Keeping Up with the (Pre-Teen) Joneses: The Effect of Friendship on Freemium Conversion

Jared Sylvester* and William Rand

Center for Complexity & Business
Smith School of Business
University of Maryland
College Park, Maryland, USA
{jsylvest,wrand}@umd.edu

Abstract

We study a massively multiplayer online (MMO) freemium educational game with an in-game social network to investigate the possible effect of social contagion on user conversion. We find that addition of new content is not linked to user conversion. In contrast, there is evidence linking certain characteristics of a user's social network with their conversion. Chief among these is the local clustering coefficient among only those of their friends who have subscribed.

1 Background

Massively multiplayer online games and apps are becoming more commonplace and are being used by an ever increasing market. Many of these games and apps have switched in recent years from a paid subscription model to a "freemium" model, where the majority of users pay nothing to participate and a small number of users pay a monthly fee in exchange for some premium benefits. A principal challenge for managers in this space is how to convince more users to pay for the premium membership since it is often a pure profit stream for them. This is especially important given recent reports that the average cost per acquisition for freemium apps has outstripped the average customer lifetime value.

The central question then is what drives conversion to premium membership. There seem to be two major competing hypotheses: (1) users convert because they like the content they see in the free version and want to try out the content available in the premium version, or (2) users convert because their friends convert. Here we explore these two hypotheses in the context of a massively multiplayer online educational game that is aimed at children ages $6-12$.

*Corresponding author: http://www.cs.umd.edu/~jared/

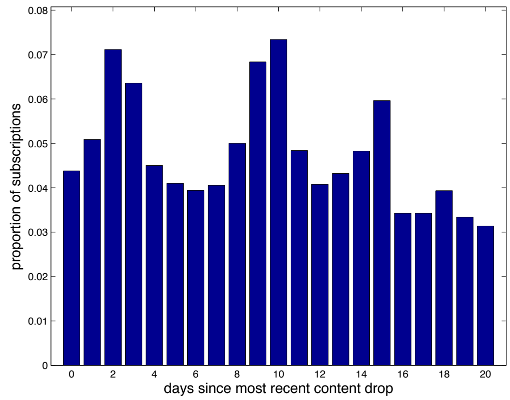

Figure 1: Fraction of conversions occurring since most recent content drop.

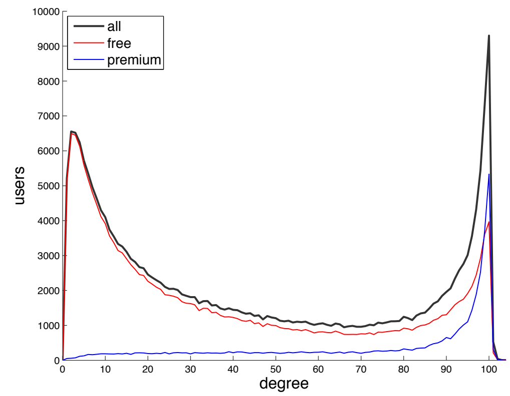

Figure 2: Degree distribution for free and premium users.

2 Initial Investigation

To study the role of social contagion in freemium environments we analyze data from 1.4 million game users, including 10 million friendships in the in-game social network and 2.1 million "e-cards" that users can send to each other.

Initial exploration of this data shows that there is no observable surge of conversions following the release of new content (called "drops," Fig. 1). If new content were the primary motivation for conversion one would see a decreasing trend as the time since last content drop increases. There is also a clear difference in friendship patterns between free and premium users, as observed in the degree density graphs shown in Fig. 2. Free users can be found on both the low-degree and high-degree ends of the graph, but premium users are only concentrated in the high-degree cluster. This lends credence to the idea that social activity is related to conversion.

Conversions tend to be preceded by the conversion of a friend (Fig. 3), with 18.6% of upgrades occurring within 24 hours of a friend's, and half happening within three days. Further supporting the importance of social relationships is the observation that having both more subscribers as friends (Fig. 4) and having a higher proportion of one's friends who are subscribers increases the likelihood that one will subscribe. While none of these observations are conclusive *eo ipso* they do point towards the social contagion hypothesis.

In addition to this exploratory analysis we adopt a predictive approach to determining which factors drive conversion. It would be useful to be able to identify users who are likely to convert 72 hours ahead of time so that appropriate promotions are enacted to ensure that conversion occurs. Such interventions may be targeted not only to effect the conversion of a particular user, but with the intention of spurring a cascade of conversions among their peers. Features considered include demographic data, weekday,

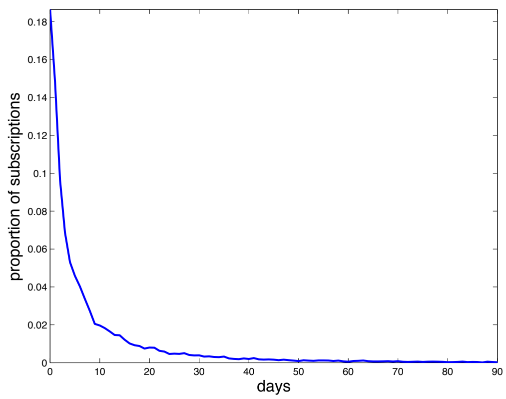

Figure 3: Days elapsed between most recent friend's conversion and a user's own.

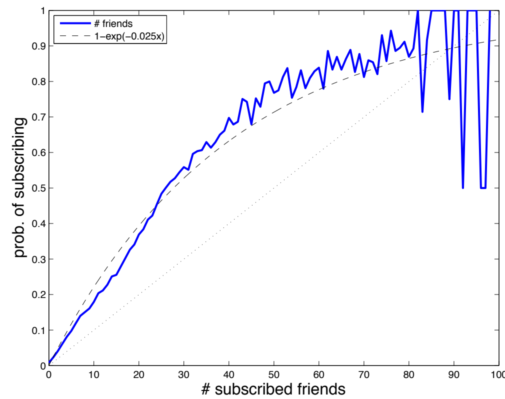

Figure 4: Probability of subscribing given the number of premium users one is friends with.

time since last content drop, a user's number of free, premium and total friends, usage of in-game e-cards, and others.

3 Results

A logistic regression reveals that recent content drops make it *less* likely that a user will convert, not more. Other factors that we had expected to drive conversion turned out to have no effect, such as a user's number of friends. The three factors with the largest positive effect on conversion rates were: (i) the local clustering coefficient of a user's premium friends (CC-premium); (ii) the proportion of their friends already subscribed (prop-premium), and; (iii) whether a friend had converted in the last three days (recent). The first of these was by far the strongest predictor. Our interpretation of this CC-premium effect is that being connected to a small group of tightly linked premium users is more influential than knowing many premium users who are not mutually acquainted. It is also interesting that the proportion of a user's friends who have subscribed is predictive, but the absolute number is not. Both of these may be related to a conspicuous consumption effect. The CFS Subset [4] and Information Gain Ratio [2] algorithms for feature selection resulted in similar findings.

Due to the imbalance in the classification problem (99% of users do not convert), and the static nature of many of the features, the prediction of conversion in three days is particularly difficult. This difficulty is greatly alleviated if one is willing to trade lowered accuracy for increased specificity. For example, analysis of the ROC curve produced by Bagging [1] yields an estimated Positive Predictive Value of 0.83. (That is, out of 100 users thought to be on the verge of converting approximately 83 of them will do so). This raises issues of weighted error and cost-sensitivity: One customer acquisition may be worth significantly more than the loss from a false-positive.

4 Future Work

Our biggest challenge is that too much of the dataset is static. To rectify this there is currently a data collection ongoing that will give us daily snapshots of the social network rather than a single image at the end of the study. In addition we will have access to dynamic data such as minutes played per day.

In the future we plan on using this dynamic social network data to build heterogenous Agent-Based Models of diffusion. Most diffusion ABMs have every node operate according to identical rules, or at most vary their behavior by drawing a parameter from a predetermined distribution (e.g. [3, 5]). We propose using values derived from empirical distributions. Furthermore, we can form more intelligent, data-driven rules for contagion, for instance by adapting conversion probabilities dynamically according observations such as those in Fig. 4. Finally, we can build effects like CC-premium in to the model directly to incorporate aspects of the local graph which appear to be influential in this real-world network.

References

[1] L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.
[2] R. Caruana and D. Freitag. Greedy attribute selection. In *ICML*, pages 28–36, 1994.
[3] J. Goldenberg, B. Libai, and E. Muller. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, 12(3):211–223, 2001.
[4] M. A. Hall. *Correlation-based feature selection for machine learning.* PhD thesis, The University of Waikato, 1999.
[5] T. C. Schelling. *Micromotives and macrobehavior.* WW Norton & Company, 2006.