DOI: 10.1145/1562164.1562199

BY MU XIA, YUN HUANG, WENJING DUAN, AND ANDREW B. WHINSTON

Ballot Box Communication in Online Communities

TIME'S PERSON OF THE YEAR: You – In 2006, the World Wide Web became a tool for bringing together the small contributions of millions of people and making them matter.

—TIME.COM, DEC. 2006

The participation of individual users in online communities is one of the most noted features in the recent explosive growth of popular online communities ranging from picture and video sharing (Flickr.com and YouTube.com) and collective music recommendation (Last.fm) to news voting (Digg.com) and social bookmarking (del.icio.us). Unlike traditional online communities, these sites feature little message exchange among users. Nevertheless, users' involvement and their contribution through non-message-based interactions have become a major force behind successful online communities. Recognition of this new type of user participation is crucial to understanding the dynamics of online social communities and community monetization.

The new communication features in online communities can be best summarized as Ballot Box Communication (BBC), which is an aggregation mechanism that reflects the common experience and opinions among individuals. By offering a limited number of choices such as voting, rating and tagging, BBC creates a new medium to effectively reveal the interests of mass population (see Table 1). Compared with traditional Computer Mediated Communication (CMC) such as email, Web publishing, and online forums,[4] BBC influences user preferences by simplifying the mass sharing of individual preferences.

These technologies offer new ways for information consumers to be involved in community activities.[3] In traditional online communities, users only have two levels of participation: "watching from the sidelines" or "playing in the game," for example, they are either passive readers or active participants in conversations. However, BBC presents a new choice – "shouting from the stands" – in which each user can express his opinion through BBC and their collective preferences can be heard as a dominant voice. For instance, Digg readers can vote on news and promote it to the front page for millions of visitors to see.

In spite of the increasing significance of non-message-based online communication, very little is known about BBC-enabled communities. As entrepreneurs build and manage new online communities, they have no choice but to look for the "right" technologies by trial-and-error. Not surprisingly, the result is hit-or-miss: some of the grandest failures of the dot com bust featured online communities.[1] Only after costly failures, it has been recognized that not all technologies can benefit the growth and sustainability of a community.

Extant theories on online communities and communication networks may offer some guidance on understanding of the emergence of new online communities (such as YouTube). Whitaker et al.[7] identify online communities as "intense interactions, strong emotional ties and shared activities." In

addition, Monge and Contractor[5] define communication networks as "the patterns of contact that are created by the flow of messages among communicators through time and space." Both study the social interaction aspect of communities such as user commenting and discussing. However, the non-social interaction aspect, which is the focus of BBC and often dominant in contemporary online communities, has not received much attention.

Characteristics of BBC

Compared with traditional online communications, BBC utilizes restricted communication to aggregate user feedback such as opinions, strategies and choices and allows users to implicitly express their preference. Simplification, the many-to-one nature, and implicit influences on users are three major characteristics of BBC compared with CMC.

Simplifying Web-based Communication. In BBC-enabled online communities, users communicate through preconfigured technologies that provide limited interaction options and lower participation costs. As a result, the communication is more detached and simplified as users no longer have to commit to composing messages. For instance, when visiting a site, a user can interact with others by voting on their posts – it is easier to click to vote on prearranged choices than to write a comment. This lightweight interaction is likely to encourage more user activities.

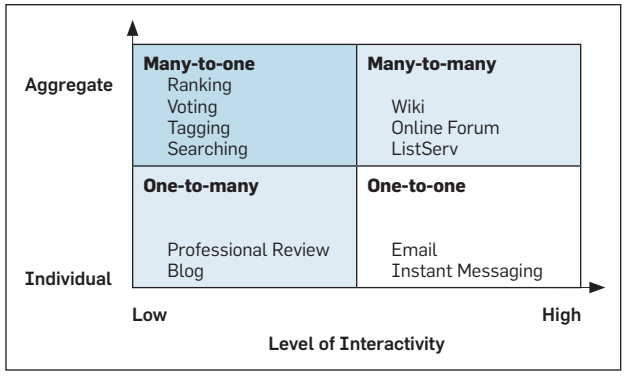
BBC also makes it easier for other users to get to know "the voice of the crowd" without incurring the high cost of, say, going through all comments. If we regard an online community as a medium that facilitates production and consumption of information, both sides now enjoy a better understanding of each other through BBC. Furthermore, by reducing communication costs, BBC also facilitates collective production.

Many-to-one Communication. Another distinct feature of BBC is its many-to-one nature, where multiple users' inputs are aggregated to form a single voice. As shown in Figure 1, many-to-one communication features a lower level of interpersonal interaction compared to many-to-many and one-to-one communications, such as online forums and email, in which information or message receivers have to understand others' messages to continue the conversation. Consequently, many-to-one communication has a unique advantage to convey many people's perceptions, preferences, and opinions on one subject, despite its inability to exchange complex semantic meanings.

Table 1. BBC-Related Techniques (In Ascending Order of Complexity for the User)

BBC Applications	Description	Examples
Access statistics	Indicating the popularity using view ranking, the number of visitors, and the number of comments	YouTube, Last.fm
Rating/Voting	Revealing users' opinions or the value of information through a poll or relevant activities such as marking as favorites and referencing	Social news (Slashdot, Digg, reddit) Blog aggregator (Technorati)
Tagging/Folksonomy	Generating metadata of content from individual labels (keywords) and publishing the outcome as various rankings, tag clouds, or search results.	Individual tagging (Youtube, Flickr), Social bookmarking (del.icio.us, Backflip), Collaborative tagging (Google Image Labeler, WikiMapia)
Searching	Recommending the most relevant results for search based on other users' search and feedback	Social network search engines (Jookster, NewsTrove)

Figure 1. Four Types of Unstructured Communication



As an aggregation mechanism, BBC focuses on revealing the common interests among users, as opposed to publishing individual thoughts in one-to-many communications (such as a blog). Essentially, the aggregation process is that many users express their opinions on a common subject matter using a very abstract language, sometimes as simple as making a binary choice. One interesting example is the famous social tagging game – "ESP Game" (www.espgame.org), in which two participants type words describing an image until they reach the same word, and then the computer system understands the image and stores the keyword as an accepted description.

Implicit Influence on Users. Even though non-message-based, BBC's impact can change usage patterns indirectly. User preferences are often swayed by the aggregate trend in the form of the most viewed or top-rated content in the community, or "trusted" individuals' plicit endorsement as expert votes. Moreover, his own consumption/voice will heighten this effect. In an online world, where users often look for guidance from others in developing their own taste, the implicit influence of BBC complements that of message-based recommendations.

Distinguishing BBC from CMC. These three distinct features of BBC offer new benefits over traditional CMC. CMC has an implicit assumption that technologies are used to facilitate the exchange of messages between users. BBC in contrast actually reduces the information richness in communications by replacing messages with a set of limited choices. Such deliberate reduction of information exchanged between users by BBC can alleviate information overloading, a more pressing issue brought about by the Internet. The communication choices BBC offers (such as, voting and tagging) are also less attached than traditional online communication (such as blogging and commenting), allowing users to participate more in communities.

The most typical BBC-enabled communities are online communities built for publishing and downloading digital

Table 2. Comparison of BBC and Traditional Message-Based Online Communications

	Online Message-based Communication	*Online BBC Communication*
User types	Contributors and lurkers	Producers and consumers
Communication richness	High	Low, mostly through observable activities
Communication cost	High cost associated with finding, reading, commenting, and posting	Low cost based on passive feedback
Role of technology	Managing messages and processing semantic content	Reducing the barrier of participation by offering a restricted set of communication choices
Community goal	To provide more content in higher quality	To aggregate user preferences
User involvement	Eyeball economy	Vote-by-foot economy
Influence on Users	Through messages and explicit	Through actions and implicit
Analogy	"The noise of the crowd"	"The voice of the crowd"

content, such as, YouTube and Flickr. On these sites, even though each picture or video clip has a section for user comments, few people choose to post anything. Contributors/sharers get to know about users' opinions on their content through aggregate measures such as total views and average rating, which are automatically generated by the system or programs.

Table 2 lists a few pronounced differences between traditional message-based online communities and BBC-enabled online communities. With simplified and many-to-one communication, BBC can efficiently reflect mass users' feedback and integrate the production and consumption processes. When content producers adjust their offerings to better user demand, the social welfare of both sides is improved.

BBC in Peer-to-Peer Music Sharing Communities

As the most popular non-message-based online communities, peer-to-peer music sharing provides a good example to illustrate BBC features. In these communities, anonymous users share and search songs using software tools and, hence, there is no direct message exchange among the participants.

We study whether such a community exhibits the three characteristics of BBC by examining Internet Relay Chat (IRC) music sharing. Although mainly used for chatting, IRC has sharing channels that allow users to share and download files in a peer-to-peer fashion. Downloaders can locate the files contributed by sharers in two ways before deciding to download. They can either send a search command with specific keywords, which all sharers will respond to, or browse the list of files available from a particular sharer.

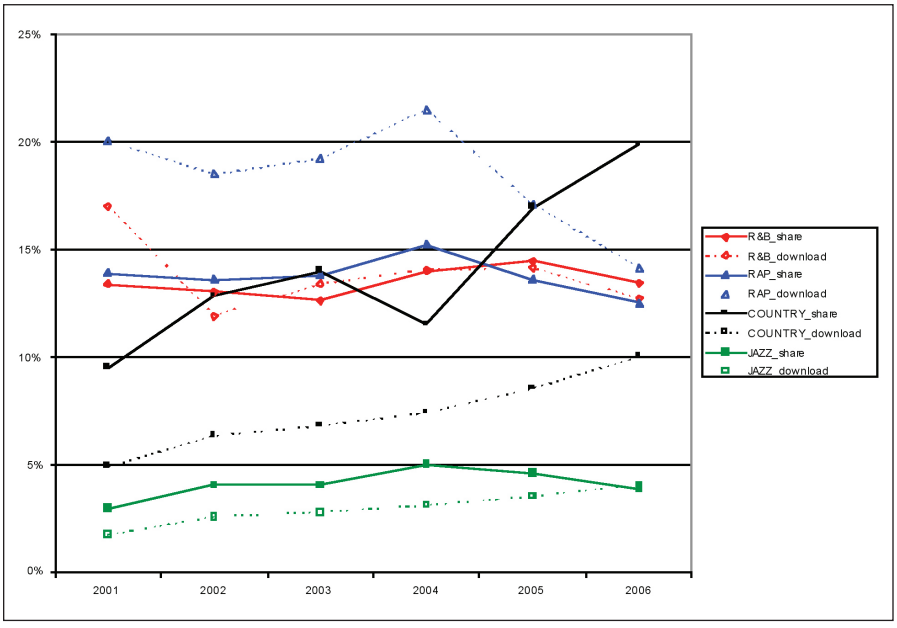
It is straightforward to confirm that the IRC channel exhibits the first two BBC characteristics. First, sharers, by making their favorite music available for download, effectively cast their vote on what music is preferred, a simplification over recommending the music in a review. It is also a many-to-one communication since multiple users' aggregate votes determine the popularity of music and it can be felt by an individual user when he searches for the music. However, whether implicit influences on users exist cannot be directly observed. To answer this question, we examine changes in aggregate and individual music preferences by analyzing a recent six-year (from 2001 to 2006) dataset, that recorded millions of IRC users' searching, browsing, and downloading activities as well as sharers' collections of files in an IRC channel #mp3passion.[a]

Aggregate Preference Changes

To test BBC's influence, we tally the numbers of songs available in the channel (supply) and actual download (demand) by genre and investigate how the genres of music changed over time. We choose five major music genres – Rock, R&B, Rap, Country, and Jazz[b] – as representatives of users' preferences in the channel. For these five genres, we select all music by 298 first-tier artists according to AllMusic's classification[c] and calculate the ratio of songs in each genre over all songs identified. We aggregate all demand and supply on a yearly basis to reduce the random impact of individual preferences. Figure 2 is a depiction of the proportion of different genres during the six years.[d]

As shown in Figure 2, supply and

Figure 2. Yearly Ratios of Sharing and Downloading Volumes (By Genre)



a While most of the music exchanged is pirated, the focus of this paper is not on the legality issues.

b Rock, R&B, Rap, and Country are the top 4 most popular music genres and Jazz represents less popular genres.

c AllMusic.com is the leading authority on music statistics.

d Even though Rock is the most popular genre and covers more than 55% share and download, because the volume does not change much, we use the other four genres to represent four degrees of freedom.

demand preferences, measured by the percentages of four genres, converged over the years. For example, RAP's share held steady at around 14% of all music sharers provided, but download percentages of this genre decreased. Since there is no message exchange among users, without BBC's influence, there are two possible scenarios of supply and demand dynamics. If individual demands were highly intrinsic and hardly influenced by sharers, the resulting proportion of download for that genre should be independent of supply. If outside forces such as music retail markets and broadcast media were influential, the demand and supply should be highly correlated between themselves and with the factors. However, neither of the two, or even any hybrid form, would be able to explain the preference convergence in supply and demand.

The most plausible explanation for the convergence of preferences, shown in Figure 2, is that downloaders were implicitly affected by the voting results, since the available music is the aggregate preference of all sharers. As a result, sharing activities have an impact on users' download. For example, as the number of country songs kept increasing and became a large proportion of songs provided, searches for country music were more likely to be successful. Indeed, this community exhibits the implicit influence of BBC.

Individual Preferences Changes

To illustrate how an individual user's demand can be affected by BBC, we show an example of a typical sharer from our data, "John Doe." Table 3 summarizes John Doe's activities during a five-week period in March and April of 2006.

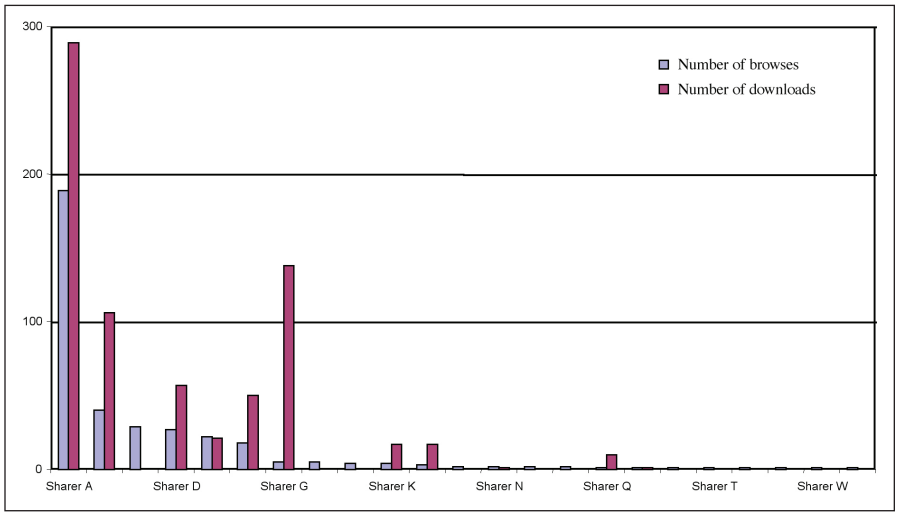
Three pieces of evidence shown in Table 3 illustrate the implicit influence of BBC. First, John Doe's browse commands led to most of the download, which is not necessarily what he had originally searched for. For John Doe, searches were mainly to identify users who had the content he might be interested in. Once such users were identified, John Doe would retrieve the complete list of their available files through browse commands. Therefore, both the number of browses and the number of browse-initiated downloads are much larger than that of searches and search-initiated downloads.

Table 3. John Doe's Activities in Five Weeks

Week	#searches (#downloads)	#browses (#downloads)	From Sharer A #browses (#downloads)	#files kept
1	28(26)	119(246)	55 (100)	165
2	18(16)	91(224)	47 (94)	162
3	5(0)	45(62)	33 (14)	28
4	10(9)	61(163)	32 (79)	84
5	6(1)	47(10)	22 (0)	11

searches (#downloads): number of John Doe's search commands (number of search-initiated downloads)
browses (#downloads): number of John Doe's browse commands (number of browse-initiated downloads)
files kept: the number of download files kept in John Doe's own collection

Figure 3. John Doe's Browse and Download Distribution



Second, we find that John Doe frequently went back to the same set of his favorite sharers and browsed their content before downloading from them. Therefore, the download is heavily influenced by these sharers' collection, that is, the content they provided in a form of implicit voting. Moreover, as shown in Figure 3, a small set of sharers account for a disproportionate amount of downloads. For instance, John Doe checked Sharer A's collection almost every day, and eventually more than 30% of his downloads came from Sharer A.

Third, John Doe not only consumed the music, but also kept and shared many of the files. This can be regarded as "voting" as John Doe implicitly endorsed the content by keeping it in his own collection, effectively increasing the likelihood another downloader would discover the content.

To summarize, John Doe discovered sharer preference through browse commands and was highly influenced by a small set of sharers. Furthermore, by replicating some files downloaded and making them available for others, he implicitly cast his vote on those files. This evidently demonstrated the power of implicit influence of BBC.

Building BBC-Enabled Communities

To build a successful community with advanced features of BBC, we believe the same lessons we learned about e-commerce (mainly through the failures) can be applied to the BBC case. Specifically, one still encounters the same challenges faced by businesses: production (content), marketing (getting people to know), and sales (having people continue to contribute to or buy products from your site).

The current Web 2.0 movement, for all its publicity and explosive growth, is a hodgepodge of implementations of often unrelated technologies, such as AJAX and RSS.[e] Yet the sustainability of most Web 2.0 sites remains a challenge for site operators due to the following three reasons related to BBC. First, the

e Web 2.0 is a marketing term coined by O'Reilly but its true meaning is often a topic of debate.

interaction between users is highly non-message-based, which may not help create the "stickiness" of the community. It also means that user population may be highly dynamic thus their collective behavior is hard to predict. Second, because individual interactions are a simplification of the real, complex user opinions and preferences, it is difficult to make any prediction by reading into their actions. Last, online communities are affected by aggregate user activities and behavior, which may entail a great degree of randomness due to their low cost for participation.

As communities increasingly build around content, it is crucial to encourage production and provision of content. However, the technologies in BBC have no built-in incentive mechanisms. Moreover, technologies may alter users' ability and their incentives to communicate. While it is easier than before to adjust interaction configuration thus change users' options, it is also not clear how these changes affect users' choices. In addition, the complex and highly dynamic interaction between different types of users and administrators of the community also makes it increasingly challenging to predict how a change is going to affect the communications.

Challenges in Understanding BBC. Although findings from extant literature can be valuable in understanding BBC, they are not readily applicable thanks to the unique features of BBC. As the users' influence on others is always imposed in a non-message-based thus implicit and collective fashion, the level of impact on individual users by collective actions of a large number of users is still not clear.

The interaction of users and the community also makes it challenging to predict the dynamics. Aggregate user communication behaviors, as detached, multifaceted, and idiosyncratic as they can be, determine the overall characteristics of the community such as total resources (total available content) and cost of using the resources (network congestions). Any individual user's behavior, in turn, is affected by these community-level characteristics.

BBC's influence may also be heavily dependent on the characteristics of evolving technologies. By making it easier to express one's preference, they also change users' participation behavior. The ensuing communications between users, therefore, are determined by the interaction of three parties: users, the community, and technologies. As a result, the outcome is difficult to characterize and its impact is even more difficult to gauge.

There are also many business-related issues in online sharing communities characterized by BBC. As many such communities have been started by entrepreneurs, there is a pressing need to identify a working business model so that the communities can be self-sustainable. While the current Web 2.0 trend values user-generated content, its sustainability as well as profitability is still a mystery. Currently, advertising seems to be the only business model available for such online communities. Viral marketing techniques take advantage of the community to promote products but the results are mixed. Nevertheless, exploring business value from online communities may distort BBC since a lot of the power resides in the community operator's hands. This study of BBC as a new communication mechanism will at least offer guidelines to answering the business-related questions. C

References

1. Bobala, B. Last breaths of theglobe.com? *The Motley Fool*, (Aug. 6, 2001); http://www.fool.com/news/2001/tglo010806.htm.
2. Duan, W., Gu, B., and Whinston, A.B. Informational cascades and software adoption on the Internet: An empirical investigation. *MIS Quarterly 33*, 1, (1999) 23-48.
3. Duan, W., Gu, B., and Whinston, A.B. Do online reviews matter? – An empirical investigation of panel data. *Decision Support Systems 45*, 4, (2008) 1007-1016.
4. Kiesler, S., and Sproull, L. Group decision making and communication technology, Organizational Behavior and Human Decision Processes 52, (1992) 96-123.
5. Markoff, J. Entrepreneurs see a Web guided by common sense. *New York Times*, (Nov. 12, 2006).
6. Monge, P., and Contractor, N. *Theories of Communication Networks*. Oxford University Press, NY, 2003.
7. O'Reilly, T. (2006). Web 3.0? Maybe when we get there. Blog post, Nov. 13, 2006 http://radar.oreilly.com/archives/2006/11/web_30_maybe_wh.html Retrieved on November 13, 2006.
8. Whitaker, S., Issacs, E., and O'Day, V. Widening the net. Workshop report on the theory and practice of physical and network communities. *SIGHCI Bulletin, 29*, 3, (1997), 27-30.

Mu Xia (mxia@uiuc.edu) is an assistant professor of Information Systems in the Department of Business Administration in the College of Business at the University of Illinois at Urbana-Champaign, Champaign, IL.

Yun Huang (yun@northwestern.edu) is a research scientist and post-doctoral research fellow in the Science of Networks in Communities (SONIC) research group in the department of Industrial Engineering and Management Sciences at Northwestern University in Evanston, IL.

Wenjing Duan (wduan@gwu.edu) is an assistant professor in the Department of Information Systems and Technology Management in the School of Business at George Washington University.

Andrew B. Whinston (abw@uts.cc.utexas.edu) is a Professor of Information Systems, Economics, and Computer Science, the Hugh Roy Cullen Centennial Chair in Business Administration, and Director of the Center for Research in Electronic Commerce, at the University of Texas at Austin, TX.