

# User-Generated Content in Mobile Digital Media: Modeling Interdependencies, Social Effects and Geographical Mobility

Anindya Ghose  
aghose@stern.nyu.edu

Sang Pil Han  
shan2@stern.nyu.edu

## 1. Overview

Over the past few years, user-generated content has proliferated rapidly on the Internet, and there is anecdotal and empirical evidence that such content has had a profound impact on electronic commerce. Scholars and practitioners alike have been interested in investigating the relationship between user-generated content and economic outcomes like sales or pricing strategies (for example, Chevalier and Mayzlin 2006, Ghose et al. 2006, Dellarocas et al. 2007, Duan et al. 2008, Forman et al. 2008). Examples of user-generated content in the online world include product reviews, reputation feedback, blog postings, audio and video clips on social networking sites, and so on.

Taking cues from electronic commerce, different kinds of user-generated content (hereinafter UGC) are becoming available in mobile digital media environment as well, spurred by rapid advances in the cellular telephony market. Besides regular content in opinions forums and social networking sites that are created through mobile media, other examples of UGC in mobile digital media settings include creation and uploading of photos, graphics, ring tones, videos, podcasts, and other kinds of multi-media content. As mobile commerce becomes an increasingly popular mode of conducting transactions in several parts of the world, newer strategies for mobile-based advertising and marketing may need to be implemented in this environment in addition to or in lieu of traditional internet marketing methods. This requires a deeper knowledge of users' content creation and usage behavior in mobile media settings as well as the potential for monetizing such content. However, little is known about how content creation by users is related to their usage of such content, or vice versa, and how these processes are affected by users' personal, geographical and social network characteristics in mobile commerce.

While most consumers tend to use content created by other users or by third-party providers, some users contribute by creating and uploading content as well. For a given user, content generation and usage may not be independent decision-making processes. Rather they are likely to be inter-related processes. Users maximize their utilities by allocating time between activities due to resource constraints (Becker 1965). They may switch their content creation and usage modes in order to maximize their utility subject to resource constraints (e.g., time and money). For example, in a given day, if a user has spent most of her leisure time in consuming content by browsing social media sites (e.g., Facebook, Epinions and other sites), that user would have only a little time left to generate content (e.g., photos, reviews, videos, and so on). It is also possible that high levels of content usage in previous periods may motivate users to continue to contribute content over time as they begin to feel an integral part of the community forged in these social media sites. Users' behavior may also be affected by the content generation and usage behavior of their network neighbors. Since users can access mobile media ubiquitously, the extent of geographical mobility may also influence their content generation and usage processes. These issues call attention for the need to better understand the emerging cellular digital media market, with a focus on users' economic and social behavior.

In this paper, we investigate the dynamics and the interdependencies between users' content generation and usage in the mobile digital media environment. We use time-series and panel data methods based econometric models to study users' content generation and content usage behavior in a mobile media setting. In addition, we map how the behavior of a given user's underlying social network affects the content generation and usage behavior of the same user. Finally, we also study how the extent of dispersion in geographical locations from where mobile calls are placed affects users' propensity to generate and use content.

## 2. Data

In this section, we describe the data that we collected from a large telecommunications service company in South Korea. Our sample is drawn from 3G mobile users who used the services of the

company between March 15, 2008 and June 15, 2008. We collected data from 3G mobile services because it enables users to upload their content faster than the conventional mobile services. Further, these services are more commonly available in the larger screen handsets that facilitate more user-friendly content generation and usage. It took us almost 2 years to negotiate access to this dataset which we eventually intend to make public (properly anonymized) for validation, verification and continuous research. The dataset consists of more than 9.5 million mobile data transaction records across 180,000 users encompassing their content uploading and downloading behavior. We also have data on 34.9 million voice calls made by the same users that enables to draw social networks. Our mobile transaction data includes individual-level information on users' content generation and usage activities include detailed information on all kinds of content generated by users through their mobile phones. In addition, we have detailed user-related demographic and geographical location data including the location from where the call was placed that provide further insights into the drivers for users' mobile content generation and usage behavior.

### **3. Econometric Analysis and Statistical Challenges**

We use both classical time-series analysis and simultaneous equations panel data modeling approaches to examine these questions. Using time-series analysis (e.g., vector autoregressive (VAR) model), we capture the evolution and interdependency between content generation and usage at an aggregate-level. Findings from the VAR model (e.g., the number of significant own lags and the lags of the other variables) provide key insights that help us specify an individual-level model. The simultaneous equations model enables us to not only quantify the dynamic interdependency between a user's content creation and usage, but also helps identify the endogenous social effect affecting one's content creation. We find that there exist positive state-dependencies within a given users' content generation and within content usage behavior, which carry forward from a previous period. We also find that there exist negative and persistent inter-temporal interdependencies between the content generation and usage behavior for a given user, which is in fact consistent with the theory of resource constraints in economics (Becker 1965). In addition, we find that the content creation behavior of a user's network neighbors is negatively associated with a given user's content creation behavior while content usage behavior of network neighbors is positively associated with a given user's content usage behavior. We use various social network measures such as degrees, diversity, centrality, and betweenness. This result supports the contagion theory (Contractor and Eisenberg, 1990). Finally, we find that while there is a positive influence of a user's geographical mobility on his/her content generation behavior, its impact on a user's content usage behavior is not always statistically significant.

A key statistical challenge we face in our estimation is that our model involves multiple equations which are simultaneously related, include lagged dependent variables as independent variables, and use a panel data set in the estimation. Additional complication arises from a sample selection issue since econometricians can observe a user's content generation and usage behaviors only when the user pushes a button in a keypad on his or her handset or by clicking an icon on a touchpad to activate a mobile session. Thus we incorporate the selection equation into the multiple-equation panel data model. Further, we aim to separate the social effect of users' being influenced by their social network neighbors from the covariance caused by both inherent similarity (i.e., homophily) and exogenous user-level factors that are not observed by the econometricians. We hope that techniques that account for these aspects in a statistically robust way would be very useful for research in this area.

### **4. Conclusion**

Implications from this paper can provide mobile commerce providers with insights into how to monetize such UGC. It can help advertisers gain an understanding of the interdependencies between users' content creation and usage behavior as well as the impact of user's social influence and geographically mobility that can be used for mobile phone-based advertising. Furthermore, this research will contribute to the emerging literature on economic value and monetization of UGC by adding to the literature that has examined these issues in online social media. We will further describe some of the econometric and statistical challenges we have faced during the course of our analysis of the data, and some research contributions we have made towards addressing these issues.

## Select References

1. Becker, G. S. 1965. A Theory of the allocation of time. *Economic Journal*, 75(299), 493-517.
2. Chevalier, J., D. Mayzlin. 2006. The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*. 43(3) 345-354.
3. Contractor, N. S., E. M. Eisenberg. 1990. Communication networks and new media in organizations. In J. Fulk and C. W. Steinfield (Eds.), *Organizations and Communication Technology*. Newbury Park, CA: Sage, 143-172.
4. Dellarocas, C., N. Awad, M. Zhang. 2007. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, 21(4), 23-45.
5. Duan, W., B. Gu, A. B. Whinston. 2008. Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems*, 45(4), 1007-1016.
6. Forman, C., A. Ghose, B. Wiesenfeld. 2008. Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets. *Information Systems Research*, 19(3), 291-313.
7. Ghose, A., P. Ipeirotis, A. Sundararajan. 2006. The Dimensions of Reputation in Electronic Markets, Working Paper, SSRN.