Relative Superiority of Key Centrality Measures for Identifying Influencers on Social Media

Yifeng Zhang, Department of Management Information Systems, University of Illinois at Springfield, Springfield, IL, USA

Xiaoqing Li, Department of Management Information Systems, University of Illinois at Springfield, Springfield, IL, USA

ABSTRACT

Marketers have been increasingly turning to social media for marketing campaigns, including viral marketing. A key step in viral marketing is to identify influencers in order to maximize the reach of a marketing message. Existing research shows that centrality measures, such as degree and betweenness, are effective methods for influencer identification. However, viral marketing models used in different studies vary greatly, making it difficult to compare findings across the studies. In this paper, the authors built an agent-based framework of viral marketing that supports different experiment settings, such as different network structures and information diffusion modes, and used it to study relative superiority of various centrality measures. The results show that relative superiority of the measures are affected by some factors, but not as much by others. Practical implications of the results are discussed.

Keywords: Agent-Based Simulation, Centrality, Influencer Identification, Social Media, Viral Marketing

INTRODUCTION

The Internet is entering a social media era. Many social media sites, such as Facebook, Twitter, and Pinterest, have emerged and become very popular with users in recent years. Users spend a large amount of time on social media, interacting with families and friends, reading and spreading news, playing online games, posting pictures etc. What makes social media different from other Internet services is that they are built upon relationships and interactions between real people. There is a complex social network embedded in each social media site (Yin, Li, & Niu, 2014).

Social media have brought great business potentials for companies. One way to harness such potentials is to conduct viral marketing...
campaigns. As an effective marketing approach to spread information about products or services from consumers to consumers, viral marketing has existed for a long time, probably as long as the commerce itself. However, social media can elevate it to a new level because it can be implemented on an unprecedented scale on social media.

A viral marketing campaign starts with selecting a group of “seed” users who will be the first to receive information about a product or a service, by letting them try the product, for example. After learning about the product, some of the seed users will, hopefully, share the information or their experiences with their families, friends, co-workers etc., who in turn will share the information with their families and friends. As a result, information about the product will reach more and more people as time passes. The success of a viral marketing campaign is determined by the number of people eventually reached by the marketing message and how fast those people are reached.

A critical success factor of viral marketing is the selection of initial seed users (Gruhl, 2004). In general, seed users need to have a large number of social connections and are trusted by other people in social networks. These people are considered influencers, i.e. people who exert great influence over others. As a result, selecting the “right” seed users for a viral marketing campaign boils down to identifying influencers on social media. The issue has attracted much attention from both practitioners and researchers. In fact, several companies, such as Klout, Tellagence, and Little Bird, built their businesses around calculating people’s influence on social media (Karif, 2012). In Klout’s case, the company calculates a person’s influence, on a scale of 1 to 100, based on her profile and activities on social media, such as the number of Twitter mentions and the number of connections on Facebook. Many companies have used Klout’s scores in their viral marketing campaigns. For instance, in one viral marketing campaign, General Motors identified 900 users who have Klout scores of 50 or higher and offered them 3-day rental of Chevy Volt for free. The campaign resulted in a large number of tweets on Twitter and blog posts about the vehicle, most of which are favorable.

Researchers have been focused on the effectiveness of using various centrality measures to identify influencers on social media. A centrality measure, such as in-degree, out-degree, betweenness, and sender rank, captures a user’s typological characteristics in a social network. For example, a user’s in-degree on a social media site is the number of connections that she has through which information flows from other users to the user under consideration. An example of this is the number of people that a user follows on Twitter. Out-degree is opposite to in-degree and is the number of connections through which information flows from the user under consideration to other users. An example of this is the number of people that follow the user on Twitter. On some social media sites, such as Facebook, information often follows in both directions in a connection. Such a connection will be counted in both in-degree and out-degree. Existing studies found that out-degree, betweenness, and sender-rank are quite effective for identifying influencers in social networks (Cha, 2010; Gosh, 2010; Kiss, 2008).

One shortcoming of existing studies of influencer identification on social media is that the models used to represent the viral marketing process differ from one study to another, making it difficult to compare and generalize findings across studies. For example, regarding the mechanism through which information spreads from one user to another, some studies adopts the informational influence mode (Bampo, 2008; Frenzen, 1993; Kiss, 2008; Watts, 2007), while other studies adopt the normative influence mode (Delre, 2007; van Eck, 2011; Watts, 2007). Regarding structure of the social networks embedded in social media, some studies use the simple random network model (Frenzen, 1993; Watts, 2007), while others use the more complex scale-free network (Albert & Barabási, 2002; Barabási, 1999) or small-world network models (Cha, 2012).

In this paper, we took an integrated approach to developing a model that encompasses a wide