubs" (Rosen, 2000). Certain recent academic studies have provided limited quantitative support for this perspective. Cha et al. (2010) found in a study of Twitter that content aggregation services and news sites (which had large numbers of followers) were the most retweeted users. The same study found that the higher the status of a user, the greater their likelihood of being 'mentioned' (replied to).

At the theoretically level, Watts argues (Watts and Dodds, 2007; Watts, 2007) and we agree, that the exact mechanisms driving the supposed positive relationship between popularity and social contagion are, at best, poorly specified in the above literature. Nonetheless, we believe that a theory can be knitted together that captures the major elements common to these accounts: what unites these theories of 'efficient hubs' is, firstly, a claim that there is a positive feedback cycle that is reinforcing the importance of popular actors. The exact mechanisms driving positive feedback cycles will varies across cases, but an archetypical example might include (a) greater contacts leading to: (b) better information, (c) greater visibility, (d) increased social status, and (e) increased trust. With each of (b) through (e) directly and/or indirectly leading to (a) greater contacts, the positive feedback cycle is completed. This generates a 'rich get richer' (Barabasi and Albert, 1999) scenario where a very small number of actors monopolize a disproportionate amount of the social contacts, information, visibility, status, and trust in a network.

The second claim of theories of efficient hubs is that these highly connected (popular) actors are the central conduits of contagious information spread. Depending on the particular author/study, this increased role in contagion is a result of (1) their increased number of contacts, and/or (2) their disproportionate possession of the means to influence any one individual–they may be a more trusted source, of higher status, have greater persuasive powers. As a result of both disproportionate friendships and means of influence, a message sent by a popular individual will, on average, be passed-on (i.e. spread, retweeted, forwarded, etc.) to a considerably larger audience than any message received by an average user. This is the theory of 'efficient hubs'.

2.1.2. Inefficient hubs

Another, more critical, section of the literature on social contagion emphasizes that highly connected individuals tend to show sub-optimal viral reproduction rates.[2] In a theoretical study, Golub and Jackson (2007) found that efficient diffusion of influence through a network is limited by the presence of highly influential, high degree nodes. Empirical studies of Twitter have similarly found that the most 'influential' actors – measured by their impact on hashtag adoption – were moderately 'sized' (followers, volume of tweets) users (Yang and Leskovec, 2011). An empirical study of an online bookstore found similar trends: high degree nodes have lower influence per recommendation; nodes tend to only be influential over their close friends; and books with higher numbers of recommendations tend to have lower viral reproduction (Leskovec et al., 2007).

While theoretically these empirical accounts differ in their hypothesized mechanisms, there is a common element to most studies: theories of inefficient hubs tend to emphasize the limits on efficient (1) growth and (2) activity which social hubs face. Limits on (1) growth occur because ties themselves are not costless: forming friendships requires time, and time is necessarily limited. Limits on (2) activity occur because communication itself is not costless: maintaining friendships requires time and energy as well. While in some non-human networks (such as hyperlinks on the internet), ties are costless, in the real human social world this is often not the case: limited time and attention place natural limits on the capacity of individual humans to meaningfully expand their social networks and to be the fulcrums of the spread of socially contagions behavior (Barabasi, 2002).

The theory of 'inefficient hubs' says that those individuals who deviate significantly from the average – who (for whatever reason) push the natural boundaries on the size of a human's social networks – will face the problem of overload and inefficiency. Because of this overload and inefficiency, highly popular individuals are expected to have a disproportionately lesser impact on social contagion than their number of social contacts would suggest. Instead, the theory of inefficient hubs predicts that real socially contagious transmission will occur amongst less 'overloaded', more 'average', individuals.

[2] Watts and Dodds (2007) provide another criticism of the influentials hypothesis: they find that they can simply not simulate a world in which influentials are important in starting contagious outbreaks. They vary a large number of parameters, but find very few situations in which influentials are substantially more influential than the average individual.