

Two-Sided Value-Based Music Artist Recommendation in Streaming Music Services

Jing Ren^{*,+}, Robert J. Kauffman⁺, Dave King^{*}

*Singapore University of Social Sciences, +Singapore Management University, •Teuvonet Technologies LLC
jingren@suss.edu.sg, rkauffman@smu.edu.sg, daveking63@gmail.com

Abstract

Most work on music recommendations has focused on the consumer side not the provider side. We develop a two-sided value-based approach to music artist recommendation for a streaming music scenario. It combines the value yielded for the music industry and consumers in an integrated model. For the industry, the approach aims to increase the conversion rate of potential listeners to adopters, which produces new revenue. For consumers, it aims to improve their utility related to recommendations they receive. We use one year of listening records for 15,000+ Last.fm users to train and test the proposed recommendation model on 143 artists. Compared to collaborative filtering, the results show some improvement in recommendation performance by considering both sides' value in conjunction with other factors, including time, location, external information and listening behavior.

1. Introduction

Music streaming is characterized by: streaming services; listeners; and music labels, indie musicians, songwriters, and producers. Streaming services are provided by middlemen that connect listeners with music. On one side, music labels and musicians use streaming services to upload their songs or albums. The services pay copyright fees based on pay-per-stream pricing, similar to paid software. On the other side, listeners subscribe to streaming services to listen to music. This enables access to tracks and albums based on subscriptions. In general, the more listeners listen, the more the music labels benefit from the streaming service revenue.

In 2016 music streaming revenues, primarily from consumer subscriptions, accounted for 51% of all recorded music revenues which totalled US\$7.5 billion, exceeding the revenues from the sale of various physical formats and digital downloads [1]. By 2017, the share from streaming services was ~65% of US\$8.7 billion or US\$5.5 billion. How much of this revenue channels to the music labels and indie musicians and writers depends on which service plays their songs. For example,

Apple Music pays \$0.0078 per play, Google Play \$0.0061, Spotify \$0.0040, Pandora \$0.0013 and YouTube \$0.0007 [2]. While these fees are fractions of a cent, they add up rapidly so a large portion of the revenues end up flowing to the major music labels, more specifically to the three major music labels (Universal, Sony, and Warner). In 2017, these labels received US\$14.2 million per day or US\$5.2 billion for the year [3]. While this may seem like good news for musicians, in general it's not. At Apple's fee rates, the highest among the top services, if a song has 100,000 plays (above the average), this translates into about US\$800. It takes an enormous number of plays/streams to produce a living wage. This may not be critical for artists with large listener bases and other sources of income, but it is serious for musicians and songwriters.

If budding or famous artists want to promote their music via streaming services, it helps to understand their listeners, who they are and when and how many times they listen to the artists' music, and the kinds of information that can be leveraged. However, there is no direct or active communication between artists and listeners in the current streaming music scenario. In the current system, music recommendations assist listeners in identifying new music and artists. They help the services expand and retain listeners, and increase revenues. Current approaches are based on a combination of theory [4] and empirical studies [5, 12], but they only work on the consumer side. They don't provide many artists, especially indie musicians, with a way to control, engage in, or benefit from the process.

We seek to improve online music recommendations by considering the value that can be obtained on both the consumer and provider sides. We develop a two-sided value-based artist recommendation method. It combines the value yielded for the music artists and consumers in an integrated model. For artists, the approach aims to increase the conversion rate of potential listeners to adopters. At the same time, for consumers, it aims to improve listening satisfaction related to recommendations they receive. It involves a new algorithm in the artist promotion context, with time, geolocation, and promotion information.

We used one year’s listening records for 15,000+ Last.fm users to train and test the proposed two-sided value-based recommendation model for 143 artists. Compared to collaborative filtering, a widely-used recommendation algorithm, our results show an increase in recommendations’ evaluation performance, including *conversation rate*, *recall*, and *value obtained*. Our work offers new knowledge and paves the way for *on-demand music promotion* for artists.

2. Literature Review

2.1. Music Recommendation

Music recommendation techniques can be categorized into four approaches: content, collaboration, context awareness, and hybrid methods:

- *Content-based methods* assume that listening preferences can be extracted from the music content of the songs in a user’s listening history. New songs are recommended to a user based on the similarity between new and the prior songs [5].
- *Collaboration-based methods* leverage users with similar listening histories who share listening preferences. New songs are recommended by referencing preference similarities. Such methods have been adopted in industry (e.g., *k*-nearest neighbors [6] and matrix factorization [4]).
- *Context-aware methods* focus on music-relevant information other than the music content. “Context” refers to the social, user, artist or usage context, including location, time, weather, etc. [7, 20]. This category is attracting more attention as big data have become available and web technology has transformed the music listening setting.
- *Hybrid methods* combine techniques such as content-based and context-aware algorithms, by merging music semantics with venue information into a latent topic model [8], or combining content and collaborative-based approaches.

With AI and deep learning, new recommendation methods are emerging. For example, researchers have recently tried to leverage the power of *convolutional neural networks* to improve recommendation methods, such as content, and collaborative filtering [26, 27].

Table 1. Comparison of Recommendation Methods

APPROACH	BIAS	TECH	ACCUR	BUS VAL
Content [5, 26]	×	×	✓	×
Collaboration [4, 6, 27]	✓	×	✓	×
Context [7, 20]	Depends	Depends	✓	×
Hybrid [8]	Depends	Depends	✓	×

Notes: BIAS = popularity bias. TECH = web technology. ACCUR = recommendation accuracy. BUS VAL = business value.

In Table 1, we compare these four categories of recommendation methods. *Popularity bias* reflects the long-tail phenomenon in music listening, and this issue exists in collaboration-based methods. *Web technology* involves the interaction of multiple channels and emphasizes context-based methods, but it may also involve hybrid methods. All four categories focus primarily on *prediction accuracy*, though none has considered the *business value* of music recommendation.

With respect to business value, in real streaming scenarios there is no guarantee that better prediction performance will translate into higher conversion rates and produce higher *average revenue per user* [11]. In most cases, either a binary variable or rating scale is used to indicate users satisfaction. In either case user preferences are short-term. Yet, music is a durable product. One listener may play the song 100 times in a month, while the other listener may only play it 5 times. The strength of the value is different for each of the listeners and also different for the music artists. Binary or rating measurement fail to capture listener utility in the longer term. Instead, business revenue patterns may offer better clues for system design in the longer term.

As listening to streaming music increases and the market expands, the industry has tried to engage in collaboration with streaming music services and to design strategies to maintain their effectiveness in the market [9]. They are now considering several aspects when designing music promotion strategies. First, pay-per-stream pricing may force them to not just consider recommendation accuracy, but also to match business value to recommendations (e.g., how many times users listened to music of a similar nature). Second, in an open environment for accessing a massive amount of music-related information, interactions among multiple channels may also affect music promotion [10, 18, 24].

In response to the limitations of recommendation approaches, our model leverages multiple factors to measure the business value and other benefits for the artist and record labels, and the listener side.

2.2. Utility Theory and Recommendation Value

In economics, *utility* is a proxy measure of one’s preference over a set of goods or services. It represents the satisfaction experienced by the consumer for a good. It is an important concept that serves as the basis for *rational choice theory* [14]. One cannot directly measure the benefits that people gain from the consumption experience. Instead, economists use indicators from a good or service that people consume to represent and measure the gained utility.

For recommender systems, past research has demonstrated the existence of important economic side-effects [12]. For example, personalized recommendations can

help increase consumer willingness-to-pay for digital music promotions. Other work has analyzed the relationships between provider profit and consumer utility surplus, as well as satisfaction and predictive accuracy. Panniello et al. showed a positive impact from the balance between accuracy and profit for online purchasing behavior, if it does not hurt the extent of consumers' trust [13]. Prior work has highlighted the necessity to balance benefits of providers and consumers in recommendation system design but has not modeled it. A reason is that, under the streaming services scenario, the characteristics of streaming music consumption and its revenue patterns may be an obstacle to calculating two-sided value.

On one side, for consumers, Varian [15] pointed out the importance of the *theory of the consumer*, in which utility describes consumer preferences. Music is a digital experience good, so consumers cannot gain any utility until they have listened to its content. According to the theory, consumer utility for a music track represents satisfaction from a listening experience, minus the search cost they expend to find the recommended music. In empirical research, determining how to measure utility is a challenge. Different listeners may experience different satisfaction levels and search costs for the same music product.

For the other side, providers, the *theory of the firm* is applicable. This is because the provider is involved in the production and marketing of music, not its consumption. Providers wish to maximize their profit earned from sales revenue beyond the cost of producing and selling music to appropriate consumers and listeners. So, the theory of the firm is applicable. Take Netflix as an example. Business analysis of its recommendations [16] have shown that personalization and recommendation can help to maintain subscribers' loyalty and reduce the number of members who decide to stop a service. Good recommendations have helped Netflix to create value by saving it more than US\$1 billion per year in its effort to acquire new consumers, simply because it has reduced subscription cancellations.

Netflix has used subscription fees it earns due to the recommendations it makes to consumers to measure their beneficial effects on business value. A reason for this approach, we believe, is that Netflix recognizes that it is hard to capture the benefits that are produced for each video that its users view. This is somewhat less true for other e-commerce products, such as clothes, cameras and food, where it is possible to count the number of units sold and the underlying cost to support such sales. In contrast, for streaming music like Netflix, it is not easy to measure revenues and average costs of supplying the music. Streaming music is not priced, nor is it obvious what the costs are since music that is acquired for streaming involves royalties for the music artists and

fees for their music labels. Therefore, for the provider side at least, it makes sense to calculate the value of recommendations that enhance consumption, while ignoring the associated costs. Costs are difficult or impossible to observe without direct access to the music labels' data sources, a roadblock for empirical research.

We next discuss how we leverage utility theory to address the two-sided recommendation value problem.

3. Two-Sided Value for Recommendations

3.1. Problem Description

The goal is to design personalized promotions for streaming services that assist labels and artists to identify listeners with the most potential. Imagine that a music label launches an artist's new song on Last.fm or Spotify to attract new listeners and wishes to increase streaming track volume in one month. This involves achieving effective listener selection for high ROI.

This is different from traditional recommendation, which seeks to find the most suitable artists, songs, or albums for a listener. Here, the target shifts from the consumers to the providers. It is useful for the music industry to promote niche music products or independent musicians. These products often are difficult to find or have not been recommended due to popularity bias. Determining how to leverage both the strength of the streaming platform and external information to assist in the search for new listeners in a short time can increase streaming volume.

The challenges are: (1) to determine how artists know who the targeted listeners are; and (2) to promote artists and delight listeners with recommendations based on artist promotion requirements, such as increasing streaming track plays or adding new listeners via broadcasting artists' future album release. This supports personalized recommendations for artists to identify potential listeners based on the two-sided value.

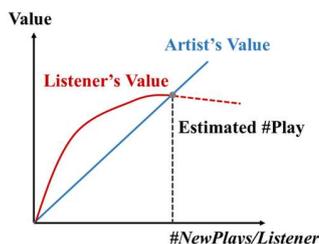
Streaming services are *semi-closed* (which encourages sharing of social information) but open to external information discovery by users [18, 24]. Two categories of external information can be leveraged to do music promotions for an artist: *Music Content* and *Non-Music Content* [22, 24]. *Music Content* includes new albums, new songs, and new music video releases. *Non-Music Content* includes artist life news, TV show appearances, live performances, festivals and related non-music release information. For each, we construct a two-sided value model for artist recommendation.

3.2. Basic Model

Figure 1 provides an overview of the two-sided

value model. The red line represents the listener's value and the blue line is the artist's value. The x-axis shows the time a listener has played an artist's music. The y-axis represents cumulative value for the artist. After time passes, the marginal value (utility) for a listener to listen to additional music will monotonically decrease (possibly to 0), and the listener may no longer add more listening content. At the same time, an artist's value is likely to continue to increase with the #NewPlays, even though for the listener, the gained value of one more listening is already decreasing.

Figure 1. Listener, Artist Value for Music Listening



By considering the value of both sides, our goal is to find the listening level for each candidate listener i to maximize the total value of the top- N recommended listeners for artist j as follows:

$$TotVal_j = \text{Max}_{TopN} \sum_i (ListenerVal_{ij}(q_{ij}) + ArtistVal_{ij}(q_{ij})) \quad (1)$$

Total value, $TotVal_j$, is a function of the quantity q of streaming music consumed that produces value. $ListenerVal_{ij}$ is listener i 's value from listening to artist j , and $ArtistVal_{ij}$ is the artist's value gained from listener i 's streaming of her music. q_{ij} represents the amount of listening by the listener i to artist j . This indicates that the selected top- N candidate listeners will yield the maximum total value by considering both sides. Next, we describe how to calculate the two sides' value.

Artist's value, $ArtistVal_{ij}(q_{ij})$. Customer satisfaction is a good predictor of firm business performance [17]. So, for the assessment of artist value, we considered two aspects. (1) Artists hope to attract loyal listeners who will continue to stream the artist's music over a period of time, and not just sample it and never return. This revenue source can be calculated with the pay-per-stream quantity as $a_0 q_{ij}$. (2) The other value source is from the potential listeners who can be affected via social influence and become aware of new artists [18]. We label this value as p_{ij} , the number of listener i 's friends who have not listened to artist j .¹ This yields:

$$ArtistVal_{ij}(q_{ij}) = a_0(q_{ij} + \text{Pr}(q_{ij}) \times p_{ij}) \quad (2)$$

¹ These two aspects represent direct and indirect listeners attracted by the recommendation. No matter which type, the value of pay-per stream a_0 is static. We couldn't access the

$\text{Pr}(q_{ij})$ is the probability for user i listens to artist j for q_{ij} times. In this case, if two listeners listen to the same amount of music, then the one who has more social friends will be more attractive for an artist's promotion, because of the long-term social effects.

Listener's value, $ListenerVal_{ij}(q_{ij})$. For the listener's value, there is no standard measure. Traditional music recommendation has used a binary or rating variable to represent the satisfaction level of the listener with the recommended music. This offers short-term feedback on the recommendation and ignores the diverse listening behavior on the recommendation. For example, when two listeners give the same rating to an artist but with different listening times (such as 100 vs. 5), the strength of the value obviously is different for each of them and also different for the artists.

Measuring the utility that a listener gains from listening to an artist q times is more difficult. It is governed by the *law of diminishing marginal utility*. So, when a person increases her consumption of a product, there will be a decline in marginal utility that she derives from each additional unit of it. This is also true for music listening. If utility does not decline as listening increases, this will be surprising, since most listeners stop listening over time. This reflects a utility decline. Our model does not calculate marginal utility directly: it estimates the likely listening quantity before a listener decides to stop.

There are various functional forms for utility. We used *King-Plosser-Rebelo utility* for e-commerce product recommendation based on Zhang et al. [19]. The listener's value is shown in Eq. 3, which yields 0 when the listening time is zero, or $ListenerVal_{ij}(0) = 0$ from:

$$ListenerVal_{ij}(q_{ij}) = a_{ij} \ln(1 + q_{ij}) \quad (3)$$

Here, a_{ij} is the weighted effect of user i 's utility for artist j 's music. It can be a binary rating or a probability for listening utility, with Eq. 1 rewritten as:

$$TotVal_j = \text{Max}_{TopN} \sum_i (a_{ij} \ln(1 + q_{ij}) + a_0(q_{ij} + \text{Pr}(q_{ij}) \times p_{ij})) \quad (4)$$

a_0 is the baseline value an artist can gain through user listening, represented by pay-per-stream revenues. We set $a_0 = 0.004$ based on the average pay-per-stream revenues generated by the major streaming music services. Detailed estimation of a_{ij} , q_{ij} , $\text{Pr}(q_{ij})$, and the recommendation approach works are presented next.

3.3. Model Specification

To calculate total value, listening quantity value q_{ij} is an important factor. Zhang et al. [19] assumed that q_{ij}

whole Las.fm dataset to test the indirect effect, so we leveraged the social relations to estimate potential artist's value.

is a random variable independent of weighted utility a_{ij} at first. But later, they used a Poisson distribution to describe q_{ij} for each consumer, as well as the collaborative filtering method to estimate utility a_{ij} . The estimation results show that the quantity is associated with the utility level. Due to the nature of the Poisson distribution, $\widehat{a}_{ij} = \overline{Q}_{ij}$. Music listening has similar but different characteristics. Listening quantity q_{ij} is correlated with utility a_{ij} , but may be affected by other factors, such as listener's listening context. The mean of the Poisson distribution is equal to its variance, which is unsuitable for music listening. Our dataset's distribution is over-dispersed, with the variance of $\#NewPlays/Listener$ larger than its mean.

Based on these characteristics, for each listener the utility and revenue estimation results can be combined to create a unified value measure. Moreover, collaboration-based and context-aware functions can be estimated, based on personal listening behavior and artist context and promotion information via Eqs. 5 and 6. Combining a_{ij} and q_{ij} controls the number of missing potential listeners, for example, when $a_{ij} = 0$.

$$TotVal_{ij} = \ln(1 + q_{ij}) + a_0(q_{ij} + Pr(q_{ij}) \times p_{ij}) \quad (5)$$

$$q_{ij} = f(a_{ij}, ListeningCharateristics_i, ListeningContext_{ij}) \quad (6)$$

Estimation of q_{ij} . As Eq. 6 shows, to allow listening quantity to vary, q_{ij} is a function of listener utility a_{ij} , listening characteristics and context (Table 2).

Table 2. Covariates Used for q_{ij} Estimation

NOTATION	CONTEXT INFO
Collaboration Estimation	
a_{ij}	Listener i 's weighted utility for artist j 's music
Listener's Listening Characteristics	
$ListeningScale_i$	# of artists user listened to
$ListeningBreadth_i$	User diversity of music listening across artists
$TasteSimilarity_{ij}$	Taste similarity of user for an artist's music
Listening Context Information	
$MajorLabel_i$	Whether artist connected to major music label
$PopLast.fm_i$	Top chart popularity, Last.fm, 2005-2013
$PopBill_i$	Chart popularity, Billboard Hot-100, 2005-2013
$ArtistExtInfoType_i$	Type of external info released on an artist, <i>Music Content</i> and <i>Non-Music Content</i> promotion info.
$Artist\#ExtInfoRelease_i$	# of artists with external info when user listened
$CtryExtInfo_i$	(1, 0) if user country is U.S.; (0,1) if it is English-speaking; (0, 0) otherwise

Note. $CtryExtInfo$ considers geolocation and language differences.

Different promotion information has a different effect on music diffusion [24]. For example, all else equal, when artist-related *Non-Music Content* is released in the

U.S., like a live TV show, listeners there may listen more times than those who are abroad. So, for each promotion category, *Music* and *Non-Music Content*, we defined new listening quantity as a function of collaborative effect, context (artist popularity, external info, geolocation), and listening characteristics (taste, scale, breadth) based on prior work [18, 22, 24].

Negative binomial regression analysis is used with multiple covariates to estimate q_{ij} . For each artist-listener pair, the unit is (q_{ij}, X_{ij}) , where $q_{ij} \geq 0$ is listening quantity and X_{ij} is a $k \times 1$ covariate vector to describe the listening characteristics and context information. When either category of promotion occurs, we focus on how the listening quantity varies with the covariates. The conditional mean μ and variance σ are:

$$\mu = E\{q_{ij}|X_{ij}\} = \exp(X_{ij}^T \beta) \quad (7)$$

$$\sigma = \mu + \alpha \mu^2 \quad (8)$$

For each external context type, β is a $1 \times k$ set of parameters estimated by Equation 7. α is a dispersion parameter of the negative binomial model. It involves a gamma density function, to represents the probability that $q = q_{ij}$ when the listener's observation is X_{ij} :

$$Pr(q = q_{ij} | X_{ij}) = \frac{\Gamma(q + \alpha^{-1})}{\Gamma(q+1)\Gamma(\alpha^{-1})} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu}\right)^{\alpha^{-1}} \left(\frac{\mu}{\alpha^{-1} + \mu}\right)^q.$$

Maximum likelihood estimation was implemented to obtain the parameters μ , α and β . For each promotion category, 80% of the artist-listener pair data were used to estimate the model's β parameters. This allows the acquisition of expectation \widehat{q}_{ij} and $Pr(q_{ij})$ for new listeners, and is used to calculate total value.

Estimation of a_{ij} . The parameter a_{ij} describes the weighted utility of listener i for artist j , so we estimated its value for each artist. For artist j , we used a listener's *k-nearest neighbors* (KNN) with a collaboration-based method [6] to estimate the potential listening utility of a new listener. KNN is a memory-based method to estimate binary or rating feedback and supports listener- and artist-based estimation. Our dataset has a relatively small number of artists though, so using artist similarity to estimate the number of listeners may result in somewhat more bias than listener similarity. Thus, we used a listener-based KNN method.²

Pearson correlation is used to calculate the similarity, Sim , between a potential listener i and another listener $v \in V$, V is set of current listeners of artist $j \in J$, J is the set of artists, with r the listening time to each artist:

$$Sim(i, v) = \frac{\sum_{j \in J} (r_{i,j} - \bar{r}_i)(r_{v,j} - \bar{r}_v)}{\sqrt{\sum_{j \in J} (r_{i,j} - \bar{r}_i)^2} \sqrt{\sum_{j \in J} (r_{v,j} - \bar{r}_v)^2}} \quad (7)$$

the small dataset, and matrix factorization will be considered in future research to further improve the model.

² Collaborative filtering (CF) and its modifications are the most widely used recommendation algorithms in industry (Amazon, Netflix, Youtube). We adopted KNN CF because of

a_{ij} is estimated based on the weighted sum of the top- k neighbors' listeners with higher similarity to listener i . This is collaborative filtering, which indicates users with similar taste will adopt similar products.

$$a_{ij} = \frac{\sum_{v \in \text{top-}k \text{ in } v} r_{v,j} \times \text{Sim}(i,v)}{\sum_{v \in \text{top-}k \text{ in } v} \text{Sim}(i,v)} \quad (8)$$

3.4. Listener Recommendations for an Artist

Traditional recommendation involves selecting a list of songs or artists, to enhance the listener's utility. We selected a list of listeners for an artist in a context by considering the consumers' and provider's value, and used a value-based ranking to realize the recommendations [20]. For each artist j , the goal is to find a set of listeners to maximize total value for both sides. The detailed procedure is: for each artist promotion, we first estimate the value for listening to the artist based on Eqs. 2 and 3 for each candidate artist-listener pair. Next, Eq. 1, two-sided value, is used to rank artist-listener pair's value, then the top- N candidate listeners that maximize total value are selected as the recommendations for the artist.

4. Research Setting and Data

We used a subset of Last.fm's user data collected by using Last.fm API. It contains listening records of 15,607 seed users during Jan.–Nov. 2013, with 1,796,932 listening records.³ We also collected the listening record of seed users' friends for *Artist Value* measurement. The listening matrix of artist-listener observations was sparse, with an average density of only 3.22%. Among the listened-to artists, we selected 143 who had external information released in the U.S. during the period for experiment. The related external information was summarized into the two promotion categories: *Music* and *Non-Music Content*.

The listening records for the three months prior to the listening observations (before the promotion was released) were used for estimating the effects of collaboration and listening characteristics for the covariates on listening quantity. We then used a month of listening records following the release of info to test the model for each promotion category.

The focus is on recommending new listeners to an artist. These listeners had not listened to the artist in the previous period but were more likely to listen to after promotion information was released. Descriptive statistics for new listeners, and the listening times for each listener and each artist are shown in Table 3. There is

³ The selection period and covariates are based on our work. Compared to 2013, the basic structure of streaming music is

diversity in attracting new listeners by different artists when promotion information is released.

Table 3. Descriptive Statistics Related to Listeners

	MIN	MAX	MEAN	SD
#NewListeners/Artist	6	681	163	152
#NewPlays/Artist	67	22,943	2,284	3,361
#NewPlays/Listener	1	4,623	14	69

Notes. Obs.: 23,309 new listeners' records on 143 artists, obs. time range: one month after promotion information released. #NewListeners/Artist: # new listeners of an artist in the time range. #NewPlays/Artist: total # times new listening occurred to on an artist. #NewPlays/Listener: # times of a new listener played an artist's music.

In this dataset, on average, an artist attracted 163 new listeners in the one-month period after new external information was released. The average number of times that a new listener listened to an artist was about 14. The listening diversity observed, based on a standard deviation greater than or equal to the mean of the distribution, indicates the effects of different types of external information. Thus, for music promotion, it is necessary to effectively identify the targeted 163 new listeners on average from the candidate pool of 15,607 listeners.

5. Experiments and Results

We next investigated the use of the proposed approach for finding new listeners to an artist when promotion information was released. For each information category, *Music* and *Non-Music Content*, the corresponding artists were randomly segmented into 5 folds. A 5-fold cross validation was run to obtain the recommendation results. In each training and testing dataset, the sizes of users' listening records on the corresponding artists are shown in Table 4.

Table 4. Music Information Training and Testing

Fold	MUSIC INFO		NON-MUSIC INFO	
	Train	Test	Train	Test
1	1,036	281	649	164
2	1,043	277	647	166
3	1,044	276	653	160
4	1,036	281	649	164
5	1,036	282	655	158

Notes: Data for 88 artists with *Music Content Info* were included in each round; 70 were for training, and 18 for testing. Also, 55 artists were included with *Non-Music Content Info* in each round, 44 for training and 11 for testing. All entries in 000s.

In the training step, for each category of promotion information, we used users' listening count for the artist's music (#NewPlays/Listener including 0) to estimate a negative binomial model, with the covariates

similar, but with more listeners now, newer data will support understanding collaboration and artist promotion better.

listed earlier. In the testing step, the model was further used to estimate the listening quantity q_{ij} and probability $Pr(q_{ij})$ for the test listener i of artist j . Finally, the top- N recommended listener list for each artist in the test dataset was proposed based on maximizing the two-sided value: listeners' utility and provider's revenues.

5.1. Evaluation Measurement

$\#NewListeners$ and $\#NewPlays$ are used as dependent variables for this evaluation. Three measures were used to test the performance of the modeling perspectives. *Conversion%* ($C\%$) measures the percentage of successful recommendations in the top- N listener list. *Recall%* ($R\%$) measures the percentage of how many new listeners were found to be in top- N listener list. And *Value%* ($V\%$) is the percentage of the total streaming plays for the value-maximizing recommendations in top- N listener list. Different N values, 100, 1,000, 2,000, ..., 7,000, were selected to observe the performance. The quantitative measures are:

$$C\% @ N = \frac{\#NewListeners \text{ in } TopN}{N}$$

$$R\% @ N = \frac{\#NewListeners \text{ in } TopN}{Total \#NewListeners}$$

$$V\% @ N = \frac{\#NewPlays \text{ in } TopN}{Total \#NewPlays}$$

A traditional collaboration-based recommendation method was selected as the baseline for analysis and comparisons. Listener-based k -nearest neighbors (KNN) was implemented, and potential listeners were ranked based on the weighted utility a_{ij} value. We used $k = 15$ as the number of neighbors parameter. This achieved stable performance, compared to $k = 5, 10, \text{ or } 20$. Besides comparing the numerical differences in the measurements, we also used t -test to compare the statistical improvement of our method with KNN.

5.2. Performance Comparison

Table 5 shows the statistical comparison between KNN and our value-based methods. For *Music* or *Non-Music Content*-related promotion, the proposed value-based method performs better than the baseline method.

For $C\%$, the value-based method is around 2 or 3 times the value of KNN in terms of conversion of customers. Although both methods found a similar number of potential new listeners when $N = 7,000$, the value-based method did so faster, while KNN was still in the process of completing its computation. For example, for the Top-1,000, the $C\%$ for the value-based method in the *Music Content* context was $\sim 2.3\%$, while for KNN

it was lower at $\sim 1.5\%$. This means that, on average, the value-based method was able to find around 23 new listeners in the Top-1,000 recommendation, while KNN was only able to find 15.

Table 5. Comparative Performance of the Methods

MUSIC CONTENT PROMOTION						
Top- N	KNN			VALUE-BASED		
	$C\%$	$R\%$	$V\%$	$C\%$	$R\%$	$V\%$
100	1.66	0.97	1.60	2.71**	1.84***	2.11
1,000	1.47	9.30	11.27	2.34**	14.88***	15.20*
2,000	1.48	18.30	20.11	2.12*	27.07***	28.83***
3,000	1.49	26.77	29.10	2.01*	37.60***	40.09***
4,000	1.52	35.75	39.41	1.93*	47.99***	51.21***
5,000	1.51	44.33	46.29	1.84	57.22***	60.57***
6,000	1.51	53.23	54.51	1.77	66.27***	70.03***
7,000	1.53	62.58	64.21	1.70	74.06***	75.81***
NON-MUSIC CONTENT PROMOTION						
Top- N	KNN			VALUE-BASED		
	$C\%$	$R\%$	$V\%$	$C\%$	$R\%$	$V\%$
100	1.35	0.82	0.45	3.96***	2.93***	2.53**
1,000	1.57	8.97	7.67	2.75***	17.75***	17.37***
2,000	1.65	18.58	16.53	2.45*	30.99***	33.28***
3,000	1.61	27.18	23.54	2.25*	41.96***	43.80***
4,000	1.60	36.27	33.48	2.11*	51.97***	52.60***
5,000	1.58	44.71	40.57	1.99	60.08***	60.86***
6,000	1.57	52.98	47.54	1.89	68.34***	69.11***
7,000	1.40	61.80	54.14	1.81	75.38***	75.30***

Notes: Statistical results are from t -test between KNN and our value-based method. Signif. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

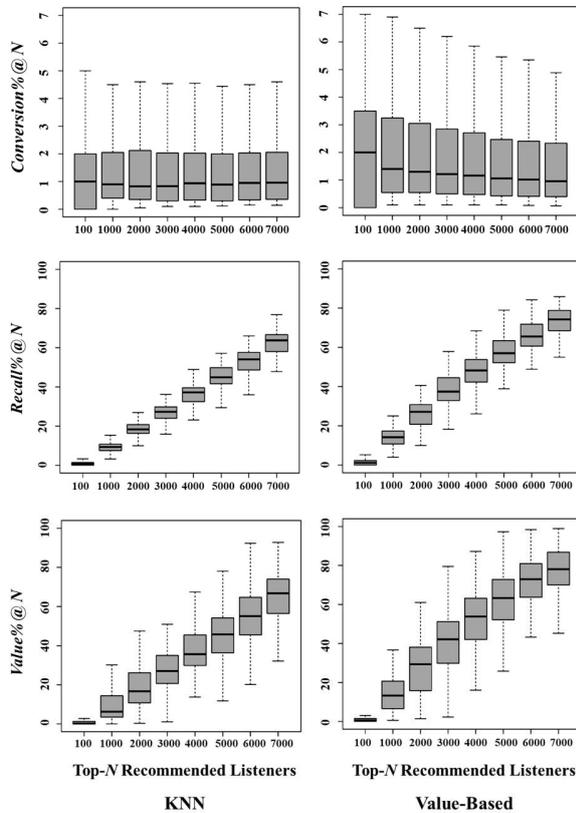
For $R\%$, the value-based method was better than the baseline method. It retrieved $\sim 74\%$ of all listeners in the top-7,000 recommendations, while KNN only obtained $\sim 63\%$. Similar conclusions for $V\%$, value ($\sim 76\%$ vs. $\sim 64\%$) were obtained. Though the $C\%$ was higher but not statistically improved for the top-7,000 ($\sim 1.7\%$ vs. $\sim 1.5\%$, and $\sim 1.8\%$ vs. $\sim 1.4\%$), our method found listeners and those who can bring more value to an artist.

These findings are useful for personalized music promotion and recommendation design for a specific artist. This can help musicians with small listener bases, and less money to invest in music promotion. For them, our model can assist the streaming services to design on-demand music promotion, which can find top- N listeners with the most potential based on artist's specific requirement, satisfy listener preferences, and maximize the possible pay-per-stream revenue value the artist can gain. At the same time, streaming services can gain ad revenues and potential subscribers through better recommendation services.

Table 5 also shows the performance of the value-based model for two categories of external promotion information. Each category of external information has subtypes, so they may exhibit diverse performance levels. Boxplots for each method are presented in Figure 2 to show the maximum, minimum, median, and standard

deviation performance of *Music Content Info* promotion.

Figure 2. Boxplots for *Music Content Promotion*



The entries in the left column are for KNN, and the right-column entries are for the value-based method. For *Music Content* promotion, there is obvious diversity in *C%* and *V%*. *R%* has less diversity though. This means different subtypes of external information may have different levels of performance when they use the same recommendation method. *Non-Music Content Info* has similar conclusion (omitted due to page limits). It is worthwhile to check the recommendation performance of each subtype. Sample results of the performance on the top-1,000 and top-4,000 for each sub-type of *Music Content Info* Promotion are shown in Table 6.

For the *Music Content Info* subtype, *Type3-Music-Video Release* had the best performance but a large standard deviation on conversion rate *C%*. it was higher because this kind of external information typically has the largest impact on attracting new listeners. A possible reason for the large standard deviation is the small number of observations – only 5 artists – in this subtype. More data are needed to solidify this result. *Type2-Album Release* had the worst performance on *C%* but smaller standard deviation. There were more observations in this subtype, with 54 artists. But, this subtype

apparently is not as effective as *Type3* in attracting new listeners, so there was less conversion. *Type1-Single Song Release* had best performance on recall *R%*, this advantage was moderated in our value-based method. For example, in top-4000 recommended listener, there was ~3.5% higher *R%* compared to *Type3*, this difference was reduced to ~2.4%.

Table 6. Performance: *Music Content Info* Promotion

KNN	N = 1,000 Mean (StDev)			N = 4,000 Mean (StDev)		
	Perf	C%	R%	V%	C%	R%
Type 1	1.47 (1.55)	9.52 (3.52)	10.51 (11.69)	1.55 (1.25)	38.04 (5.96)	38.13 (15.90)
Type 2	1.40 (1.65)	9.30 (4.77)	11.60 (13.70)	1.43 (1.60)	34.59 (6.86)	38.12 (15.88)
Type 3	2.20 (2.55)	8.02 (3.29)	12.04 (8.93)	2.23 (2.55)	34.56 (6.75)	43.12 (8.50)
Value	N = 1,000 Mean (StDev)			N = 4,000 Mean (StDev)		
	Perf	C%	R%	V%	C%	R%
Type 1	2.41 (2.01)	15.61 (5.16)	15.72 (11.08)	2.04 (1.73)	49.45 (7.90)	50.77 (16.71)
Type 2	2.20 (2.50)	14.48 (5.00)	14.83 (11.62)	1.79 (1.99)	47.30 (9.65)	50.75 (20.34)
Type 3	3.48 (3.79)	14.95 (7.19)	16.25 (12.79)	2.73 (2.99)	47.03 (12.19)	58.77 (5.27)

Notes: Obs.: 88 artists with *Music-Content External Info*. Perf: Performance. *Type1-Single Song Release*: 29; *Type2-Album Release*: 54; *Type3-Music Video Release*: 5.

In this case, although the artists' characteristics (e.g., popularity, major label) were considered when estimating the model, there were not enough observations to estimate a balanced model for every subtype and artist. This can be addressed if we model each subtype separately when a larger dataset is available. A similar conclusion can be reached for *Non-Music Content* promotion.

6. Discussion

Most existing music recommendation approaches adopted by streaming music services, such as Last.fm and Spotify, are focused on the listener side. None has tried to design *personalized artist music promotions* for a specific artist. This is useful, as streaming music has generated more and more market revenue, especially for indie musicians and song writers. Their music is usually not that easy to find among the available music choices, even when there is support available. So, there is value in understanding personalized artist promotions.

We showed that a value-based music promotion method can be used to assist the artist to target potential listeners when an artist's external information is released. It considers the listener's utility, with artist revenue from customer listening and pay-per-stream as a basis for measuring value. The findings provide new de-

sign thinking for music recommendations and personalized artist promotions in online music platforms.

First, considering two-sided value and external promotion information can improve an artist's music recommendation accuracy in the streaming scenario. The increases in $C\%$, $R\%$ and $V\%$ confirm that the proposed value-based method can efficiently assist artists to find new listeners when they use various promotion strategies, including new album releases, or news about upcoming concerts. This method is different: it identifies new listeners who never listened to the artist and can reactivate intermittent listeners who listened before but stopped for a spell. Finer-grained modeling for different types of external information or a specific music genre can improve the performance further. This paves the way for *on-demand music promotion*.

Second, another stream in music recommendation research is *influential user identification*. The goal is to identify the "big" users who can influence others to adopt a product or information in the presence of some social structure [25]. However, this is still an indirect effect. In this study, we worked toward a combination of direct and indirect promotion by considering listener utility and the number of her social friends. The latter is a proxy to measure potential revenue a listener may generate over time. Results confirmed that combining the direct and indirect approach will be more effective.

Third, through estimating potential two-sided value and then making a comparison with realized value [23], on-demand recommendation design can address all three key elements of streaming music. Artists can pay some amount of money to the services for a specific music promotion strategy. Streaming services can gain ad and subscription fees, while listeners can enjoy better recommendation services. Our approach sheds light on how to balance value segmentation among the key elements in the streaming music ecosystem and, consequently, warrants deeper study.

Last, we apply explanatory econometrics. Such modeling and estimation have uncovered useful insights by considering many possible factors to analyze an event. However, industry applications usually focus on a single perspective due to the complexity, data and modelling costs, and the difficulty of achieving causal explanations, among other reasons. This work demonstrates the combination of statistics and econometrics with traditional music recommendation methods from computer science. This kind of blended empirical data analytics can be extended for other state-of-the-art recommendation methods, such as matrix factorization.

7. Conclusion

We proposed a two-sided value-based artist promo-

tion method for recommending streaming music, and report some statistical gains compared to collaborative filtering. There are several limitations.

First, in the proposed two-sided value-based model, we did not consider the two sides' costs or the weight setting for each side when estimating total value. Consumer utility was measured using listening quantity, but we ignored the search cost for finding the music. It is hard to calculate the search cost for each listener because there are too many ways they can access music information before deciding to listen to it. For the provider's value, we borrowed the idea of "value in the presence of information less value in the absence of information": the *economic value of information* [21]. We think of this in terms of availability of recommendation information, or the lack of it, to calculate the potential value that the providers can gain from recommendation. This ignores the investment they need to make for music production and promotion. Although we mentioned that the platforms can price the promotion, determining how to measure and combine the two-sided costs entails additional investigation. Similarly, even though we set the same weights for the two sides, finer grained analysis is needed to tune them.

Second, the model is in its infancy. The estimated expectations of listening quantity and probability were used for recommendation, but our approach was not rigorous. Such an expectation may be insufficient to describe how much a potential listener really will listen. The amount of listening that has the largest probability of being observed may be better to explain audience listening and estimate its value. Thus, improving the estimation of future listening quantity is on our list.

Third, our dataset is limited, so memory-based music recommendation was selected as the base model on which to make improvements. This may not be workable for a very large dataset. Also, we did not consider the CF cold start problem. So, further exploration is required to understand: (1) how to build a scalable algorithm that can be easily implemented in industry environments; and (2) how to leverage the hybrid (content + CF) or the Bandit algorithm to address the cold start problem. Also (3) given that historical data were used to train and test the proposed model, we may need to do a randomized experiment or a user study to test the approach with unique new data.

We plan to collect more data, explore how to achieve scalability for the proposed method, and transfer the base model from memory-based KNN to model-based matrix factorization (MF) via the hybrid method. MF can help improve the estimation of a_{ij} by embedding listener and artist value-aware attributes into the model-based recommendation algorithm; and support exploring the two-sided costs.

8. References

- [1] D. Sanchez, "What Streaming Music Services Pay (Updated for 2017)", *Dig. Music News*, January 16, 2017.
- [2] D. Sanchez, "What Streaming Music Services Pay (Updated for 2018)", *Dig. Music News*, January 16, 2018.
- [3] P. Parisi, "Copyright Royalty Board Boosts Songwriters' Streaming Pay Nearly 50%", *Variety*, January 27, 2018.
- [4] Y. Koren, R. Bell, C. Volinsky, "Matrix Factorization Techniques for Rec. Sys.", *Comp. J.*, 42, 2009, pp. 42-49.
- [5] P. Cano, M. Koppenberger, N. Wack, "Content-based Music Audio Recommendation", in *Proc. 13th ACM Intl. Conf. Multimed.*, ACM Press, New York, 2005, pp. 211-212.
- [6] C. Desrosiers, G. Karypis, "A Comprehensive Survey of Neighborhood-Based Recommendation Methods", in *Rec. Sys. Handbook*, Springer, Boston, 2011, pp. 107-144.
- [7] P. Knees, M. Schedl, "A Survey of Music Similarity and Recommendation from Music Context Data", *ACM Trans. Multimed. Comp., Comm. Appl.* 10(1), 2013, pp. 1-21.
- [8] Z. Cheng, J. Shen, "On Effective Location-Aware Music Recommendation", *ACM Trans. Info. Sys.* 34(2)13, 2016.
- [9] IFPI (International Federation of the Phonographic Industry), *Global Music Report*, 2017, London, UK.
- [10] J. Salo, M. Lankinen, M. Mäntymäki, "Use of Social Media for Artist Mktg.: Music Industry Persp. and Consumer Motivations", *Intl. J. Med. Mgmt.* 15(1), 2013, pp. 23-41.
- [11] T. Belluf, L. Xavier, L. R. Giglio, "Case Study on Business Value Impact of Personalized Recommendations on a Large Online Retailer", in *Proc. 6th ACM Conf. Rec. Sys.*, ACM Press, New York, 2012, pp. 277-280.
- [12] G. Adomavicius, J. Bockstedt, S. Curley, J. Zhang, "Effects of Online Recommendations on Consumers' Willingness to Pay", *Info. Sys. Res.*, 29(1), 2017, pp. 84-102.
- [13] U. Panniello, S. Hill, M. Gorgoglione, "The Impact of Profit Incentives on Relevance of Online Recommendations", *Elec. Comm. Res. App.* 20, 2016, pp. 87-104.
- [14] S. John, Rational choice theory, in G. Browning, A. Halckli, F. Webster (eds.), *Understanding Contemporary Soc.: Theories of the Present*, Sage, Thousand Oaks, 2000.
- [15] H.R. Varian, *Intermediate Microeconomics: A Modern Approach* (8th ed.), W.W. Norton, New York, 2010.
- [16] C.A. Gomez-Urbe, N. Hunt, "The Netflix Recommender System: Algorithms, Business Value, and Innovation", *ACM Trans. Mgmt. Info. Sys.* 6(4)13, 2016.
- [17] N.A. Morgan, L.L. Rego, "Value of Different Customer Satisfaction and Loyalty Metrics in Predicting Business Performance", *Mktg. Sci.* 25(5), 2006, pp. 426-439.
- [18] R. Garg, M.D. Smith, R. Telang, "Measuring Information Diffusion in an Online Community", *J. Mgmt. Info. Sys.*, 28(2), 2011, pp.11-38.
- [19] Y. Zhang, Q. Zhao, Y. Zhang, D. Friedman, M. Zhang, Y. Liu, and S. Ma, "Economic Recommendation with Surplus maximization", in *Proc. 25th Intl. Conf. Comp. on WWW*, ACM Press, New York, 2016, pp. 73-83.
- [20] S. Tan, J. Bu, C. Chen, B. Xu, C. Wang, X. He, "Using Rich Social Media Information for Music Recommendation via Hypergraph Model", *ACM Trans. Multimed. Comp. Comm. Appl.*, 78(1)22, 2011.
- [21] J. Demski, *Information Analysis*, Addison-Wesley, Reading, MA, 1980.
- [22] R.J. Kauffman, K. Kim, S.Y.T. Lee, A. Hoang, J. Ren, "Combining Machine-based and Econometrics Methods for Policy Analytics Insights", *Elec. Comm. Res. Apps.*, 25, 2017, pp. 115-140.
- [23] M.J. Davern, R.J. Kauffman, "Discovering Potential and Realizing Value from IT Investments". *J. Mgmt. Info. Sys.*, 16(4), 2000, pp. 121-143.
- [24] J. Ren, R.J. Kauffman, "Understanding Streaming Music Diffusion in a Semi-Closed Social Environment," in *Proc. 22nd Pac. Asia Conf. Info. Sys.*, 2018, Assoc. for IS, Atlanta, pp. 206-219.
- [25] J. Ren, Z. Cheng, J. Shen, H. Miao, "Influences of Influential Users: An Empirical Study of Music Social Network", in *Proc. 6th Intl. Conf. on Inet. Multimed. Comp. Serv.*, ACM Press, New York, 2014, pp. 411.
- [26] A. Van den Oord, S. Dieleman, B. Schrauwen, "Deep Content-based Music Recommendation", in *Adv. Neur. Info. Proc. Sys.*, MIT Press, Cambridge, 2013, pp. 2643-2651.
- [27] H. Wang, N. Wang, D.Y. Yeung, "Collaborative Deep Learning for Recommender Systems", in *Proc. 21th ACM SIGKDD Intl. Conf. Knowl. Disc. Data Min.*, ACM Press, New York, 2015, pp. 1235-1244.