Product Adoption in Online Social Network

Yunjie (Calvin) Xu  
National University of Singapore, xuyj@comp.nus.edu.sg

Cheng Zhang  
Fudan University, zhangche@fudan.edu.cn

Ling Xue  
University of Scranton, xuel2@scranton.edu

Lay Lan Yeo  
National University of Singapore, laylan@alumni.nus.edu.sg

Follow this and additional works at: http://aisel.aisnet.org/icis2008

Recommended Citation
Xu, Yunjie (Calvin); Zhang, Cheng; Xue, Ling; and Yeo, Lay Lan, "Product Adoption in Online Social Network" (2008). ICIS 2008 Proceedings. Paper 200.  
http://aisel.aisnet.org/icis2008/200

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2008 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
PRODUCT ADOPTION IN ONLINE SOCIAL NETWORK
L’adoption de produit dans les réseaux sociaux en ligne

Completed research

Yunjie (Calvin) Xu
Department of Information Systems
National University of Singapore
3 Science Drive 2, Singapore, 117543
xuyj@comp.nus.edu.sg

Cheng Zhang
School of Management
Fudan University
670 Guoshun Road
Shanghai, China, 200433
zhangche@fudan.edu.cn

Ling Xue
Kania School of Management
University of Scranton
Scranton, PA 18503
xuel2@scranton.edu

Lay Lan Yeo
Department of Information Systems
National University of Singapore
3 Science Drive 2, Singapore, 117543
laylan@alumni.nus.edu.sg

Abstract

In product-oriented online social networks, members post product reviews as well as maintaining a list of friends. How to measure the influence a member receives from her friends? This study compares four models to measure the influence based on different theories which make various assumptions of member behavior in a social network. An empirical study was conducted based on 2324 social network members of a cosmetic website. We found the influence from a member’s ego-centric network is best measured by a model that incorporates both the frequency and valence of interactions among members.

Keywords: online social network, product diffusion, product adoption, viral marketing, frequency-rating

Résumé


Introduction

Online social networks such as Myspace.com and Facebook.com have spurred enormous interest among practitioners and researchers. The basic assumption is that consumers in these social networks can be leveraged to promote products for firms. For example, Amazon’s product reviews have been found to influence consumer purchase (Chevalier & Mayzlin, 2006). According to Emarketer.com, 37% of U.S. adults and 70% of online teens...
used online social networks every month in 2007 (Williamson, 2007). Another survey by PowerReviews (Tar
er, 2007) indicated that 65% of respondents used online reviews for product purchase decision making, 78% of them spent more than 10 minutes to read reviews, 86% of them found customer reviews extremely or very important, and 64% of them used online reviews even though they may purchase from other channels. Emarketer.com predicted that the advertising expense on social networking sites in the U.S. will reach $1.6 billion in 2008.

It is not surprising that products spread via word-of-mouth (WOM) (Brooks, 1957). In the offline context, WOM has been extensively studied as means to promote products (Godes & Mayzlin, 2004). For example, Bass’ (1969) product diffusion model is based on the assumption that some consumers (potential adopters) adopt products because of the influence of their friends and direct contacts who have already adopted the product (adopters) rather than the influence of marketers. What is new in the online social network is that the network structure is now observable: Information on who is whose friend is made explicit. An immediate reaction of firms to this golden opportunity is to identify the influential members who often manifest as earlier adopters or well-connected hubs in a social network, so that products can be promoted through them (Keller & Berry, 2003; Kumar & Benbasat, 2006; Rogers, 2003). For example, firms can offer discounted product to them, or motivate them to invite friends by giving them incentives.

However, an aggressive strategy to involve influential product adopters as product ambassadors is not always possible because (1) not all online social networks have a convenient function to email friends, (2) adopters are unwilling to “spam” their friends, (3) when they do so for the sake of monetary incentives, they tend to promote products to weak ties upon which they have little influence, and (4) the online social network platform provider may not allow marketers to do so for privacy concerns.

An alternative strategy is not to ask current adopters to “spam”, but to directly market to the most likely potential adopters, for example, by displaying the most relevant advertisement to the most susceptible members. The notion of stealth marketing is “attempts to present a new product or service by cleverly creating and spreading ‘buzz’ in an obtuse or surreptitious manner” (Kaikati & Kaikati, 2004, pp.6). Instead of aggressively shouting to everybody at the same time, stealth marketing relies heavily on the power of word of mouth to encourage customers to feel they just “stumbled” upon the product or service themselves. For example, the Google AdSense feature customizes the advertisement displayed to a webpage based on the content of the webpage.

Because WOM is a powerful mechanism for stealth marketing, firms need to identify members who are most likely to adopt by quantifying the influence of their social network on them. Motivated by this need, the research question of this study is to propose and compare a set of models that quantify the influence of a member’s friends on her product adoption. This study is important both to firms who are promoting products in online social networks and to the online social network platform provider. An effective model to quantify the influence of one’s friends on her can help firms place more relevant product promotions. It can also be combined with other prediction models (e.g., based on the webpage content, the member’s past purchase etc.) for better prediction.

We address our research question in such a context: In our online social network, members post reviews on products they have adopted. Each member also keeps a list of friends. The friendship relationship is not necessarily bi-directional. When one adds the other as a friend, the reverse is not necessarily true. Rather than being “spammed” with product information, members freely browse the network for product information as well as visiting other members’ profile. Members who have adopted a product are termed as the adopters of that product. Since one is mostly influenced by immediate friends, members who are directly connected to an adopter are potential adopters. In other words, potential adopters are the first-order friends of adopters. We term the remaining members who do not have a direct connection with any adopter non-adopters. This study tries to model the influence of a potential adopter’s ego-centric social network (Figure 1) on her conversion to an adopter in a given time period. Therefore, the dependent variable is whether the potential adopter adopts the product in a particular period of time. Based on various theoretical assumptions, a set of different models are proposed to represent the influence of the ego-centric network. Our empirical study indicates that the influence from a member’s ego-centric network is best measured by a model that incorporates both the frequency and valence (i.e., positivity of WOM) of interactions among members.
Product diffusion in online social network

This study is organized as follows. First, we review the literature related to product diffusion with an emphasis on WOM and studies that focus on product diffusion in a social network context. After that, a set of theories on people’s behavior in social networks are reviewed, which serve as the foundation of our modeling of potential adopters’ adoption behavior. Based on that, we propose a set of models. An empirical study is then reported. We conclude with discussions and implications.

Conceptual Background

Product Diffusion

The product diffusion literature suggests that individual characteristics and network influence are the main categories of factors that influence the adoption of new products (Rogers, 2003). Individual characteristics include innovativeness in product adoption, socioeconomic status and personality variables. Network influence refers to the influence of an individual’s social network on her product adoption. This study focuses on network influence. Network influence can be studied at the dyadic level or at the network level. The stream of research on WOM focuses on the dyadic influence on product diffusion. It explains the mechanism of how one consumer influences another. Another stream of research goes beyond the dyadic level and explicitly focuses on the influence of a larger social network on product diffusion (e.g., Rogers & Kincaid, 1981).

Word-of-Mouth and Product Diffusion

WOM involves informal communications among consumers regarding a brand, a product, a service or an organization (Anderson, 1998; Buttle, 1998). It is seen as a powerful factor in the introduction and diffusion of new products and services (Brown & Reingen, 1987; Chevalier & Mayzlin, 2006; Price & Fedick, 1984). There are several types of WOM marketing, including community marketing whereby people with common interest get together to share their experience, viral marketing whereby entertaining or informative messages are designed to be passed along via electronic means, and referral programs whereby customers are motivated to refer their friends.

A consumer is informed of product information through WOM passively or actively (Rogers, 2003). An individual could passively discover a product in an accidental manner when an adopter casually mentions about it. An active individual continually searches for information and WOM. WOM not only serves to inform a potential adopter with a low cost (Coleman, 1988), but also to help decision making because of the trustworthiness of information in WOM (Silverman, 2001). Besides the demand for WOM by potential adopters, adopters can also initiate WOM. Adopters can be motivated to promote a product when there is benefit (e.g., discount, reference credit) provided by the product seller or when there is a network externality in the use of the product (Subramani and Rajagopalan, 2003). As mentioned above, our focus is on information seeking behavior initiated by potential adopter.
At the dyadic level, the effect of WOM on a potential adopter is a function of both the volume and the valence of WOM (e.g., Mahajan, Muller, and Kerin, 1984; Mizerski 1982, Neelamegham and Chintagunta 1999). Volume measures the total amount of WOM interactions and it has been suggested to create consumer awareness (Godes & Mayzlin, 2004). It has been found that extremely dissatisfied adopters engaged in a greater amount of WOM than satisfied adopters (Anderson, 1998). Valence refers to the nature of the information communicated, whether it is positive or negative. Positive WOM is expected to improve consumers’ attitude towards a product and negative WOM worsens it (Arndt, 1967). A recent study by Chevalier and Mayzlin (2006) found that a negative WOM impacts consumers more than a positive WOM. Liu (2006) examined both volume and valence in his study and found that volume is the dominating factor in predicting box office revenue.

However, consumers depend on WOM to different degrees. Opinion leaders are individuals from whom other people (opinion followers) ask for information and they have the ability to exert a disproportionate degree of influence on people’s behavior in certain areas (Rogers & Cartano, 1962). In general, opinion leaders are more innovative than their followers (Rogers 2003). This characteristic has been confirmed in many different contexts such as cosmetics (Coulter et al. 2002), fashion (Summers 1970), household appliances, travel, politics and automobile (Myers and Robertson, 1972). The number of opinion followers is often used as a measure to indicate one’s opinion leadership (Weimann, Tustin, Vuuren, & Joubert, 2007). Significant correlation between the number of opinion followers and self-reported opinion leadership were observed in past studies (Rogers & Cartano, 1962; Weimann et al., 2007).

While WOM explains dyadic product information transmission, it has a macro level ramification on product diffusion. The famous Bass model of product diffusion (Bass, 1969; Bass et al. 2000) assumes that consumers adopt a produce through two factors: personal innovation and WOM. The former includes the influence from mass media and marketers’ persuasion effort. However, the Bass model does not investigate the individual adoption behavior. In addition, it presumes perfect social mixing, that is, everyone interacts with all the rest (Granovetter, 1978; Van den Bulte & Lilien, 1997). Obviously, consumers often have only a limited exposure to other adopters as constrained by their social network (Burt, 1987; Rogers, 2003). Therefore, studying the influence from one’s social network helps to open the black-box of the Bass diffusion model, as we shall discuss next.

**Product Diffusion in Social Networks**

A social network is a set of actors (e.g., consumers) connected by a set of ties (Wasserman & Faust, 1994). Ties between actors can be directed (e.g., advice relationship) or undirected (e.g. marriage). In addition, ties can be dichotomous or valued. An ego-centric network in a social network refers to a focal actor, called ego, and a set of alters, who have ties to the ego, and the measurements on the ties from egos to alters and on the ties between alters (Wasserman & Faust, 1994). Ego-centric network is often the unit of analysis in social network analysis (Garton et al., 1997) as we shall adopt in our study.

Past ego-centric social network studies found that ego-centric network exerts significant influence on one’s adoption of new products. Past ego-centric network studies typically used survey method (Valente 1996) to collect information on respondents’ social contacts and relationships. Respondents were more likely to recall strong ties than weak ties. It was shown that shared attitude develops from social proximity which is defined as the distance between individuals. This distance can be measured using the number of nodes that need to be traversed from one individual to another. The line of argument is as follows: When facing an uncertain situation that has no direct answer, ego will approach people whom she knows and discuss the issue with them. Eventually, they share the same perspective of the situation after coming to a common understanding (Burt, 1987).

Social networks can be broadly classified into offline and online. The difference between them is that the latter represents a set of actors and their ties digitally. Recently, there has been a proliferation in the number of online social networking sites and users. In those sites, the core information is user profiles and an explicit representation of relationships (O’Murchu, Breslin, & Decker, 2004). Unlike offline social network, geographical boundaries are no longer a constraint in the development of social proximity (Hampton & Wellman, 2000; Wellman et al. 1996). In the online world, members base their feelings of closeness on shared interests rather than physical proximity, and may even consider each other as their closest friends although they seldom or never met before (Hiltz & Turoff, 1993).

Unfortunately, although there are some studies on product diffusion in offline social networks, there are very few empirical studies on product diffusion in online social network. Table 1 summarizes some representative studies of...
online social network. One salient observation of the literature is that although there are a few theories to explain product adoption, there was no study that has compared the applicability of these theories to online social networks. Based on literature review, we summarize five major theories used to explain consumers’ adoption of product in social networks, based on which, we propose four models to describe the influence of an ego-centric network on ego’s adoption probability, two of which are directly based on past theories and two of which are newly developed by drawing upon the tenets of both epidemic diffusion and WOM research.

**Theories of Behavior in Social Networks**

The literature has proposed the following theories to explain product adoption in a social network: (1) threshold model, (2) personal network exposure model, (3) degree of adopter friends model, (4) network closure model, and (5) epidemic diffusion framework. We shall elaborate on each except for the network closure theory. Network closure theory (Coleman, 1988) postulates that the first-order contacts that are connected with each other can potentially influence the individual more than when the first-order contacts are not connected. However, in our empirical study, because more than half of the potential adopters (57.53%) had only one adopter friend, we omit this theory.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Research context</th>
<th>Major findings / Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Katona, Zubcsek and Sarvary (2007)</td>
<td>• <strong>Product type:</strong> Membership in a Central European social networking site</td>
<td>• The probability that a potential adopter adopts a product is a positive linear function of the number of adopter friends that the potential adopter has. (Degree of adopter friends)</td>
</tr>
<tr>
<td></td>
<td>• <strong>Objective:</strong> Examine membership growth of the site which was based entirely on referral in the first 3.5 years of service</td>
<td>• The probability of adoption increases with the number of edges present between her adopter friends. (Network closure)</td>
</tr>
<tr>
<td></td>
<td>• <strong>Type of WOM:</strong> WOM marketing (Referral programs)</td>
<td>• The incremental probability of adoption increases with the number of edges present between her adopter friends. (Network closure)</td>
</tr>
<tr>
<td></td>
<td>• <strong>Dataset:</strong> 138,964 registered users of the site</td>
<td>• The ability to influence potential adopters decreases as the number of an adopter’s friends increases. (Personal network influencing)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Model based on the degree of adopter friends is shown to be a generalization of the Bass model.</td>
</tr>
<tr>
<td>Hill, Provost and Volinsky (2006)</td>
<td>• <strong>Product type:</strong> Telecommunications service</td>
<td>• Potential adopters who have direct communication with an adopter are more likely to adopt.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Objective:</strong> Examine adoption of a new communications service in a large direct-mail marketing campaign</td>
<td>• Potential adopters who communicated with a greater number of times or longer with existing adopters are more likely to adopt.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Type of WOM:</strong> Naturally occurring</td>
<td>• Adding individual characteristics of potential adopters to social network characteristics does not help to enhance the prediction of adoption.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Dataset:</strong> Undisclosed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• <strong>Potential adopters:</strong> People who received a targeted mail from the firm</td>
<td></td>
</tr>
<tr>
<td>Leskovec, Adamic and Huberman (2007)</td>
<td>• <strong>Product types:</strong> Book, DVD, Music, Video</td>
<td>• Probability of purchase increases with the number of recommendations received by a potential adopter until a saturation point is reached.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Objective:</strong> Examine the impact of an online recommendation referral program on product adoption over a period of 2 years</td>
<td>• Probability of purchase decreases as the number of recommendations from the same person increases.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Type of WOM:</strong> WOM marketing (Referral programs via email and with incentives)</td>
<td>• Probability of purchase decreases when the potential adopter gets more than one recommendation.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Dataset:</strong> Nearly 4 million members of the retailer’s web site</td>
<td>• Average rating of the product does not predict probability of purchasing.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Potential adopters:</strong> Recipients of recommendations</td>
<td>• Probability of purchasing increases with the price of product recommended.</td>
</tr>
<tr>
<td>Trusov, Bucklin and Pauwels (2006)</td>
<td>• <strong>Product type:</strong> Membership in a social networking site</td>
<td>• There are more adopters during weekdays and summer break than on weekends and school time.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Objective:</strong> Examine membership growth of the site over 36 weeks</td>
<td>• The number of adopters does not increase with more media appearances.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Type of WOM:</strong> WOM marketing (Referral programs)</td>
<td>• The number of adopters increases with number of promotion events and WOM. WOM had a greater positive impact on the number of adopters than promotion events.</td>
</tr>
<tr>
<td></td>
<td>• <strong>Dataset:</strong> Undisclosed</td>
<td></td>
</tr>
</tbody>
</table>
Threshold Model

The threshold model of collective behavior assumes that an individual will join an activity only after a certain number or proportion of other individuals in the social system has already engaged in that activity (Granovetter, 1978). A threshold is the proportion of individuals who have already engaged in the activity before the given individual would engage in that activity. Individuals with a low threshold engage in the activity before the majority of other individuals do and become early adopters. In contrast, individuals with a high threshold are later adopters (Rogers 2003).

However, there are two difficulties in the application of the threshold model to product adoption. Firstly, some product adoption may not be directly observable, for example, the use of family planning method (Kohler et al. 2001). Secondly, even if one’s product adoption is observable, it might not be observable to all other people in the social system. Hence, the measurement of threshold should consider only people with whom a potential adopter has direct communication with (e.g., first-order contacts). Valente (1995), thus, suggested defining individual threshold with respect to her ego-centric social network, leading to the personal network model.

Personal Network Exposure Model

Personal network exposure (PNE) considers only a potential adopter’s ego-centric social network which she has direct contact, rather than the entire social system. A study conducted by Valente et al. (1997) found that women were more likely to adopt contraceptives if they perceived that people in their personal network were already using it. PNE relies largely on the idea of connectedness\(^1\). Connectedness is the number of people that the individual is connected to and it measures exposure. When the individual is connected to a large number of adopters, exposure is high. PNE is measured by standardizing the connectedness measure:

\[
\text{Network influence} = \frac{\text{Number of direct contacts who are adopters}}{\text{Total number of direct contacts}}
\]

When the potential adopter’s direct contacts adopt the product, they expose her to the product via WOM. Furthermore, the higher the network exposure, the more likely will the adoption be regarded as a norm, and the higher will be the volume of WOM to the potential adopter. Therefore PNE serves as a measure of network influence. In a model to predict a potential adopter’s adoption probability, if the influence of an ego-centric network is measured with PNE, we call it the PNE model.

Degree of Adopter Friends Model

When PNE measures the influence of an ego-centric network by the percentage of adopters, it assumes the potential adopter has the knowledge of product adoption behavior of all members in the ego-centric network. It is reasonable to argue that the potential adopter does not have full knowledge of others’ adoption behavior. Rather, only the adopter friends are observed by her, maybe because the adopter friends advocate the adoption of a product. This explanation leads to a simpler degree model. For a particular individual, degree refers to the number of people in the ego-centric network (Marsden, 1990; Wasserman & Faust, 1994, p.100). In this study, the degree of adopter friends refers to the number of adopter friends in the ego-centric network. As the number of adopter friends increases, the chances of exposure to the product via direct communication increases. Like PNE, the reasoning of degree of adopter friends relies on the volume of WOM as the underlying mechanism of influence. Katona et al. (2007) found that the likelihood of a potential adopter joining a social networking site increases with the number of adopter friends. Mohammed (2001) also found similar results in a study on listeners and non-listeners of a radio soap opera.

\(^1\) The usage of the term connectedness here is not to be confused with its definition in graph theory. In the latter, connectedness measures the degree to which nodes within a graph can be reached from any other node in the same graph (Wasserman & Faust, p.109).
Those who had more listener contacts were more likely to be listeners themselves. Thus, we propose that the degree of adopter friends is positively related to the network influence one receives and we call such a model the degree of adopter friends (DAF) model (Equation 2).

\[
\text{Network influence} = \text{Number of direct contacts who are adopters}
\tag{2}
\]

**SIR Framework**

Epidemic diffusion theories are often used to explain product diffusion. Different from the reliance on the amount of exposure as an explanation of adoption, these theories consider the life cycle of a disease with an individual. Susceptible-infectious-recovery (SIR) framework is the classical disease-propagation framework in epidemiology (Bailey, 1975). According to this framework, an individual is first susceptible to a disease. When she gets the disease, she becomes infectious. Finally, she recovers from the disease or dies.

What is interesting to product diffusion is that the SIR framework suggests some insightful factors that may affect how a susceptible is converted to an infectious. This probability of conversion depends on many factors. First, the SIR framework highlights that infection is not so much as the number of friends who were virus carriers, but the degree that one gets into touch with carriers. Second, infection is a function of the virulence of the virus (i.e. the ability of the organism to cause disease). Third, it is a function of the extent to which the person’s own biological and biochemical defenses can deal with the incoming organisms. Applying the above concepts to product adoption, the disease is a product and the infectious contact is a potential adopter’s visit to an adopter. The final adoption decision will also be affected by personal considerations, i.e., the “resistance” factors.

**Frequency model**

Based on the SIR framework, we propose a frequency model to quantify the network influence by the frequency of contacts rather than by the number of adopter friends. The influence from the ego-centric network hinges on the total number of times that a potential adopter comes into contact with all her adopter friends per unit of time. If a potential adopter frequently comes into contact with adopters in her ego-centric network, she is more likely to adopt the product. This reasoning is more refined than the degree model in the sense that it considers not only who adopter friends are, but also how close the relationships are. In the online context, the infectious contacts are one’s visit of other’s web page where product reviews are posted.

The frequency model is not only consistent with the SIR framework, but also with the WOM research. The effect of WOM on a potential adopter is a function of both the volume and the valence of WOM (e.g. Mahajan, Muller, and Kerin, 1984; Mizerski 1982, Neelamegham and Chintagunta 1999). Frequency of contact better corresponds to the volume of WOM than the absolute number or relative proportion of adopter friends. For the \( n \) adopter friends in a potential adopter’s network, we could measure the influence of an ego-centric network as:

\[
\text{Network influence} = \sum_{i=1}^{n} \frac{\text{Number of visits to } i^{th} \text{ adopter's profile in a period}}{\text{Number of days in the period}}
\tag{3}
\]

**Frequency-rating model**

The frequency model assumes the virus in every contact is potent. However, this assumption may not hold in online social networks where consumers can pass around both positive WOM and negative WOM. If a positive WOM helps spread a product, a negative one discourages it. Based on the SIR framework and WOM literature, we propose that a network influence model should factor in both the frequency of contact and the valence of contact (e.g. Mahajan, Muller, and Kerin, 1984; Mizerski 1982, Neelamegham and Chintagunta 1999). Chevalier and Mayzlin (2006) found that sales in Amazon.com were improved when books were rated more highly, although such effects are not always present (Liu, 2006). When product rating is factored into the influence of an ego-centric network, we call such model the frequency-rating model (Equation 4).
Product diffusion in online social network

\[
\text{Network influence} = \sum_{i=1}^{n} \frac{\text{Number of visits to } i^{th} \text{ adopter's profile in a period} \times \text{Rating}}{\text{Number of days in the period}}
\]  

Comparing the four models, the PNE model and the degree of adopter friends model both rely on the static network structure more than the interaction among members. In contrast, the frequency model adopts a dynamic view of member interactions on top of the structural constraints. The frequency-rating model further factors in the valence of interaction. The latter two are closer to the actual member behavior in a social network and could be expected to predict better in product adoption behavior.

Research Methodology

Data Collection

To compare our models, we examined the adoption behavior of members of a product-oriented cosmetics online social network in Taiwan, UrCosme.com. The primary purpose of the site was for people to share their experience in using various cosmetics products. At the time of writing, the site had about 120,000 registered members, 16,000 products and 250 brands. A major difference between the site and other product review sites was that it incorporated the notion of social network and enabled members to link up with each other. We consider cosmetics as appropriate products for testing our models because they are experience goods. This increases the need for people to turn to their friends who have tried the product before for advice to reduce the risk of purchase. UrCosme.com had three major categories of products – face care, cosmetics and body care. Under face care, there are 19 subcategories, which included make-up removal product, facial cleansing product, toner, face mask, eye care and lip care. Similarly, there were many other subcategories for cosmetics and body care.

Instead of studying the diffusion of one product, we focused on the diffusion of a subcategory of products. We define a subcategory of related products as products of the same brand name and are complementary, for example, facial cleansing product, toner, and face mask. Each product in a subcategory may have variations to serve different skin types. Product subcategory is used as the unit of analysis also because the WOM effect can naturally spill over to related products.

For this study, one product subcategory with 9 products of the same brand was chosen. In this subcategory, there were two popular products with more than 600 reviews; therefore it was not dominated by one product. For the rest products, there were at least 70 reviews each.

Potential adopters were identified as follows. First, adopters of a product subcategory were operationalized as those who wrote a review of any product in the subcategory. Based on the lists of reviews for all products in the subcategory, we compiled a list of adopters. Then, potential adopters of a product subcategory were operationalized as those who were the first-order friends of at least one adopter but had not adopted any product in the subcategory. We collected the friends of these adopters who had not adopted any products in the subcategory. A member could have three types of friends, namely link-in friends, link-out friends, and mutual links. All three types of links between adopters and potential adopters were included. Finally, for potential adopters, we further collected their friends of all three types. A potential adopter’s friends often were non-adopters.

A potential adopter was considered converted when she submitted a review of product in the subcategory or added the product to her wish list. In order to build a model of network influence, we needed to monitor potential adopters for a period of time and observe their conversion. In this study, we adopted a look-back strategy. Based on the time stamp for each link and adoption, we reconstructed the network of adopters and potential adopters 210 days back to the point of data collection. Because a significant portion of members seemed to be students, we chose 7 months as our prediction period to cover one semester and two holidays when students often had more time to browse and purchase. Potential adopters and adopters in the reconstructed network served as our sample. We measured the influence of ego-centric network based on the data 7 months before the data collection. We term the 7 months the prediction period and the time before that the observation period. The observation period was not necessarily the same for members because they might join the network or adopt the product at a different time.
A crawler was developed to collect the data. As a result, 27 (1.15%) adopters were discarded as the web crawler encountered errors in parsing web pages. The collected social network for the target product subcategory comprised of 5,012 members and 23,507 arcs (excluding arcs between non-adopters). There were 1,474 adopters, 2,324 potential adopters and 1,214 non-adopters just before the prediction period.

**Operationalization of Network Influence Models**

Given the different ways to measure the influence of an ego-centric network on a potential adopter’s product adoption, we propose a more general logistic regression model to predict one’s probability of adopting a product in a certain period of time. Our model is:

\[ L_{it} = \ln \frac{p_{it}}{1 - p_{it}} = f(\text{network, innovation, control, } \varepsilon) \]  

(5)

where:

- \( p_{it} \) = potential adopter \( i \)’s probability of adopting the product in time period \( t \), i.e., the prediction period,
- \( L_{it} \) = potential adopter \( i \)’s logit of adopting the product in time period \( t \),
- \( \text{network} \) = the influence of an ego-centric network as measured by Equation 1-4,
- \( \text{innovation} \) = individual innovativeness in product adoption,
- \( \text{controls} \) = control variables, and
- \( \varepsilon \) = error term.

The network effect was calculated based on Equation 1-4 corresponding to different theories. Personal innovativeness was approximated by number of link-ins a potential adopter had at the end of the observation period. We included individual characteristics as control variables. The control variables comprised of skin type, age, number of days as a member, and brand preference. We ignored gender because the majority was female.

A potential adopter’s brand preference \( p \) for a brand was calculated as follows:

\[ p = \frac{\sum_{i=1}^{n} \text{Rating } i}{\text{total count of reviews}} \]  

(6)

In other words, for each potential adopter and for each brand, the brand preference was the average rating of the brand. Subcategory rating was calculated in a similar way covering only the 9 products of interest. All independent variables were based on the data in the observation period.

The actual conversion of a member to adopter was based on whether she posted a review or added a product to her wish list in the predication period.

**Data Analysis**

The average age of the 2,324 potential adopters was 23.16 (SD=8) years old and the average number of days as a member before the prediction period was 478.61 (SD=282). A potential adopter was linked to an average of 8.05 friends. In addition, each potential adopter was linked to an average of 2.95 adopter friends. A potential adopter made an average of 0.1140 visits per day to adopters’ pages during the observation period. At the end of the prediction period, 79 potential adopters had adopted the product, representing 3.40% of the total potential adopters. Figure 2 illustrates the distribution of total number of friends members kept. It illustrates the typical long tail phenomena.
We first fit our models to the whole dataset of 2,324 potential adopters. Table 2 reports correlations among variables and Table 3 reports the regression result. It indicates that although PNE was significant, the sign was not correct. Only the frequency-rating model discovered a significant positive network influence on product adoption. The rest models found no significant network influence. The degree of link-in as an indicator of personal innovativeness showed significant effect in all models. Age and days as a member had a significant negative effect. Brand preference and skin types were insignificant.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Brand preference</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Age</td>
<td>-0.103**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Days as member</td>
<td>0.026</td>
<td>0.073**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Link-in</td>
<td>0.122**</td>
<td>0.082**</td>
<td>0.154**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) PNE</td>
<td>-0.166**</td>
<td>-0.046</td>
<td>-0.233**</td>
<td>-0.242**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) DAF</td>
<td>0.212**</td>
<td>-0.016</td>
<td>0.020</td>
<td>0.421**</td>
<td>-0.220**</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) Frequency</td>
<td>0.093**</td>
<td>-0.018</td>
<td>-0.116**</td>
<td>0.173**</td>
<td>-0.099**</td>
<td>0.566**</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>(8) Frequency-rating</td>
<td>0.097**</td>
<td>-0.019</td>
<td>-0.116**</td>
<td>0.174**</td>
<td>-0.099**</td>
<td>0.569**</td>
<td>0.996**</td>
<td>1.000</td>
</tr>
<tr>
<td>(9) Adopt</td>
<td>0.090**</td>
<td>-0.064**</td>
<td>-0.072**</td>
<td>0.045**</td>
<td>-0.032</td>
<td>0.040</td>
<td>0.076**</td>
<td>0.080**</td>
</tr>
</tbody>
</table>

*p<0.05, **p<0.01.

Was the weak network effect due to the existence of inactive users? Because both PNE model and DAF model assume all potential adopters were active, the network effect so measured might be upward biased if some potential adopters were actually inactive. In contrast, the frequency and frequency-rating model would automatically factor in the inactivity of some potential adopter, hence were less likely affected. To explore such possibility, we re-checked the dataset. We defined a potential adopter to be inactive if she did not write any product review nor visit any of her friends for the past one year before the prediction period. There were a total of 363 inactive potential adopters, which comprised of 15.62% of the total number of potential adopters. Removing such inactive potential adopters, we had 1961 potential adopters. The correlation among variables is reported in Table 4.
### Table 3. Fitted logistic regression models for all potential adopters

<table>
<thead>
<tr>
<th>Variables</th>
<th>Baseline</th>
<th>PNE</th>
<th>Degree</th>
<th>Frequency</th>
<th>Frequency-rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand preference</td>
<td>0.749</td>
<td>0.69</td>
<td>0.755</td>
<td>0.731</td>
<td>0.727</td>
</tr>
<tr>
<td>Skin type dummy (1)</td>
<td>17.963</td>
<td>18.048</td>
<td>17.964</td>
<td>17.919</td>
<td>17.915</td>
</tr>
<tr>
<td>Skin type dummy (2)</td>
<td>17.786</td>
<td>17.904</td>
<td>17.786</td>
<td>17.755</td>
<td>17.75</td>
</tr>
<tr>
<td>Skin type dummy (3)</td>
<td>17.42</td>
<td>17.467</td>
<td>17.422</td>
<td>17.336</td>
<td>17.327</td>
</tr>
<tr>
<td>Skin type dummy (4)</td>
<td>17.718</td>
<td>17.794</td>
<td>17.717</td>
<td>17.692</td>
<td>17.687</td>
</tr>
<tr>
<td>Age</td>
<td>-0.099**</td>
<td>-0.099**</td>
<td>-0.099**</td>
<td>-0.098**</td>
<td>-0.098**</td>
</tr>
<tr>
<td>Days as member</td>
<td>-0.002**</td>
<td>-0.002**</td>
<td>-0.002**</td>
<td>-0.001**</td>
<td>-0.001**</td>
</tr>
<tr>
<td>Link-in</td>
<td>0.032**</td>
<td>0.028**</td>
<td>0.033**</td>
<td>0.027*</td>
<td>0.027**</td>
</tr>
<tr>
<td>PNE</td>
<td></td>
<td></td>
<td>-0.76*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAF</td>
<td></td>
<td></td>
<td>-0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
<td>0.259</td>
<td></td>
</tr>
<tr>
<td>Frequency-rating</td>
<td></td>
<td></td>
<td></td>
<td>0.054*</td>
<td></td>
</tr>
<tr>
<td>Negelkerke R²</td>
<td>0.089***</td>
<td>0.094***</td>
<td>0.089***</td>
<td>0.093***</td>
<td>0.094***</td>
</tr>
<tr>
<td>-2Log likelihood</td>
<td>636.053</td>
<td>632.793</td>
<td>636.025</td>
<td>633.359</td>
<td>632.975</td>
</tr>
</tbody>
</table>

*p<0.05, **p<0.01, ***p<0.001.

### Table 4. Correlation among variables for active potential adopters

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Brand preference</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Age</td>
<td>-0.095**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Days as member</td>
<td>0.066**</td>
<td>0.058*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Link-in</td>
<td>0.107**</td>
<td>0.085**</td>
<td>0.178**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) PNE</td>
<td>-0.153**</td>
<td>-0.052*</td>
<td>-0.272**</td>
<td>-0.244**</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) DAF</td>
<td>0.202**</td>
<td>-0.014</td>
<td>0.067**</td>
<td>0.421**</td>
<td>-0.226**</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) Frequency</td>
<td>0.084**</td>
<td>-0.016</td>
<td>-0.099**</td>
<td>0.173**</td>
<td>-0.099**</td>
<td>0.561**</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>(8) Frequency-rating</td>
<td>0.088**</td>
<td>-0.016</td>
<td>-0.099**</td>
<td>0.174**</td>
<td>-0.099**</td>
<td>0.565**</td>
<td>0.996**</td>
<td>1.000</td>
</tr>
<tr>
<td>(9) Adopt</td>
<td>0.081**</td>
<td>-0.060**</td>
<td>-0.055*</td>
<td>0.044*</td>
<td>-0.038</td>
<td>0.034</td>
<td>0.073**</td>
<td>0.077**</td>
</tr>
</tbody>
</table>

Table 5 reports the fitted logistic regression models based on active potential adopters. The result indicated that the frequency-rating model is still the only model that indicated a significant positive effect. In summary, it seems that the frequency-rating model is the best to measure the influence of an ego-centric network on product adoption.

### Discussion and Implications

In online social networks where one can easily add links to other members, if links represent friendship, such friendship is “cheap” and not very suggestive of mutual influence. If a marketer is to utilize links to locate potential product adopters, a more accurate measure of influence is needed. In this study, we compare four measures of network influence based on different behavior theories in social network. Among them, the frequency-rating model makes two more realistic assumptions than other models: (1) links can be either active or inactive, and (2) not all WOM can be either positive or negative.
Table 5. Fitted logistic regression models based on active potential adopters

<table>
<thead>
<tr>
<th>Variables</th>
<th>Baseline</th>
<th>PNE</th>
<th>Degree</th>
<th>Frequency</th>
<th>Frequency-rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand preference</td>
<td>0.653</td>
<td>0.596</td>
<td>0.659</td>
<td>0.635</td>
<td>0.631</td>
</tr>
<tr>
<td>Skin type dummy (1)</td>
<td>18.038</td>
<td>18.172</td>
<td>18.039</td>
<td>17.99</td>
<td>17.986</td>
</tr>
<tr>
<td>Skin type dummy (2)</td>
<td>17.991</td>
<td>18.175</td>
<td>17.99</td>
<td>17.958</td>
<td>17.953</td>
</tr>
<tr>
<td>Skin type dummy (3)</td>
<td>17.59</td>
<td>17.686</td>
<td>17.593</td>
<td>17.498</td>
<td>17.488</td>
</tr>
<tr>
<td>Skin type dummy (4)</td>
<td>17.885</td>
<td>18.026</td>
<td>17.884</td>
<td>17.858</td>
<td>17.853</td>
</tr>
<tr>
<td>Skin type dummy (5)</td>
<td>17.018</td>
<td>17.156</td>
<td>17.02</td>
<td>16.978</td>
<td>16.968</td>
</tr>
<tr>
<td>Age</td>
<td>-0.092**</td>
<td>-0.093**</td>
<td>-0.093**</td>
<td>-0.091**</td>
<td>-0.091**</td>
</tr>
<tr>
<td>Days as member</td>
<td>-0.001*</td>
<td>-0.001**</td>
<td>-0.001*</td>
<td>-0.001*</td>
<td>-0.001*</td>
</tr>
<tr>
<td>Link-in</td>
<td>0.029**</td>
<td>0.029*</td>
<td>0.03*</td>
<td>0.024*</td>
<td>0.024*</td>
</tr>
<tr>
<td>PNE</td>
<td>-0.048*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree of adopter friends</td>
<td>-0.003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td>0.259</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency-rating</td>
<td></td>
<td></td>
<td></td>
<td>0.054*</td>
<td></td>
</tr>
<tr>
<td>Negelkerke $R^2$</td>
<td>0.069***</td>
<td>0.077***</td>
<td>0.069***</td>
<td>0.074***</td>
<td>0.075***</td>
</tr>
<tr>
<td>-2Log likelihood</td>
<td>636.053</td>
<td>594.512</td>
<td>598.562</td>
<td>595.883</td>
<td>595.508</td>
</tr>
</tbody>
</table>

PNE and DAF models suffer the unrealistic assumption that all links are active. Besides the presence of inactive potential adopters, even active potential adopters were not exploring all their relationships. In our dataset, only 3,476 (47.51%) out of the 7,316 links between all potential adopters and adopters were activated by the potential adopters after the establishment of links. In that same period, only 9,129 links (45.74%) were activated out of the 19,958 links that were present between all potential adopters and their friends. The large number of inactive links explains why some measures of the network effect such as PNE and DAF were less effective.

Our results revealed that PNE is significant but of opposite sign. Higher proportion of adopters in the ego-centric network in fact decreased the odds of adoption. The ineffectiveness of PNE was probably due to the long tail nature of online social network. The majority of the potential adopters were loners with very few friends. We suspect that loners were likely to have high PNE. To test it, for potential adopters, we calculated the correlation between one’s total number of friends and the number of adopter friends, which was 0.936. In contrast, the correlation between total number of friends and PNE was -0.349. This difference suggests that for loners who tended to have few total friends, their PNE could instead be high. This was not because they had more adopter friends, but because they had a smaller denominator. In other words, active adopters were likely to link to many loners. The presence of such loners made PNE an ineffective indicator of network influence one was subject to. This observation contradicts the offline social network studies where PNE was found effective. PNE was effective in offline studies because subjects were more likely to recall only strong and active links whereas in online contexts they keep a large number of weak or inactive ties in the ego-centric network.

The frequency model suffers an unrealistic assumption of positive interaction. As the research in WOM has suggested, WOM has both volume and valence aspects. This is a plausible reason for its inability to manifest a positive network effect.

**Theoretical Implications.** The comparison of four different ways to measure influence of an ego-centric network on product adoption offers important theoretical implications for future research on online social network. First, it shows that the frequency-rating model is probably the best behavioral model among the four to explain the mechanism in online social networks. While all four models have their root in offline social network studies, it is inadequate to assume that offline behavioral theories could be directly applied to online social networks. In the context of product diffusion in online social network, we show that the often effective DAF or PNE model suffer unrealistic assumptions of member behavior in the online setting. We show that the traditional volume and valence explanation of WOM remains effective in the online setting as manifested via the frequency-rating model.
Second, this study suggest that social influence in online social network is less likely to be determined by the network structure one is embedded, but rather the dynamic interactions on top of the network. In other words, it is the activation of one’s social network or part of it that produces an impact, not the network itself.

Third, although the frequency-rating model has its root in both epidemic diffusion framework and WOM research, it extends them. For example, epidemic diffusion framework ignores the valence of contacts. While the WOM research recognized both volume and valence, it has not been tested in an ego-centric network, neither in an online social network context. Our study bridges these gaps.

**Practical Implications.** This study offers a few important practical implications. First, this study proposes a new way of marketing in online social networks, i.e., the identification of potential adopters for targeted marketing. Past literature adopted a pushy strategy that relied on influential adopters. Ours does not. Rather, it offers a practical method for stealth marketing in e-commerce setting by utilizing online social networks. The proposed models, especially the frequency-rating model, can be used for targeted advertising and product recommendation agent design.

Second, this study suggests what information the platform provider of an online social network needs to capture and keep. Some online social networks do not record the interaction among members to reduce the burden of data storage on the system. This study demonstrates that interaction information is invaluable to marketers, more important than the static links. At the same time, our study also suggests a practical way for data reduction. After the recording of member interaction for a certain period, the platform provider can consolidate the relationship between each pair with a single influence score, which can be easily stored and integrated with later observations.

However, the above findings and implications should be interpreted within its limitations. First, we defined adopters as those who had posted a review of a product. This set of adopters underestimated the true number of adopters as some might have adopted the product but did not post a review. Similarly, this study overestimated the number potential adopters as some of the potential adopters might have adopted the product but did not submit any review. Second, given the relatively large sample size, although the frequency-rating model indicated significant network effect at 5% confidence level, this significance should be considered weak. Third, we chose only one product subcategory due to resource constraints. Our study is also confined to one social network. Generalization to other product types and other online social networks should be treated with caution. Especially, besides the difference between search and experience products, products could differ in the financial risk involved, social risk, and purchase cycle, which make them subject to different degrees of social influence. Finally, the choice of prediction period was also subjective.

**Conclusion**

In summary, with an objective to compare different ways to measure the influence a member receives from her ego-centric social network regarding a product, this study explored four different measures: the personal network model, the degree of adopter friends model, the frequency model, and the frequency-rating model, based on various behavioral theories of social network. Our empirical study of 2,324 online social network members in a cosmetic network indicated that the frequency-rating model found the most significant social network influence. It suggests that the frequency-rating model might be a better behavioral model than other models in explaining product adoption in a social network. This model also provides an alternative to explain the formation and dynamics of the product diffusion network in online social networks.

**Acknowledgement**

This project is supported by School of Computing, National University of Singapore, research grant R253-000-057-112.

**References**


