Influence of Social Networks in Target Marketing

Dept. of CIS - Senior Design 2009-2010

Aman Nalavade
nalavade@seas.upenn.edu
Univ. of Pennsylvania
Philadelphia, PA

Lyle Ungar
ungar@seas.upenn.edu
Univ. of Pennsylvania
Philadelphia, PA

Vanditha Srungavarapu
vanditha@seas.upenn.edu
Univ. of Pennsylvania
Philadelphia, PA

Shawndra Hill
shawndra@wharton.upenn.edu
Univ. of Pennsylvania
Philadelphia, PA

ABSTRACT

Online social networks such as Facebook have shown a huge increase in the number of users. In this work, Facebook will be used as a dataset to see how social networks can be used to provide better target marketing strategies. In our work, a Facebook application was built and products were recommended by users to friends over their social networks. These recommendations are considered to be a form of advertisement. It was shown that: (1) Users are more likely to search for and recommend products with high brand equity, whereas acceptance of a recommendation depends upon complexity of the product that was recommended. (2) Genuine recommendations that are sent out after considering product characteristics such as newness, complexity, risk, and brand equity, have a higher acceptance rate as compared to random recommendations or banner advertisements. This research is important because it will enable marketers to provide more effective marketing strategies by showing that product recommendations play a vital role in increasing the product acceptance rate. It also helps them evaluate the products that do and do not work well over social networks by looking at what product characteristics cause that product to be recommended.

1. INTRODUCTION

A social network is a network consisting of individuals who are connected to other individuals through social links such as friendship, professional or family links. Social networks have always been of interest to researchers as they encode how people interact with each other. Previously, social networks were limited by proximity as forming bonds with people who were located further away required more effort than forming bonds with proximal people. With increased access to the Internet, social networks have become more prominent and diverse because people can now connect more easily with less effort.

Although people can have a diverse set of connections, it is usually seen that many of these connections tend to be homophilous [9]. Homophily is defined to be the tendency for individuals to associate with people of similar interests and demographics such as age, gender, hometown, etc [9]. This property of networks has been of interest to marketers because similar people tend to have similar interests and thus have similar purchase patterns. Marketers use this property to predict the product adoption probability, which is the probability of a user’s friend considering purchase of a given product assuming that the user recommends that product. Network-based marketing uses this concept of product adoption probability and provides a marketing technique where social networks and the links between customers are used to improve marketing and sales [8]. A link is defined as a tie between two individuals based on one or more types of interdependencies such as friendship, kinship, professional, etc. Facebook is primarily a social network consisting of friends and family where users can add friends, send them messages, write on their walls, and join applications [10]. In Facebook, a link is formed between two users if one accepts the “friend request” of the other [10]. Target marketing on the other hand, does not use a social network and markets a given product to a specific group of people who can be defined possibly by age, gender, or race.

A recommendation is a form of advertisement where one user shares the product information with other users and suggests that product to them. A random recommendation in this context is when a user randomly recommends a product to different friends, but those friends are not aware that the recommendations are random and thus are perceived to be genuine by them. A genuine recommendation in contrast is when a user recommends a selective product to a specific friend, after considering the friend’s probable interest in that product. In this case, the product characteristics will play a role in determining which product is picked for the recommendation. In a recommendation, the recommender or the user suggests products to the recommendee or the user’s friend. Click rate is defined as the probability that a user will click on and accept a recommendation. This metric will be used to test if recommendations provide a better form of marketing.

Product characteristics form an important role in determining whether a product is recommended over a social network. The product characteristics that were taken into account in the analysis are defined below: (1) Risk is defined as the monetary amount that the purchaser pays in order to buy or use the product [2]. For example, purchasing electronics has higher risk than purchasing groceries. (2) Complexity is the difficulty that the user faces while using of the product and is defined as whether the user needs prior
knowledge of how the product works [2]. Automotive tools have high complexity due to its need of specialized knowledge. (3) Involvement refers to the capital value of a good, where a high involvement good is purchased only after long and careful consideration such as a car or a laptop [2]. (4) Brand equity is a brand’s power derived from name recognition that it has earned over time, and thus gives higher sales volume [2]. (5) Experience products are those which are difficult to observe in advance but can be evaluated upon consumption [2]. For example, music albums are experience goods whose value is determined only after listening to it. (6) Newness refers to how different the product is compared to the already existing products [2].

Sociologists and marketers are interested in this work because it looks at the interactions between social networks and products. This will enable them to understand about what products get recommended over social networks and thus help provide better marketing strategies. This will also give insights into how people behave over a network with respect to products. Computer scientists play a major role in this domain by building the application to collect large sets of data, store the data efficiently in a database, then analyze the data and finally build a predictive model to provide better marketing strategies.

The purpose of this work is to look at social networks and understand how they function and how they can be used for marketing. Social networks can provide a better marketing tool when certain factors such as the characteristics of a product, the user’s demographic data and the user’s interest data are considered. Using this additional data, the acceptance of a recommendation can be increased significantly. The characteristics of a product like newness, risk, complexity, etc. form an important part in determining whether the advertisements for that product will be accepted and also whether the product will be recommended to other friends.

The project is essentially divided into two different sections: Data acquisition where the data is collected by creating a Facebook application, and data analysis which entails looking at the data collected and building a model to provide effective target marketing strategies.

2. RELATED WORK

2.1 Homophily

Social networks have been found to display homophily. Homophily in general can be based on a wide range of attributes like gender, age, and hobbies. Homophily has been found to exist in a vast majority of social networks and can be used for target marketing.

The earliest studies of homophily looked at small social groups, where an observer could see the links between the members [9]. The observations were based on behavioral cues as well as explicitly stated data from the individuals about their close friends. The first evidence of homophily came from informal ties between school children, college students and small urban neighborhoods. These initial studies showed homophily based on demographics such as age, sex, race and education [9].

During the 1970s and the 1980s, the study of homophily changed due to the use of modern sample surveys. Recent research looks more at the organizational contexts of networks. There are two kinds of homophily: status homophily in which similarity is based on social status or one’s position in society and value homophily where similarity is based on values, attitudes and beliefs and does not depend on social status [9]. This work will focus mainly on value homophily to figure out whether people tend to recommend products to other people who are homophilous with respect to their values. Another important factor affecting homophily is proximity. With the use of the Internet, although proximity is not crucial in forming ties, it has been found that residential proximity is the single best predictor of how friendships are formed [9].

2.2 Tie Strength over Social Networks

In addition to homophily, relationships need to be measured in terms of tie strength so that strong ties can be distinguished from weak ties in order to better use social networks to market products. Tie strength is measured by four strength dimensions: (1) the amount of time spent together, (2) the emotional intensity, (3) the intimacy and (4) the reciprocal services between the users [5]. A study [5] looked at a dataset of 2000 Facebook users and classified their friendships as strong or weak ties with an accuracy of 85%. This tie strength has been used for studying individuals and organizations to figure out how social networks work. Banks that form the right proportion of strong and weak ties with other firms get better financial deals. Weak ties on the other hand also benefit certain groups of people such as job-seekers [5].

2.3 Network-Based Marketing

This use of links gives rise to network-based marketing. There are three different modes of network-based marketing: explicit advocacy wherein individuals openly recommend certain products to their friends or acquaintances, implicit advocacy wherein individuals do not explicitly recommend products but they do it implicitly through their actions such as adopting the product. The third mode of network-based modeling is network targeting wherein the firm targets the prior purchasers’ network based neighbors [8]. A key assumption that network-based marketing makes is that consumers propagate only positive information about the products that they buy and no negative information [8]. Using these assumptions, the firm will value a certain group of people more highly as compared to the others as they are the influencers. These influencers promote the product better than other groups of people [8]. Prospective targets in this type of marketing are found based on demographic data and customer relationship data. It was found that network neighbors, or consumers that are linked to a previous customer, usually adopt a service 3 to 5 times higher than the control groups that were not related in a network. It was also found that analyzing the network allowed firms to acquire new customers who otherwise would not have been recognized using the traditional marketing techniques [8].

2.4 Viral Marketing

Another interesting aspect of social networks is the ability to look at viral marketing. Viral marketing is defined as the use of pre-existing social networks to increase brand awareness and sales using self-replicating viral promotions such as video clips, images, interactive Flash games, etc [1]. Given a model of a social network, there is a well-defined optimization problem that chooses the set of customers to
market to, so as to maximize net profits. This has been shown to be an NP-Hard problem but can be approximated quite well [1]. With the use of online social networks, tracking this problem becomes much easier as the interactions between people are known. In addition, if one could segregate the users, a company could market to those people who are more influential.

Companies are finding that traditional marketing is in crisis, because customers are increasingly becoming habituated with the traditional means of advertisements and hence are not paying attention to them. In spite of this, some companies like Amazon and Google succeed with virtually no marketing because of word of mouth and network effect wherein the value of the product increases as more people use it. A recent study found that positive word of mouth among customers is by far the best predictor of a company’s growth [1]. The problem that these marketers now face is to figure out how to do this effectively and to understand how viral marketing works. There are still many questions regarding which products work well with this type of marketing. While this is known at a high level for some product segments, many startup companies have failed because they invested heavily in creating network effects that never materialized [1].

2.5 Product type

A target based marketing system predicts the content that is relevant to a given person, and advertise only this information to that person. To make those predictions, the systems take in large amounts of data about the user and filter it based on previous product interaction history [3]. Another way to find data was to look at a user’s social network. A study that looked at the advantages of using social networks found that by creating sub-groups of people based on interaction data, they were able to increase the quality of advertising [3]. Starting with an entire graph of all the users and then partitioning the users into groups, the researchers saw that their algorithms predicted the library books used by different groups. This algorithm performed quite well as compared to randomly choosing books [3]. Although the study ended there, the valuable information that was found about social networks would enable us to better target products to each of these sub groups. This study however looked solely at library books, and so the extrapolation is quite limited in terms of which products work well in a social network. Also the study used only user interaction data, and did not take into account demographic or basic profile information about the users. The combination of both the social networks as well as user data could prove to be a better advertising system than the use of only network data.

Other studies have also looked into social networks using a wide range of products. A study [6] that looked at movie recommendations saw that the use of networks allowed better recommendations. The algorithm traversed the graph and ranked the “closeness” of users and their friends and then used that information to suggest how much a given user’s friend may be interested in a movie given that the user already saw the movie. This study compared the results with a simple person correlation recommender and saw much better results of product adoption [6]. With this study, the strength of links was derived based on the network itself, and this was a key input in determining the recommendations [6]. This study however, did not look at user profile data and did not create sub groups within the network itself. The study did raise an interesting point about trying to make accurate recommendations when the number of links to a given person is small or close to zero [6]. This often created inaccurate results, and thus the degree of each node (the number of friends for each user) was considered to be an important attribute to consider.

3. SYSTEM MODEL

The primary data set for this work was collected using an application that was created on Facebook. Considering that Facebook has a social network, the information about users’ demographics, interests and friends is readily available and easily accessible [10] and thus it forms the data set for this work.

3.1 Data Acquisition

A Facebook application was built to collect the required data. Facebook allows developers to build custom applications that can automatically collect the required demographic and interest data from the user’s profile. This allows developers to ask other interesting questions such as to recommend a product, rate a product, etc. and not explicitly ask questions about their demographics or interests. Although the benefits of using Facebook to collect this data are clear, the more pressing issue was finding incentives for users to actually use this application.

To overcome this problem, coupon codes were used instead of products. Coupon codes are “codes” that users can apply to certain products in order to receive a financial discount or rebate. Coupons are based on products and have the same product characteristics and also provide the user an incentive to use them, in the form of a discount. This enables users to gain monetary benefits by using the coupon, while providing information about the user’s preferences and the friends to whom the user recommends the products to.

In order to provide users with quality coupons, a company called CouponCabin that had a number of online coupon
codes was contacted. These coupons were then categorized based on the Keller Kay Group’s [4] categorization, which is a common classification used in the marketing community. Some of these categories include apparel, food and dining, home goods, sports, electronics, etc. Each of these chosen categories have unique product characteristics (experience, risk, complexity, involvement, etc). Along with these product characteristics, products are also categorized across different brands, which play an important role in the spread of products.

Users sign up for the Facebook application and their profile information is accessed in the background. Their profile information includes demographics, their interests, their educational background, their work history, etc. Link information regarding the user’s friends is stored as well. Information about the recommendations (product, recommender, recommended users and specific link type between the users) are stored once a user recommends a coupon. In addition, how a user enters the application and the user’s movements throughout the application such as clicking on a specific category, searching for specific coupons, using of a coupon, rating of a coupon, and inviting of friends to use the application are stored. The flow of this back-end processing has been shown in Figure 1. This information will be useful to test our hypotheses.

Once users access the application, they can browse through the different categories and select a category and view all the coupons in the category. The coupons appear in a random order to prevent any influence on the user to pick only the first coupons that were presented on the screen. A user can use the coupon by going to the merchants website, can rate or share the coupons. When a user shares a coupon, the coupon is posted on the friend’s wall and serves as a recommendation. A user also specifies the exact link type (friend, family, or colleague) that characterizes their relationship with that Facebook friend. Finally, the user has the option of searching for coupons, stores and categories as well as seeing the top coupons and the top “recommenders”. Figure 2 shows the user’s movement throughout the application.

### 3.2 Data Analysis

The data from the data acquisition section was used to understand how to market a product over a social network. The two main hypotheses that were tested are: (1) Users are more likely to search for and recommend products with higher brand equity, whereas acceptance of a recommendation depends upon the complexity of the product that was recommended. (2) Genuine recommendations that consider product characteristics such as newness, complexity, risk, and brand equity have a higher acceptance rate than random recommendations or banner advertisements. The steps used in the data analysis have been shown in Figure 3.

In order to test these hypotheses, homophily is a prerequisite and needs to be tested. To show homophily, the rates of “sameness” among different demographic attributes like age, gender, religion, etc. and different interest attributes such as favorite movies, books, music, etc. were calculated for a given user and his friends. Given that homophily exists, and that similar people have similar interests [9], it is now possible to suggest that possibly purchase patterns of these homophilous users could be similar.

Recommendation data is then considered to check if there was increased homophily between a recommender and a recommendee, than compared to a user and his friends. Product characteristics were considered along with the recommendation data to test if they influenced: (1) use of a coupon by a user, (2) coupon recommendation by a recommender, (3) acceptance of a coupon recommendation by a recommendee.

Finally, the click rates of the recommendations posted to the recommendee’s walls were calculated to check if there
was an increased click rate from: (1) banner advertisements, to (2) random recommendations, to (3) genuine recommendations.

Research has looked at specific products to see how they spread over social networks and the results suggest that certain specific products are spread more successfully across networks [9]. Given the productivity of these research results, this work looks at a wide range of and more general product categories and analyzes the product characteristics that cause a given product to be used and recommended by a given user. This will define how or if a social network can be used as a marketing strategy tool depending upon the product. Looking at whether product characteristics influence the acceptance of a recommendation helps understand whether the recommendation was accepted solely based on the user's recommendation or if the type of the product also played a role. This allows marketers to decide which products do and do not work well over a social networks.

4. SYSTEM IMPLEMENTATION

4.1 Data Acquisition

The Facebook application was developed using the Facebook Developer platform, PHP, MySQL, JavaScript and AJAX. The Facebook developer platform that was used consists of an API, a markup language, and a query language. The Facebook API allows access to the profile information of the users who have signed up for the application. FBML which is the Facebook Markup Language provided additional graphical features while the Facebook Query Language or FQL provided analysis and retrieval of the Facebook data in an SQL-style interface. The data collected was stored in a MySQL database which integrates well with PHP and is convenient to use along with a Facebook application.

The Facebook application that was created is called "the couponbook", which is an interactive application allowing users to easily search for and use coupon codes while sharing them with their friends. The coupon codes are pulled in from an XML file provided by the Coupon Cabin, a large independent company which runs its own coupon code website. When a user signs up for the application, the user's demographic data, interests and friend information is accessed by the application automatically by using preexisting queries provided by the FQL. This data was stored in a MySQL database and the main components were stored in the structure shown by Figure 4.

As soon as the user enters the application, a screen shows up that allows him to pick a category and to look for specific coupons in that category as shown by Figure 5. The user can also search for coupons in different brands, stores, or categories of their interest. The searching algorithm, parses the search terms and removes any basic articles such as "the" or "a" or "an". It also considers singular and plural searches as equivalent by aggregating results from both these searches. Additionally, the search term is delimited by the space character and searching occurs based on all the different strings obtained from the search term. The results are displayed based on a store match, a category match or a coupon description match. The user can also look at (1) the coupons that have been recommended the most (2) users who have recommended the most number of coupons.

![Figure 5: Initial Screen where user can pick a category](image)

The user can select a category and see all the relevant coupons in that category as shown by Figure 6. In this page, the user has an option of (1) using the coupon, (2) sharing the coupon with his friends or (3) rating the coupon. To use the coupon, the user can click on the coupon code and a new tab redirects him to the retailer's website where the coupon can be used. The user can rate the validity of the coupon by clicking on the "Like" or "Dislike". As stated in the system model section, these coupons were presented in a random order to prevent skewing of the data. The user also has the option of looking at only the top brands or looking at all the coupons in that category.

The user can share a coupon with his Facebook friend by clicking on the share button as shown in Figure 6 and this is considered to be a "recommendation" as defined earlier. This information is tracked and stored in the database. The user can select up to ten friends to share the coupon and must select the type of link (family, friend, or colleague) that they share with that Facebook friend as shown in Figure 7.

When a user shares a coupon with a friend, a wall post is created on that friend's profile page as shown in Figure 8 and a notification, that a recommendation is made, is sent to them. A wall post is also made on the recommender's wall stating that this user made a recommendation. This wall post contains a link to the application that the friend can use to sign up for the application and to view the coupon code.

![Figure 4: Entity-Relationship diagram representing what data was stored for the application](image)
This allows the spread of the application, thus providing more data than can be analyzed in the data analysis section. These wall posts are also tracked to see how a given user’s friends respond to the recommendations. The user also has a choice of commenting on the recommendation or seeing other coupons that are available in the application.

As the user moves through the application, their movements such as the category they clicked on, the friends with whom they shared the coupon, etc. are tracked. All of this data was stored in a different tables in a MySQL database. Although this information could be essentially stored in different text files, using databases was chosen to be a better option as they allow for easy storage using PHP and also easy retrieval by using SQL queries.

Apart from looking at coupons and using coupons for oneself like in other websites, this application additionally allows: (1) sharing coupons with one’s friends, (2) friends to see which coupons the user recommended to his other friends, thus letting them use the coupon as well. This application is also easy to find and access because it leverages Facebook’s popularity [10].

4.2 Data Analysis

The data that was collected in the data acquisition section using the Facebook application was first tested for homophily. The user and his link information (friends’ information) was the specific data that was used for this analysis. This information was stored in text files and was parsed to find the different attributes of user’s demographic and interest data. The parser was implemented using JAVA. The user’s (1) demographic attributes such as gender, home city, college name, religion, etc. and (2) interests attributes such as books, music, television shows, hobbies, etc. were compared individually with their friends’ attributes. To establish a baseline, users were compared to all existing application users (regardless of whether they were connected by a link). This allowed an evaluation of whether homophily among a user and his friends in fact exists.

Then the recommendation data was analyzed to test for increased homophily specifically to test whether a user’s recommendees are more similar to a user across different attributes compared to the rest of the user’s friends. To implement this process, recommendation data was read from the MySQL database and the text files containing recommendees’ information were parsed using JAVA. This analysis was done for each user and the results were averaged across all of the different users.

To test the hypothesis that product characteristics play a significant role during the recommendation process, risk, complexity, involvement, brand equity, experience and newness were determined for each of the categories as shown in Table 1. Due to the presence of a large number of coupons, each with different product characteristics, coupons within a given category were grouped together to allow for easier analysis. Each of the categories were assigned a rating between 1 and 10 for each of these different product characteristics based on the kind of coupons that were present in that category. These ratings were then standardized by calculating a z-score to prevent skewing towards one of the extreme values. This regression testing was done to check whether these product characteristics were significant for the following events: (1) The number of times coupons in a given category were used, (2) the number of times coupons in a given category were recommended, and (3) the number of times coupon recommendations from a given category were accepted.

Finally, the hypothesis that click rate increases from banner advertisements to random recommendations to genuine
Table 1: Ratings on a scale of 1-10 for each of the different categories based on product characteristics

Table 2: User application entry mode statistics

Table 3: Category clicks statistics

(2) The application was released to the public and by word of mouth many college students started using the application, with students from the University of Pennsylvania being the majority of the sample. The application was also advertised on Facebook and it was seen that the users who signed up through these advertisements were mainly users over the age of 40 years. Figure 9 shows a frequency distribution of the ages of the application users.

(3) It was seen that there were more female users than male users of the application. There were 59% female users and 41% male users. This could be attributed to the fact that more of the coupons were dedicated to women’s interests like apparel, flowers and gifts, jewelry, etc. (4) Users first select a category that matches their interests. Categories such as women’s apparel, electronics and men’s apparel were clicked on the most as shown in Table 3.

(5) Users first look for coupons that interest them and then only recommend those coupons to their friends. (6) Users do not tend to rate all the coupons they use.
Users tend to search for coupons that interest them through the search box. Analyzing their search terms shows that the searching algorithm needs to use a more sophisticated tool like stemming. Searching is a valuable functionality to have as in this case where there are a large number of coupons and it is difficult to find the desired coupon easily.

5.2 Data Analysis

5.2.1 Homophily between users and their friends

Firstly, a given user’s demographic information was tested against all of his friends for homophily. The demographic attributes that were tested were: (a) last name, (b) gender, (c) birthday month, (d) birthday year, (e) home city, (f) home state, (g) home country, (h) religion, (i) current city, (j) current state, (k) current country, (l) college name, (m) college year, and (n) college concentration. As shown in Figure 10 users are more similar to their friends as compared to random users for all of the demographic attributes.

![Figure 10: Homophily seen in application users’ demographic data](image)

<table>
<thead>
<tr>
<th>User’s Demographic Attributes</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>College Name</td>
<td>1.601E-10</td>
</tr>
<tr>
<td>College Year</td>
<td>8.681E-12</td>
</tr>
<tr>
<td>College Concentration</td>
<td>0.0012</td>
</tr>
<tr>
<td>Last Name</td>
<td>0.0066</td>
</tr>
<tr>
<td>Gender</td>
<td>3.222E-07</td>
</tr>
<tr>
<td>Home City</td>
<td>7.981E-08</td>
</tr>
<tr>
<td>Home Country</td>
<td>5.507E-08</td>
</tr>
<tr>
<td>Home State</td>
<td>3.956E-13</td>
</tr>
<tr>
<td>Birthday Year</td>
<td>8.558E-07</td>
</tr>
<tr>
<td>Birthday Month</td>
<td>1.527E-01</td>
</tr>
<tr>
<td>Religion</td>
<td>7.102E-02</td>
</tr>
<tr>
<td>Current City</td>
<td>8.681E-06</td>
</tr>
<tr>
<td>Current State</td>
<td>5.076E-01</td>
</tr>
<tr>
<td>Current Country</td>
<td>5.43E-07</td>
</tr>
</tbody>
</table>

Table 4: z-test to test for significance in user’s demographic data when n = 80

A z-test was conducted to test for significance and the results are shown in Table 4. The p-value is significant (less than 0.05) for most of the demographic attributes except birthday month, religion, and current country. Birthday month was expected not to be significant as a given user does not tend to make friends based on similarity of their birthday month. Similarly, religion also did not show a significant difference, but its significance was greater than that of birthday month. Therefore, this shows that homophily does in fact exist among users and their friends when demographic attributes are considered.

Next, a given user’s interest information was tested against all of his friends for homophily. The interest attributes that were tested were: (a) hobbies, (b) books, (c) music, (d) television shows, and (e) movies. As shown in Figure 11, users were more similar to their friends as compared to random users for all of the interest attributes.

![Figure 11: Homophily seen in application users’ interest data](image)

However, when a z-test was conducted to test for significant differences among the attributes, only television shows showed a significant p-value (less than 0.05) as shown in Table 5. This could be because larger number of users list television shows in their profile as compared to music and hobbies. Therefore, although users’ interest data showed homophily, it was significantly lower as compared to the homophily shown by the demographic data.

![Table 5: z-test to test for significance in user’s interest data when n = 80](image)

<table>
<thead>
<tr>
<th>User’s Interest Attributes</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hobbies</td>
<td>0.056</td>
</tr>
<tr>
<td>Books</td>
<td>0.162</td>
</tr>
<tr>
<td>Music</td>
<td>0.062</td>
</tr>
<tr>
<td>TV Shows</td>
<td>0.0451</td>
</tr>
<tr>
<td>Movies</td>
<td>0.138</td>
</tr>
</tbody>
</table>

5.2.2 Homophily between users and their recommended friends

Recommendation data was then tested for homophily to check whether people recommended coupons to friends who were more similar to them than their average friend. Different demographic attributes such as gender, college name, college year, current country, current state, and current city were considered. As shown earlier, there was more homophily for a friend than a random user. The recommendation data shows that this already existing homophily increased when recommended friends were considered as shown in Figure 12.

Age was one of the demographic attributes that was considered, but unlike other attributes, rather than testing for
similarity in ages, difference in ages was calculated. The same homophily principle that was shown for the other demographic attributes was seen in this case as well as shown in Figure 13.

The p-values were calculated for the different demographic attributes when user’s friends and recommended friends were compared. The p-values were significant for only gender, college year and current city as shown in Table 6. Although the graphs seem to show a significant increase in homophily for college name, current state, current country, and age, due to the large variance of the data as well as the small size of the data set, there was no significance when regression testing was done.

Users first search for coupons that interest them, and then recommend those coupons to their friends. Categories like apparel, jewelry & watches, tools & automotives, health & beauty, etc., are gender biased, and hence when people search for coupons for themselves in these categories, they also recommend these coupons to the same gender. Gender is also an attribute that is listed most often on a user’s profile and so a large enough dataset was available to show significance for gender.

Similar to the demographic attributes, interest attributes for recommended friends were also considered as shown in Figure 14, there was increased homophily for hobbies and television shows, but slight decrease in homophily for music. When regression testing was conducted, it was found that the p-value was significant for only television shows as shown in Table 7. This is probably due to the fact that there are a fewer number of television shows as compared to the number of available musical selections as well as the fact that more people tend to list television shows on their profiles compared to music and hobbies.

<table>
<thead>
<tr>
<th>Demographic Attribute</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>0.27756</td>
</tr>
<tr>
<td>College Name</td>
<td>0.19798</td>
</tr>
<tr>
<td>College Year</td>
<td>0.03865</td>
</tr>
<tr>
<td>Current Country</td>
<td>0.04180</td>
</tr>
<tr>
<td>Current State</td>
<td>0.06439</td>
</tr>
<tr>
<td>Current City</td>
<td>0.35635</td>
</tr>
<tr>
<td>Age</td>
<td>0.31282</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interest Attribute</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music</td>
<td>0.12882</td>
</tr>
<tr>
<td>Tv Shows</td>
<td>0.00086</td>
</tr>
<tr>
<td>Hobbies</td>
<td>0.12882</td>
</tr>
</tbody>
</table>

Table 6: p-values shown for different demographic attributes when a user’s friends and recommended friends are compared

Table 7: p-values shown for different interest attributes when a user’s friends and recommended friends are compared

5.2.3 Product Characteristics

Now that homophily was found between a user and his recommendees, product characteristics were tested for correlation with (a) the number of times a user clicked on a given category, (b) the number of times a coupon in a given category was recommended by a user, and (c) the number of times a coupon recommendation from a given category was accepted. A regression testing was conducted to check for significance among the different product characteristics.
namely (1) risk, (2) complexity, (3) involvement, (4) brand equity, (5) experience, and (6) newness.

Three results were concluded: (a) Brand equity was the only product characteristic that showed a significant p-value when compared against the number of times a user clicked on a given category as shown in Table 8. Users only tend to search for coupons that they have heard of and according to the definition of brand equity, coupons in categories with high brand equity are marketed the most. Therefore, as expected brand equity shows a significant correlation when users look for coupons. (b) Table 8 shows that brand equity is again the only product characteristic that shows a significant p-value when compared against the number of times a user recommends coupons in a given category. Users only recommend coupons that they search for and from the earlier result, brand equity was the only product characteristic that was correlated with the search for coupons. Therefore, brand equity is again the only product characteristic that was correlated with the number of times coupons in a category were recommended. (c) When acceptance of coupon recommendations was considered, complexity was the only product characteristic that was significant as seen in Table 8. The coefficient of the complexity variable was negative showing that as complexity increased, users were less likely to accept the recommendation.

<table>
<thead>
<tr>
<th>Product Characteristics</th>
<th>Coupon Use</th>
<th>Coupon Recommendation</th>
<th>Coupon Recommendation Acceptance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk</td>
<td>0.8859</td>
<td>0.6451</td>
<td>0.0214</td>
</tr>
<tr>
<td>Complexity</td>
<td>0.9025</td>
<td>0.4512</td>
<td>0.0212</td>
</tr>
<tr>
<td>Involvement</td>
<td>0.9714</td>
<td>0.2811</td>
<td>0.1947</td>
</tr>
<tr>
<td>Brand equity</td>
<td>0.9713</td>
<td>0.9710</td>
<td>0.8659</td>
</tr>
<tr>
<td>Experience</td>
<td>0.9717</td>
<td>0.9700</td>
<td>0.5135</td>
</tr>
<tr>
<td>Newness</td>
<td>0.8592</td>
<td>0.2985</td>
<td>0.9008</td>
</tr>
</tbody>
</table>

Table 8: p-value shown after regression testing for product characteristics

5.2.4 Click rates

Finally, the hypothesis that click rate increases from a banner advertisement to a random recommendation to a genuine recommendation was tested. Normally, products are advertised through banner advertisements, which are displayed on websites. These advertisements are often random and not tailored to individual users. It is known that the click rate for these kinds of advertisements is approximately 0.5% [7]. The random recommendation click rate was found to be approximately 3%, which is higher than that of the banner advertisements. This shows that regardless of product preferences, the recommendations are accepted solely based on the fact that a friend recommended it. The genuine recommendation click rate was also calculated and it was found to be approximately 11% as shown in Table 9. This rate is much higher than that of random recommendations and banner advertisements. This shows that apart from a friend recommending a product, when the recommendation is genuine (the user considers the friend’s product preferences) the click rate increases drastically. Therefore looking at these results, genuine recommendations are a better form of advertisements, especially over social networks.

Table 9: p-value shown after regression testing for product characteristics

<table>
<thead>
<tr>
<th>Type of Advertisements</th>
<th>Click Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banner Advertisements</td>
<td>0.56%</td>
</tr>
<tr>
<td>Random Recommendations</td>
<td>2.55%</td>
</tr>
<tr>
<td>Genuine Recommendations</td>
<td>11.32%</td>
</tr>
</tbody>
</table>

6. FUTURE WORK

This Facebook application was built for the purpose of finding better marketing strategies over social networks. The capabilities of this application allow for data collection to help track and understand users’ product preferences and who they would recommend products to. The final results from this data collection can range from general statistics about products or users to predictive models of how the user is expected to act. The main obstacle to finding these results is getting a large number of users to actively use the application, and so as the number of users increase, so will the accuracy of the results and possible models.

Some of the interesting questions that can be solved in the future with a sufficiently large dataset are: (1) The effect of creating a chain message over a social network. This would entail sending out recommendations where the sole purpose is to propagate a certain product by creating a chain incentive hence increasing the "word of mouth" of a certain product. (2) Understanding how link type factors in with recommendations. (3) Calculate the link strength between friends on a social network to test what kind of product recommendations are sent over strong links and what are sent over weak links. 4) Creation of a model to predict who would be most likely to purchase a certain good with certain product characteristics. Social networks have a unique property in that one can track user’s data as well as friends’ data; and thus perhaps knowing users’ shopping habits would help predict what their friends would like. The world of network marketing has many possibilities and there are many questions that can be asked. This application was built to be versatile and so with it, many other questions as mentioned above can be solved.

7. CONCLUSION

Earlier research has been conducted to find better target marketing strategies [2], but limited research has been done using social networks to see whether they improve product acceptance rate. Hill’s study [8] used social networks to show homophily. Hence, using this homophily concept and social networks, there was some untapped potential that marketers could use to improve target marketing. This work aimed to show that social networks can in fact be used to improve target marketing. Also, research has been done previously based on specific products like movies [6], but no research has been done on general product categories such as apparel, electronics, home & garden, etc. There was also limited research based on relationship between product characteristics and social networks [8]. Therefore, this work looks at a wide range of product categories with different product characteristics and sees how coupons in each of these categories are recommended over a social network.

The Facebook application was created to collect data about
users’ demographic information, interest information, friends’ information and recommendations that they sent to their friends. This data helped show that: (1) Homophily increases between a user and (a) a random user to (b) a friend to (c) a recommended friend. Therefore a user is most similar to a recommended friend compared to all of his friends. From a marketers perspective, if a user likes a certain product, they can extrapolate that their friends will also probably like that product and thus advertise the product to these friends. (2) Users are more likely to search for and recommend products with higher brand equity whereas products with a lower complexity are more likely to be accepted. Therefore marketers should target products which have high brand equity when they want a user to use or recommend a product to a friend. However when they want a recommendation to be accepted, lower complexity products work well. (3) Genuine recommendations show a higher acceptance rate as compared to random recommendations, which show a higher acceptance rate than banner advertisements. The best way that a marketer can get a product to be accepted is by finding a user who is interested in the product and have that user recommend that product to his friends whom he thinks might like the product.

However, the limitations of this work are: (1) The number of users using the application is currently not large enough and given enough time for the application to spread on Facebook, a sufficiently large dataset can be collected. (2) All of the different available coupons are not recommended. Users first look for coupons that interest them, and then recommend only those coupons. Therefore data about all of the coupons is not available. (3) Some stores sell a wide range of products and it is not defined as to how these stores are classified into different distinct categories. (4) Brand strength cannot be completely calculated based on the number of coupons recommended from a given brand, as coupon’s monetary value influences it as well. After considering these limitations, there is some convincing evidence that shows that social networks can in fact improve target marketing.

8. REFERENCES