The Promotional Value of Peer-to-Peer Networks[i]

Ramayya Krishnan,[*] Alan Montgomery,[†] Michael D. Smith[*]

[*] H. John Heinz III School of Public Policy and Management, Carnegie Mellon University

[†] Tepper School of Business, Carnegie Mellon University

{ rk2x, alm3, mds}@andrew.cmu.edu

April 2004

Abstract

A major criticism of peer-to-peer (P2P) networks, such as Napster and Gnutella, is that their main purpose is the pirating of intellectual property. However, a redeeming aspect of P2P networks is that they can perform a valuable promotional function by making users aware of a product that they would not otherwise have considered. Specifically, peers can recommend a song to another peer. Music tends to have a great deal of heterogeneity in tastes, and if peers are connected to peers with similar preferences then the recommendations serve a valuable promotion function. In this research we seek to understand how P2P networks could aid information producers in increasing the user base of their products and their profitability. Traditionally producers have relied upon broadcasting playlists that are targeted towards a homogeneous set of users, which can limit their value. A potential benefit of P2P systems is that they can serve a collaborative filtering function by correlating preferences across users and providing playlists customized to an individual user's tastes.

1. INTRODUCTION

The frequent discussion of P2P networks like Gnutella, Napster, Kazaa, and Morpheus in the trade press has been about the pirating of intellectual property. However there is a promotional benefit from P2P systems that is often ignored. Traditionally, music has been promoted through centralized partners, like radio stations, that must to cater to large sets of consumers, resulting in more homogenous promotions. Further, because traditional broadcast radio stations cannot customize their playlists to individual customers, their recommendations much appeal to relatively large, heterogeneous audiences.

In contrast, P2P networks potentially could isolate much smaller, more homogeneous groups. P2P allows users to take advantage of the heterogeneity in preferences and implement recommender systems to share new music discoveries in a large market of albums. In this paper we discuss how such a system could be built and show what the promotional benefits of such a system would be.

A natural question is why P2P and why not distribute music through a centralized channel. There are two reasons. First, privacy and control as peers retain control over how their preference information is shared with other peers. Second, bandwidth and cost as: 1) privacy and control, and 2) bandwidth and the cost of promotion.

A basic difference with the P2P system that we analyze is the assumption that music can be freely shared, but can only be sampled in its entirety a limited number of times by users who have not purchased listening rights. The rights holder, therefore, can use digital rights management (DRM) techniques to limit how many times to allow a song to be shared and over what period of time. Allowing sharing an unlimited number of times over an unlimited time horizon would correspond to current "copyright unfriendly" P2P networks. Prohibiting all sharing would correspond to the record industry's apparent desired outcome for P2P networks. The basic research question is existence and nature of a profitable system of promotion to the music label using intermediate values.

The remainder of this paper proceeds as follows. In section 2 we provide background on the decision flow associated with our proposed P2P promotional model and a brief overview of the DRM infrastructure necessary for this system. In section 3 we develop a model of P2P promotion and discuss initial analysis of this model. In section 4 we conclude and discuss future research directions.

2. BACKGROUND

2.1 User Decision Flow

The user decision flow for our model is shown graphically in Table 1. The basic idea is to allow an intelligent agent acting on behalf of individual P2P users to recommend songs to a user based on the availability of songs on the network, and the similarity between the learned preferences of the user and the other peers offering these songs for promotion.

Specifically, in step 2, the intelligent agent determines the set of songs available for trial from the set of songs offered by other peers on the network that have been enabled for trial use (and are not already owned by the user) along with a rating for the song. This rating could either be implicit (e.g., based on the play count) or explicit (e.g., based on a manual rating provided by the user). In step 3, the agent uses information about the similarity in preferences between the user and the offering peer to predict the utility the user will get from each piece of trial-enabled content on the network. In step 4 the agent selects a playlist containing songs with the highest predicted utility. In step 5, the user listens to a subset of these songs, reports its responses to the agent, and decides whether to purchase any of these songs. The agent in turn uses this information to update both its library of the user's content and information on the user's preferences and the process starts over again.

2.2 Digital Rights Management

Currently, most digital music files fall into one of two categories. The most common category does not feature DRM and will (ideally) play on devices with the appropriate hardware and software (e.g., standard MP3 files).[ii] The second category includes some level of DRM preventing the file from being played on any device unless the correct permissions can be presented (e.g., AAC files used by Apple's iTunes Music Store).

Our promotional system relies on a middle ground between these two extremes where a customer with the correct permissions can play the song, but can also forward that song to friends for them to trial. Specifically, our system requires that customers who do not have permissions to play a particular song file can play that song for a period of time (where t=0 is equivalent to most current DRM systems, and t=∞ is equivalent to current non-DRM systems). This capability could be provided by standard DRM systems and is part of the XrML schema adopted by the in the MPEG-21 REL specification.

3. UTILITY MODEL (SINGLE VALUE, DISCRETE TIME, INCOMPLETE AVAILABILITY)

In our framework we are interested in modeling a user who gains utility from listening to music. An intelligent agent acting on behalf of the user (a peer) compiles the playlist to maximize the user's utility. This intelligent agent is working within a peer network to share and gather information that is used in making decisions about what music to transfer to the user. The user in turn interacts with the agent by providing information about likes and dislikes, whether to purchase a specific song, and how much time to allocate to listening to the agent's recommendations. If the user has good experience with his agent, he is likely to allocate more time to listening to the agent's recommendations. Finally, we have the music label, which is attempting to promote its music to maximize profit. In this section we introduce a formal model to describe the user's utility, the intelligent agent, the peer network, and the promotional decisions faced by the music label.

3.1 Intelligent Agent

In our framework music comes from two sources, either a user owned library or a set of promotional music. The source of the promotional music is either another peer or a direct promotion by the music label (perhaps through a radio station).

During each discrete period user i will listen to a set of songs, which we call the playlist. This playlist is created by an intelligent agent that acts on the user's behalf to collect the best music possible. We assume that the utility of the sth song is a linear function of its k observable attributes (a_{sk}), such as the genre, artist, period, album, and vocal versus instrumental; potential time trends, e.g., the more the song is heard the less benefit it brings; and a random effect:

$$U_{ist} = \sum_k \beta_{ik} a_{sk} + \varphi_i U_{is,t-1} + \varepsilon_{ist}, \qquad (1)$$

where ε_{ist} is independently and identically distributed. The consumer's utility from a playlist becomes:

$$U_{it} = \sum_{s \in \Pi} U_{ist}, \qquad (2)$$

where the playlist (Π_{it}) is defined as the set of songs selected by the agent to be played.

The dimension of the playlist (P_{it}) is selected by the user, but the agent optimizes this by choosing the best songs. Since utility is additive and the errors are i.i.d., this set is defined by the first P_{it} order statistics:

$$\Pi_{it} = \{U_{i,1:N_{it}}, \ldots, U_{i,P_{it}:N_{it}}\},\ \Theta_{it} = \{U_{i1}, \ldots, U_{i,N_{it}}\}, \quad (3)$$

where the Θ_{it} is set of all songs available from the user's library and promotional songs at time t and its dimension is N_{it}. Promotional songs are available from two sources. The first is directly from the music label, which we denote by $\overline{\Gamma}_t$, with dimension $\overline{\gamma}_t$. The second are those that are referred from friends, which we denote by $\widetilde{\Gamma}_t$, with dimension $\widetilde{\gamma}_t$. Hence, the set of promotional songs (Γ_t) is constructed from the union of these two sets:

$$\Gamma_{it} = \overline{\Gamma}_t \cup \widetilde{\Gamma}_{it} \quad (4)$$

We assume that regardless of source, promotional music is only available for the current period.[iii]

The set of songs in user i's library in period t is defined by Λ_{it}, and is made up of the songs in the previous period and those that are bought ($\mathbf{B}_{it}$) in the current period:

$$\Lambda_{it} = \Lambda_{i,t-1} \cup \mathrm{B}_{it},\quad \Lambda_{i0} = \varnothing \quad (5)$$

Notice that the user's initial library is assumed to be empty. Additionally, we assume that the library is always persistent. In other words once a user purchases an album it is always available. During each period a user generates an indicator variable that records their evaluation of the song (e.g., e_{ist}=1 if liked, =0 if not), and whether they want to buy the song (e.g., b_{ist}=1 if it is purchased, =0 otherwise). Hence, the set of purchased songs is defined by:

$$\mathrm{B}_{it} = \{s : b_{ist} = 1\}, \quad (6)$$

where the dimension of $\mathbf{B}_{it}$ is η_{it}. In summary, the choice set for the agent deciding on the recommended playlist defined as:

$$\Theta_{it} = \Gamma_{it} \cup \Lambda_{it} \quad (7)$$

Finally, a user's expected utility is:

$$U_i = \sum_t^{\infty} e^{-rt}\left(U_{it} - \nu(X_{it} - \rho\eta_{it})\right) \quad (8)$$

where X_{it} measures the consumer's total expenditures at time t, r is the discount rate, ρ is the price of a song, and $\rho\eta_{it}$ represents expenditures on music. In the present model, the user's problem at time t is to decide whether to purchase a song by deciding if their utility is increased by purchasing a song:

$$U_i(b_{ist} = 1) \geq U(b_{ist} = 0) \quad (9)$$

3.2 Network Effects

In our framework we assume that a user is participating in a social network. We assume that the ties in this social network are exogenous. This assumption is consistent with the present design of most hybrid P2P networks (e.g., Kazaa, Gnutella 0.6). However, some authors have proposed endogenously specified connections based on utility models and club goods (Asvanund et al. 2004). Including endogenously specified links between users based on mutual utility provided by the association would likely increase the benefits by our proposal, and would be a useful area for future research.

Formally, we assume that the ith peer is connected to a set of $\mathbf{Y}_i$ other peers, whose dimension is υ_i. During each period the peer will refer (or essentially promote) the music that she has purchased during the past L periods to other peers it maintains connections to. It does so by maintaining a list of songs and a rating of how often these songs have been listened to by the user. Thus, the promotional music available to a peer is equivalent to all music purchased by a user in the previous L periods:

$$\widetilde{\Gamma}_{it} = \bigcup_{j \in \mathrm{Y}_i} \bigcup_{l=1}^{L} \mathrm{B}_{j,t-l} \quad (10)$$

Notice that if L=1 then peers only share the last period's music, but if L is large then the peers are essentially sharing their entire libraries. L can be thought of as the flexibility that the music label provides its customers in distributing music, and can be enforced within several existing DRM systems as noted above.

Since peers do not have identical preferences, we introduce a similarity measure between peers. The similarity of the *i*th peer to the *j*th peer is defined as σ_{ij}, where $0 \le \sigma_{ij} \le 1$. This similarity measure is defined as the probability that *i* will rate an album favorably given that *j* has purchased the album. For simplicity, in this initial model we assume that this probability does not vary with the song's attributes.

3.3 Promotional Decision

Promotion results in awareness, and without awareness a user will never purchase an album. However, simply being aware of an album does not guarantee purchase. Instead purchase depends upon the expected value of the song. We assume that there are two methods of promotion available to the music label. The first is a direct promotional channel that allows the music label to send a message to all users, where this promotional decision is represented as an indicator variable:

Error! Objects cannot be created from editing field codes. (11)

However, the music label cannot selectively target users. Rather, it is forced to broadcast to all users at a charge of ω per user. Hence, the promotional set of music is defined as:

$$\overline{\Gamma}_t = \{s : \delta_{st} = 1\} \quad (12)$$

The second promotional alternative is to spend nothing and hope that the peer networks will help promote the album through word of mouth.

This broadcast results in an album that is only available for one period to the user, but otherwise is identical to the original product. The user does not know what the music label will promote, but it does have an expectation about the quality and quantity of the music that will be promoted. Hence, the user has to balance the amount of music to purchase with the guarantee that the music will be available, against listening to promotional music with uncertain availability and utility.

The profits for the music label of song *s* equals the total number of albums bought less the costs incurred for direct promotions appropriately discounted through time:

$$\Omega_s = \sum_t^{\infty} e^{-rt} \left(\rho \sum_i^M b_{ist} - M\omega\delta_{st} \right) \quad (13)$$

4. DISCUSSION AND EXTENSIONS

Our model envisions an environment where sharing is allowed and explicitly encouraged by rights holders because of its promotional value. Rights holders use standard DRM techniques to determine how many times and over what time period users can trial songs before making a purchase decision. Most current P2P networks facilitate unlimited trials and the record industry has advocated for an environment with no sharing at all.

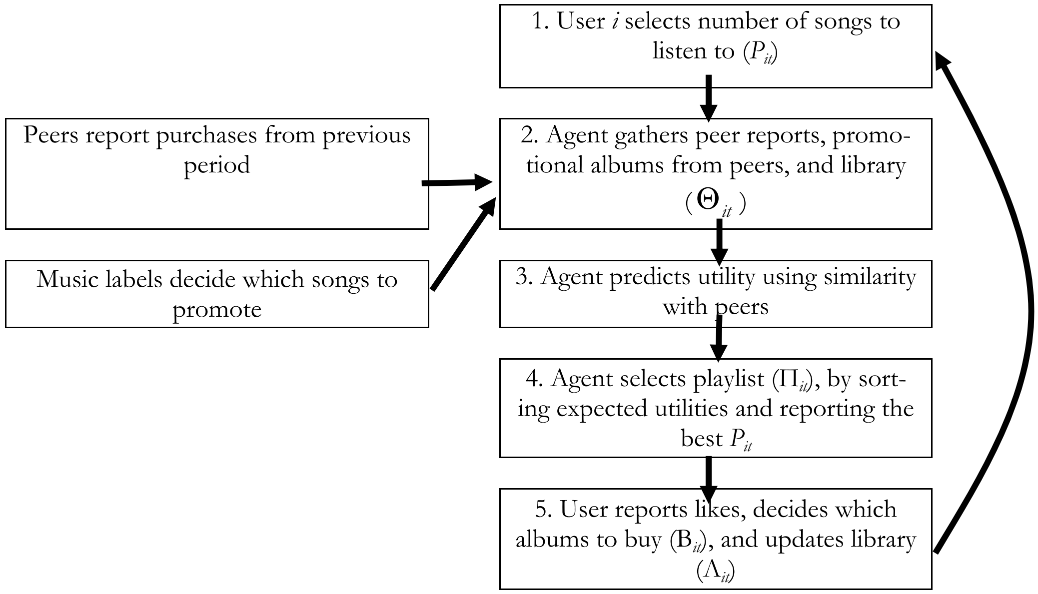
Our model allows us to analyze whether there are intermediate solutions to these licensing terms that would be provide rights holders with more profit over either of the extreme positions currently being advanced. Our initial simulations suggest such an intermediate optimal solution exists. We will be able to characterize these solutions in more detail in any potential workshop presentation.

Our research also raises other interesting questions that can be explored in the future. First, the use of anonymous peer ratings raises the possibility of malicious or self-interested ratings; for example, by record companies to promote their content at the expense of other companies' content. This could be a topic for mechanism design analysis to ensure accurate reporting. In this regard, accurate reporting by peers could be assisted through the use of implicit ratings (which are difficult to fake). Additionally, for a malicious strategy by record companies to be effective, the malicious peer would have to have a reasonable estimate of the target peer's utility function, which is difficult to obtain in our distributed setting. Another possible extension could involve how to manage distributed reputations in this environment or extending our exogenous ties between users to endogenous ties similar to those proposed in Asvanund et al (2004).

5. BIBLIOGRAPHY

Asvanund, Atip, Ramayya Krishnan, Michael D. Smith, Rahul Telang. 2003. Intelligent Club Management in Peer-to-Peer Music Sharing Networks. Working Paper, Carnegie Mellon University, Pittsburgh, PA.

Table 1. Flow diagram that illustrates steps made by a user and his intelligent agent each period.



Appendix: List of Notation

Category	Sym.	Description	Dim.	Category	Sym.	Description	Dim.
Indices	i or j	Index for user	M users	Users	β_{ik}	Part worth of attribute	
	s	Index for song	S songs		φ	Autoregressive coefficient	
	t	Index for period	T periods		ε_{ist}	Error utility	
	k	Index for attribute	K attributes		U_{ist}	Utility of song	
	l	Index for license	L periods		U_{it}	Utility for time period	
Song	a_{jk}	Attribute			U_i	Total Utility	
User Sets	Π_{it}	Playlist	P_{it}		r	Discount Rate	
	Λ_{it}	Library			X_{it}	Total Expenditure	
	$\overline{\Gamma}_t$	Public Promotion List	$\overline{\gamma}_t$		$v()$	Value function of outside good	
	$\tilde{\Gamma}_{it}$	Peer Promotion List	$\tilde{\gamma}_t$	Network	Υ_i	Peer set for user	υ_i peers
	Γ_t	Promotion List	γ_{it}	Promotional	ϱ	Song price	
	Θ_{it}	Choice set for Playlist	N_{it}		ω	Cost per user for direct promotion	
	$\mathbf{B}_{it}$	Buy decision	η_{it}		δ_{st}	Indicator variable for direct promotion	
					Ω_s	Total Song Profit	

[i] Financial support was provided by the National Science Foundation through grant IIS-0118767.

[ii] Note that in March 2004 Thompson/Fraunhofer III has recently announced a DRM-enabled version of the MP3 standard.

[iii] This could easily be extended within the model (and within current DRM systems) to allow the song to be available for a specified period of time. For simplicity, we only consider the single time period case in the present model.