The Gift of Gab: Evidence TelE-Commerce Firms Can Profit from Viral Marketing

Shawndra Hill
New York University
44 W 4th St., New York, NY
10012-1126, USA
shill@stern.nyu.edu

Foster Provost
New York University
44 W 4th St., New York, NY
10012-1126, USA
fprovost@stern.nyu.edu

Chris Volinsky
ATT Labs Research
Florham Park, New Jersey
07934, USA
volinsky@research.att.com

ABSTRACT

Viral or buzz marketing takes advantage of communication linkages to propagate positive influence regarding a product or service. TelE-commerce is an ideal domain within which to study viral marketing, because communication linkages can be observed. In this paper, we follow a new telE-commerce service. In particular, we observe how the communication networks of existing customers influence the rate of product diffusion. The main contribution of this paper is evidence that consumers are more likely to purchase a service if they have previously spoken to a person who has the service. In addition, we offer the following three contributions: 1) the clarification that this need not be evidence of viral influence, we suggest different explanations; 2) we also describe the relation of these explanations to theories of purchasing behavior; and 3) we present some evidence to discern from among the explanations.

1. INTRODUCTION

How many times have you bought a product, signed up for a new service, or spread the word about a new product after you have been told about it by a friend, colleague, acquaintance or someone in passing? For many of us, the number is high. We have participated in *viral marketing*. It is this type of social network behavior that viral marketing campaigns seek to exploit. Firms believe that positive viral marketing will lead to increased profit and brand recognition.

Models that consider a consumer's network have traditionally been referred to as *word-of-mouth marketing*, *buzz marketing*, and *network marketing*. But, for e-commerce consumers, the ability to pass on a message about a product or service to others via email or a weblink is negligible. Enabling such channels for sharing product information gives influential consumers far greater reach in less time than can be achieved with traditional word of mouth channels. Viral marketing, as the process is typically called, generates rapid exponential growth in a product's exposure and influence as consumers themselves are passing on product information. Firms believe viral marketing is potentially more profitable than traditional marketing: the take rates are higher by those marketed to and viral marketing is cost effective. In addition, traditional marketing methods don't appeal to some segments of customers. For various reasons, customers value the appearance of being on the cutting edge or "in the know," and therefore derive satisfaction from promoting new, exciting products. In fact, the firm BzzAgents [4] managed to entice voluntary marketing of new products.

Firms assume word-of-mouth marketing exists, is prevelant, and is beneficial. However, they have not been able to measure the extent to which network attributes influence sales because they cannot observe their customers' communication linkages; therefore, they cannot target consumers based on the networks they belong to. We offer that telecommunications networks present a natural testbed for viral marketing models because the communication linkages and patterns of complete consumer networks can be observed and evaluated over time.

Our research utilizes telecommunications networks to build probabilistic models of product adoption. In this paper, we test our methods for predicting customer adoption on network data generated by the communication patterns from all customers who signed up for a new telE-commerce service, which is an Internet-based phone service. Unlike other data sets used in prior research to study viral marketing, this unique dataset enables us to monitor the adoption of the new service from its inception. In addition, we are able to observe consumer response rates to large direct marketing campaigns.

In the following sections we present evidence that viral marketing consumers, those potential customers who have previously spoken to a person with the service, respond to direct mailers at a higher rate than non-viral marketing consumers. We also can attain greater lift in profit when using network attributes derived from a consumer's social network compared to traditional customer segmentation data.

Figure 1 illustrates a simple viral marketing social network. The nodes, labeled A-O, represent customers, and the links between them represent influence as indicated by the directed arrow. A viral marketing node is a node that appeared on the network prior to purchasing the service. Node K spoke with node A, while node A was active, prior to signing up for service; likewise, node M spoke with node D. From our organization's standpoint, nodes D and A are influential customers because they spoke with a number of potential customers, A with 7 and D with 3, that later became customers.

Another example of a network attribute is *degree*, which for our target problem is the total number of active consumers a potential consumer is connected to. In Figure 1, we see again that A is connected to 7 people and therefore, her degree is 7. B is connected to 2 other people, and his degree is 2. Another attribute, the *clustering coefficient*, measures, on average, how similar potential customers are to the active customers around them based on who they communicate with. These informative attributes, and attributes like them, only arise from our ability to observe the network.

2. METHOD

In this study, we compare and contrast predictive models built with combinations of four sets of attributes: 1) customer attributes including demographics and preferences; 2) network transactions including the frequency , recency and duration of

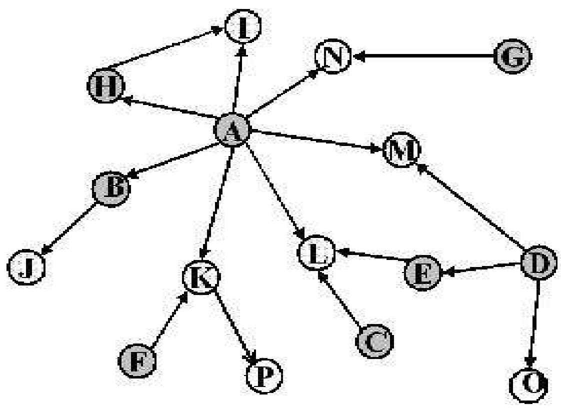

Figure 1: An example of a viral marketing influence subgraph. Nodes represent active telephone numbers and arrows indicate the direction of influence between the nodes. The nodes are labeled alphabetically according to tenure.

calls; 3) network structure attributes including network position and proximity to influencers and other similar individuals; and 4) experience attributes generated from the list of services potential consumer can interact with.

We use logistic regression to compare the influence of attributes used in our models. For evaluation, we rely on cost sensitive evaluation measures such as AUC and Brier Score. However, we are able to go beyond traditional target marketing methods because we can take advantage of reliable network attributes.

In the next section, we present evidence of viral marketing from a target marketing campaign directed to a substantial number of consumers including viral targets.

3. EXPERIMENT

In late 2004, we sent out a large direct mail marketing message to potential customers of this new service. The recipients of the mail piece were broken into 22 different marketing segments, based on typical marketing attributes such as services the customer had at the time of the marketing, characteristics of their calling behavior, demographics and other classical marketing segments. We created a list of potential viral customers, who had current users of the service in their calling neighborhood. Where our list had overlap with the marketing list, we looked to see how the sales rate (after one month) for our viral customers compared to the non-viral customers. Table 3 shows the sales rates for our marketing segments (for space reasons we restrict ourselves to the 5 segments with the largest number of viral customers marketed to). Due to proprietary restrictions in reporting the data, all sales percentages are normalized by the value of Segment 1 non-viral sales. The ratios of viral to non-viral sales show that the viral group always does better, in fact this is true across 18 of the 22 segments. Overall the viral group's sales outperform the non-viral group by about 2 to 1. One interesting fact is that the groups where the overall non-viral sales rate is the lowest (in this table, Segments 4 and 5) is where the benefit of the viral group is the strongest. This suggests that viral marketing tends to help most where traditional marketing does worst.

4. DISCUSSION

One of the main concerns for any firm is when, how and to whom they should market their products. Based on how much a firm knows about their target customer and potential customers, they may choose to mass market when they don't know much or to target market based on some desireable observed characteristics of current or potential customers or, more recently based on the network that they may have influence on [2]. We take a network marketing approach to this problem and provide evidence on real world data that there is indeed information in communication links.

Segment	N-V Sales%	V Sales%	Ratio(V/N-V)
1	1.00	1.47	1.47
2	0.75	1.01	1.34
3	1.21	1.67	1.38
4	0.22	1.42	6.53
5	0.45	1.10	2.45

Table 1: ***Comparison of sales rates for viral (V) and non-viral (N-V) customers from a direct mailing.***

Our preliminary results indicate, we can benefit from the use of social networks to predict purchases. However, viral marketing is not the only possible explanation for our result. So, we offer other explanations for this phenomenon and build different models of behavior informed by three theories of purchasing behavior: 1) homophily [1]; 2) collective behavior [3]; and 3) customer influence [2]. We evaluate each model of behavior against our data as a first step toward separating out the effect of viral marketing on profit.

Whether or not this is evidence of viral marketing is interesting from a research perspective, but does not necessarily bear on the importance to the firm – for example, if the reason is purely homophily based on some hidden variable, the firm can still use the network to improve marketing. For our research however, our immediate goal is to develop methods for separating viral effects from other market effects.

5. FUTURE WORK

Additional next steps include: utilizing telecommunications networks to build models of customer value in addition to predicting the likelihood of customer adoption, developing effective tools for dynamic network visualization, and building models to predict when viral marketing has negative effects on customer retention and attrition.

6. ACKNOWLEDGEMENTS

The authors thank Deborah Swayne for thoughtful comments on dynamic graph visualization and Deirdre Paul for help with data collection.

7. REFERENCES

[1] P. Blau. *Inequality and heterogeneity: A primitive theory of social structure*. Free Press, New York, 1977.

[2] P. Domingos and M. Richardson. Mining the network value of customers. In *Proceedings of the Seventh International Conference on Knowledge Discovery and Data Mining*, pages 57–66. ACM Press, 2001.

[3] M. Granovetter. Threshold models of collective behavior. *The American Journal of Sociology*, 83(6):1420–1443, 1978.

[4] R. Walker. The hidden (in plain site) persuaders. *New York Times Magazine*, pages 69–75,104, December 5 2004.