Testing contagion in small social groups via direct marketing

Johannes Bjelland, Geoffrey Canright, Kenth Engø-Monsen, Asif Iqbal, Pål Roe Sundsøy
Telenor Research and Future Studies
Telenor ASA
Oslo, Norway
Geoffrey.canright@telenor.com

Rich S. Ling
IT University / Telenor ASA
Copenhagen, Denmark / Oslo, Norway
rili@itu.dk

David Lazer
Northeastern and Harvard Universities
d.lazer@neu.edu

***Abstract*—In this paper we report on two experimental tests of contagious adoption in small connected groups. We define a "segment" as a group of customers with an expected affinity for the marketed product, and perform marketing experiments within the segment. Target groups include: "isolated" individuals (no social contacts in the segment); "pairs" (based on selecting a strong contact in the segment); and (in one experiment) "triplets"—cliques of three within the segment. Our experiments test the following questions: (i) how do adoption rates depend on the size of the targeted group? (ii) how do multiple offers to a pair affect the adoption rate?**

Social Network Analysis; Small clusters; Viral Marketing; Telecom

I. Introduction

It has long been known among marketers that our social network matters when we make purchasing decisions, and that having positive word of mouth about a product can be a key to success; see e.g. [1] for a review of studies on social networks within marketing. Traditionally, data on social networks have been difficult to collect, but in recent years researchers have gained access to massive social network data from e.g. online instant messaging services [9][5] and phone log data [2][4][3] [6][8]. Such data has made it possible to study e.g. social churn [3], service uptake [2] among telecom customers, and product adoption on an Instant Messaging network [9]. These studies confirm that consumer behavior is dependent on the communication network.

We have in a recent study [6] shown how the structure of the *adopter network*—the social network of adopters—develops over time, and how social spreading can be measured by studying this network. In this paper, we focus on a different aspect of social adoption. We wish to study the very local social effects that arise when the target unit for the marketing is a small, connected, 'molecular' group, rather than an 'atomic' individual. The key question is whether it is sensible to cluster targeting. Here we build on the ideas around complex contagion [7], which posits that the probability of adoption increases more than linearly with exposure. Traditional viral marketing approaches assume simple contagion—in which case clustering marketing information would be inefficient.

II. Method

Our social network is built by collecting call data records, aggregated over a 3-month period, and then using the communication links (voice and sms) as proxy for the social relationships. To remove error sources due to 'non-personal' relationships, we have applied some filtering of the dataset. E.g. we see that some customers have thousands of contacts during the three-month period. This can be machines, set up to automatically send SMSs, company call-centers, or other forms of extreme calling behavior. Such outlier nodes are filtered out based on combinations of extreme usage and degree (number of unique contacts). Only traffic between Telenor customers is used; calls to other operators are excluded.

We define the "segment" for the product to be offered based on a simple predictive model—which has been tested, and shown to have good predictive power. The segment is then, intuitively, those individuals who are interested in the product. More precisely, we build the segment from all individuals giving a minimum affinity score or higher. Isolated individuals are then those without a "strong" bond (over a minimum threshold) to anyone else *within the segment.* Similarly, pairs and triplets are all defined with respect to the subgraph defined by the segment.

Experiment 1 was performed in Asia, and is complete. In this experiment we compare: i) adoption rates within and outside the segment, ii) adoption rates of isolates and pairs within the segment and iii) the adoption rate of pairs in the segment where one or both members were seeded. Since there was no attempt to study triplets, "pairs" were defined by the presence of *at least one* strong enough bond. More precisely, pairs were defined by choosing non-isolates one at a time, and then choosing the strongest bond in the segment for this individual, until the desired target number of pairs was attained. For pairs, there were two randomly assigned treatment conditions: making the offer to only one of the pair ('molecular group 1', or MG1), versus making the offer to both ('molecular group 2', or MG2).

Experiment 2 (performed in Europe) compares isolates, pairs, and triplets. Since we wish here to compare pairs and triplets, it was necessary to ensure that the pairs chosen were in fact isolated (as pairs) in the segment—that is, that they were not actually part of larger clusters, and hence of a triplet. Defining triplets was simpler—we simply found cliques of three lying in the segment, without testing whether these cliques were or were not part of a larger cluster (since there was no comparison with larger clusters).

Pairs in experiment 2 were randomly divided into two treatment groups (one or two offers per pair), just as in Experiment 1. To distinguish these pair treatment groups from those in Experiment 1, we call them 'Pair1' and 'Pair2'.

There are in principle three possible treatments for triplets that are cliques (and even more for open "chains" of three). However, we have found that, starting with a customer base of millions, the segment definition, followed by application of our triplet criteria, gave in the end fewer than 3000 potential triplets. Thus, to ensure adequate statistics, we implemented only one triplet treatment group (composed of 1900 triplets), in which one randomly chosen individual from each triplet was given the product offer.

III. Results—Experiment 1

Figure 1 gives the main results from Experiment 1. This figure shows the adoption rates for different types of 'seeds' (those individuals receiving the direct offer). There are three results of particular interest in this figure:

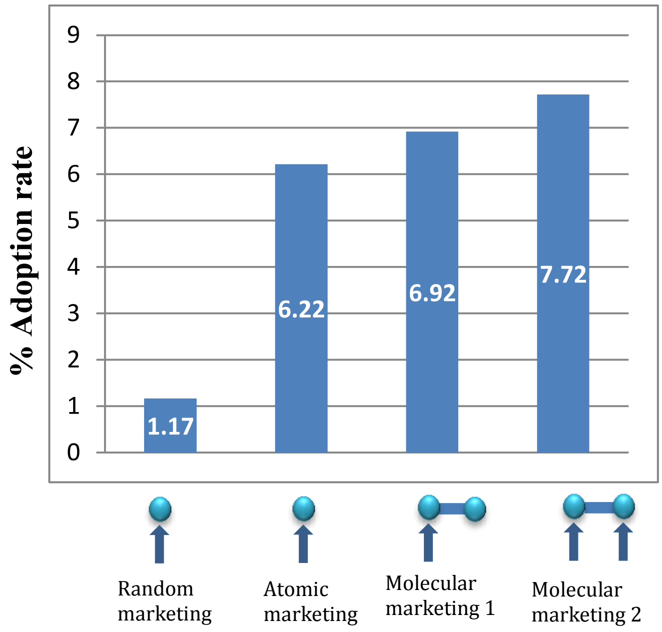

Figure 1. Here we show the adoption rate for seeds ('direct' hits) in percent. "Random marketing" (RM) shows the rate for randomly selected targets, while "Atomic marketing" (AM) shows the rate for isolated seeds in the segment. "Molecular group 1" (MG1) and "Molecular group 2" (MG2) show the adoption rate for seeds lying in pairs within a segment, where either one or both members in the pair were seeded.

First, we see that the adoption rate of randomly chosen targets is ca. 1% ("RM"), whereas it is above 6% for isolated seeds within the segment ("AM"). This difference gives an indication of the predicative value of the rules which define the segment.

Second, we compare adoption among isolated seeds ("AM") to that of seeds that have at least one friend in the segment ("MG1"). We see that when the seed lies in a pair its adoption rate is boosted by over 10% (compared to isolates). Our results do not give us the mechanism for this effect—it could be that those most susceptible to the product tend to have friends also in the segment, or that the friends in the pair discuss the offer together such that the adoption rate is increased—or both of these effects may be in play here.

Another thing worth noting is that the adoption rate of alters in MG1—that is, the nodes in pairs who did not receive the direct offer is 0.95%. Because there was no other way to receive the offer, this percentage should be viewed as a contagion from the focal node to the alter.

Third, we compare pairs where one or both nodes were seeded (MG1 and MG2). Our primary interest here is the adoption rates of focal nodes which received the marketing offer. In MG1 the adoption rate was 6.92% and 7.72% for MG2. This 0.8 percentage points difference is significant on a 5 percent level and should be interpreted causally as contagion—that is, subject to being exposed to the direct marketing, having an alter who is also exposed increases adoption by about 12%.

We then measure adoption in the pair – both seeds and non-seeds. This means including the ‘indirect’ hits in MG1 that are due to spreading. When comparing the net, per-pair, per-offer adoption rates the result is less clear however: 7.87% and 7.72% for MG1 and MG2, respectively—a difference that is neither substantively nor statistically significant. One might normally assume that the extra offers within the pairs in MG2 were “wasted”, since the viral spreading within the pairs should do the same job. Here we find that these extra offers were neither wasted nor advantageous—the two marketing approaches to pairs give essentially the same net result. We can say that the “resonance” effect of giving two offers within the molecular pair just barely makes up for the double cost in resources.

In the next iteration we will examine contagion beyond the pairs—the question is whether since more pairs adopted in MG2, the effects on third parties were amplified.

IV. RESULTS—EXPERIMENT 2

Experiment 2 has been carried out in a European country, and included a number of treatment groups not used in Experiment 1. Figure 2 shows the ‘direct’ hitrates (adoption rates for seeds receiving the offer) for all treatment groups.

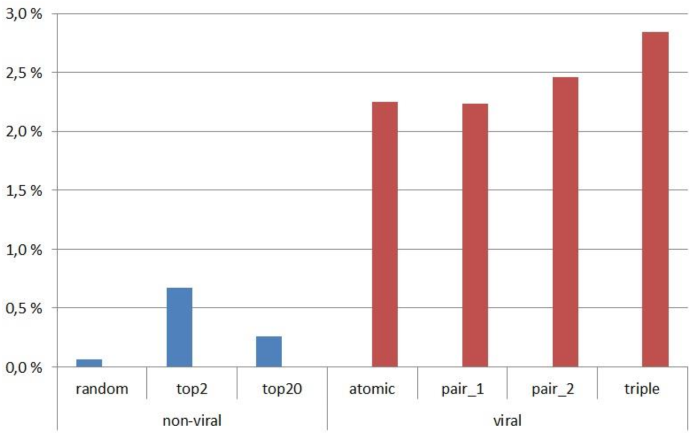

Figure 2. Here we show the adoption rate for seeds in Experiment 2. “random” show the adoption rate for randomly selected targets, whereas top2 and top20 show the rate for the top 2 and 20 percent of the customers ranked by affinity. “atomic” shows the adoption rate of isolated seeds in the segment. “pair1” and “pair2” are pairs within the segment where one or both in the pair were seeded. “triple” show the adoption rate of seeds in a triplet where only one member of the triplet was seeded.

First we note that overall adoption rates for Experiment 2 were significantly lower than those for Experiment 1—less than 3% for all groups. We assume that this is due to differences in the distinct combination of market, culture, and price.

We see again that the predictive model defining the segment works well here: both the top 2% of customers (as ranked by affinity) and the top 20% adopt at much higher rates than randomly selected customers. (These rates are, respectively, 0.68%, 0.26%, and 0.06%--extremely low!)

Secondly, we note that (in contrast to the result from Experiment 1) the pair1 group does not adopt more often than ‘atomic’ customers who are isolated in the segment. We do see however, in Figure 2, that the ‘resonance’ effect found in Experiment 1 is also present in Experiment 2. That is, the significant increase in seed adoption rates, from having a friend who has also gotten the offer, is seen again in these results: 2.23% for pair1, vs 2.52% for pair 2 (an increase of about 13%).

Unfortunately, we cannot compare the net adoption rates, per pair, per offer, in experiment 2; due to a technical failure, many of the adoption codes for viral adopters were not correctly recorded. This means that we only have a very weak basis for estimating the alter adoption rate for pair1, and so we cannot draw any conclusions here.

Next we observe that the direct hitrate for seeds in triplets is higher than that for any other group. We note that, on average, we expect that the seeds in triplets are more central (with respect to the segment) than those in pairs—which in turn are more central than isolates. Thus, we see support for the idea that more segment-central nodes are more likely adopters. This is a distinct idea from the more common one: that more central nodes in the segment are better *spreaders* of the product.

V. SUMMARY

We have reported method and results for a set of experiments which were designed to test the efficacy of targeted marketing to small, socially connected groups of individuals. Based on these results, we find a clear causal effect in marketing to pairs: sending the offer to both individuals in the pair significantly increases the adoption likelihood of each seed. However, our ‘bottom-line’ result is not decisive: the one-offer pairs and two-offer pairs have very nearly the same net adoption rate, in terms of total adopters per pair per offer. We also find, in Experiment 2, that the direct adoption rate is largest for seeds lying in the largest cluster (triples). This is consistent with the qualitative idea that more central nodes in the segment are not only better spreaders, but also more likely to adopt themselves. Finally, our interesting qualitative result from Experiment 2 shows an effect of viral campaigning on the seeds themselves: being given a ‘secret’ code may strongly enhance the probability of adoption. This result merits further study. We plan to test this result quantitatively, with a suitable experimental design, in the near future.

[1] C. Van den Bulte and S. Wuyts,"Social Networks and Marketing", Marketing Science Institue 2007

[2] S. Hill, F.Provost and C. Volinsky, “Network-Based Marketing: Identifying Likely Adopters via Consumer Networks”, Statistical Science. 2006, Vol. 21, No. 2, 256-276

[3] Dasgupta, K., Singh, R., Viswanathan, B., Chakraborty, D., Mukherjea, S., Nanavati, A. A., and Joshi, A. 2008. “Social ties and their relevance to churn in mobile telecom networks”. In Proceedings of the 11th international Conference on Extending Database Technology: Advances in Database Technology (Nantes, France, March 25 - 29, 2008). EDBT '08, vol. 261. ACM, New York, NY, 668-677.

[4] J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész and A.-L. Barabási,"Structure and tie strengths in mobile communication networks." Proc Natl Acad Sci U S A. 2007 May 1; 104(18): 7332–7336.

[5] Sinan Aral, Lev Muchnik, and Arun Sundararajan. "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks." Proceedings of the National Academy of Sciences, 106(51):21544–21549, December 2009.

[6] P Sundsøy, J Bjelland, G Canright, K Engø-Monsen, R Ling, “Product adoption networks and their growth in a large mobile phone network”, IEEE Advanced in Social Network Analysis and Mining (ASONAM 2010).

[7] Centola, D., The spread of behavior in an online social network experiment, Science 329, 1194-1197, 2010

[8] N Eagle, A. Pentland, D Lazer, “Inferring friendship network structure by using mobile phone data”, National Academy of Sciences 106.36 (2009) 15274-15278, 2009.

[9] Rushi Bhatt, Vineet Chaoji, and Rajesh Parekh. 2010. Predicting product adoption in large-scale social networks. In *Proceedings of the 19th ACM international conference on Information and knowledge management* (CIKM '10). ACM, New York, NY, USA, 1039-1048.

[10] P. Sundsøy, J. Bjelland, K. Engø-Monsen, G. Canright, R. Ling, "Comparing and visualizing the social spreading of products on a large-scale social network", In: “The Influence of Technology on Social Network Analysis and Mining”, Lecture Notes in Social Networks Volume 6, 2013, pp 201-225, Springer, 2013.