

New Recommendation Agent to Identify Innovators Utilizing User-Generated Content

Tae-Hyung Pyo
University of Idaho

Chanchal Tamrakar
University of North Texas

Sanjay Sisodiya
University of Idaho

Jae Young Lee
Yonsei University

Consumers who shop online are overwhelmed by the unlimited choices and options presented to them. Companies use different recommendation systems (agents), such as collaborative filtering that processes large volumes of data, to provide personalized and relevant recommendations for their customers to alleviate this problem. This paper explores a novel recommendation system that identifies consumers' unique unobserved preferences by utilizing user-generated content (UGC), such as social tags or bookmarks. We propose better metrics and algorithms that can be used as recommendation agents for a successful new product campaign; our proposed method outperforms current popular recommendation algorithms (RA).

Keywords: recommendation system, preference, user-generated content, social tag, e-commerce

INTRODUCTION

Consumers are increasingly turning to online platforms to make most of their purchases, as evidenced by e-commerce sales of \$1.11 Trillion in 2023, an increase of 7.6% from the previous year (U.S. Department of Commerce, Census Bureau, 2023). While e-commerce sales accounted for about 15.4% of total retail sales (U.S. Department of Commerce, Census Bureau, 2023), for some other product categories, such as music, online purchases accounted for as much as 84% of total revenue in 2021 (Wilson, 2022). As online consumer purchases have increased, so has the time consumers spend conducting online research. 50% of consumers spend more than 75% of their total shopping time conducting research online, while 15% spend more than 90% on online research (Vaughan, 2017). With an abundance of information sources, many choices, and multiple online options, companies are utilizing Recommendation Agents (RAs) to provide personalized suggestions to aid consumers in their decision-making. RAs are widely used across many

online platforms ranging from e-commerce (Amazon, eBay) to online playlist generators (YouTube, Netflix) (Wu, Sun, Zhang, Xie, & Cui, 2022). A trusted RA system provides value-added services to customers by reducing search costs (Pathak, Garfinkel, Gopal, Venkatesan, & Yin, 2010) while helping companies earn business; e.g., 60% of Netflix's streaming traffic and 35% of Amazon's sales come from recommendations (Hosanagar, Fleder, Lee, & Buja, 2014).

Recommendation Agents are software or algorithms that identify customers' preferences, interests, and behavior to generate personalized recommendations. RAs play an essential role in reducing information overload, search costs, and uncertainty, thereby aiding consumers' decision-making (Lu, Wu, Mao, Wang, & Zhang, 2015; Wu et al., 2022). Because of the exponential growth of online information sources, RAs are required to generate the best and most relevant recommendations from a seemingly infinite number of choices, in real-time, for a diverse set of customers, using applications such as E-commerce platforms, social networks, video or music sharing platforms (J. Chen et al., 2023; Jannach, Resnick, Tuzhilin, & Zanker, 2016; Liebman, Saar-Tsechansky, & Stone, 2014; Thorat, Goudar, & Barve, 2015; Wu et al., 2022).

Due to its importance and decades of research on recommendation systems, companies use various algorithms, data sources, and methods to refine their RAs (Ekstrand, Riedl, & Konstan, 2011).

One widely used RA approach is content-based filtering, which generates customized recommendations for individuals based on their previous choices. Consumer profiles, past purchases, clicks, ratings, etc., are important in content-based filtering (Thorat et al., 2015). RAs use their understanding of attributes and keywords to match items with consumers' preferences to generate recommendations.

Another widely used RA approach is collaborative filtering, which generates personalized recommendations based on the choices of other users with similar preferences. The RAs recommend products that other users with similar preferences have already tried or liked (Z. Chen & Wang, 2022; Hosanagar et al., 2014; Lu et al., 2015; Zhou, Leng, Liu, Cui, & Yu, 2022). Amazon.com's use of "Customers who bought this also bought" is a good example of collaborative filtering.

Another widely used RA approach is hybrid recommendation systems, which combine collaborative filtering, content-based filtering, and other methods to enhance the accuracy of recommendations. Various permutations of different RA systems are combined to form a hybrid RA. For example, content-based and collaborative filtering recommendations could be generated separately and then combined, or adding some content-based capabilities to collaborative RA, or combining different RAs to provide a recommendation. Netflix is a good example of a hybrid RA.

Despite their effectiveness, most RAs face significant challenges are associated with data sparsity, scalability, and cold start problems. RAs may face a cold start problem while recommending solutions for new users, new products, and items with limited data. The user-item matrix used by RAs can become very large and sparse, leading to the degradation of the recommendation process as well. Also, with an increase in data availability, the computational complexity of RAs can lead to scalability issues (Najafabadi, Mohamed, & Onn, 2019; Natarajan, Vairavasundaram, Natarajan, & Gandomi, 2020; Thorat et al., 2015). Matrix completion techniques have been introduced as a complementary approach to collaborative filtering to address many of these issues (data sparsity, missing entries, etc.) (Z. Chen & Wang, 2022; Jannach et al., 2016).

In this study, we propose an alternative novel approach to accurately quantify customer preferences by utilizing social tags or social bookmarks, which are user-generated content (UGC) and are widely adopted in online stores and digital/social platforms (Zheng & Li, 2011). We believe that identifying each customer's unique preferences is key to successful RA development.

Social tags can be easily observed and are a popular categorization method used by users of online websites such as Instagram product tags, X hashtags, Pinterest hashtags, Tumblr tags, etc. These UGC elements can then be aggregated to extract meaningful characteristics of those online content, e.g., the genre of songs defined by its users. Unlike professionals or website managers, online users create social tags (keywords) to classify or categorize content such as music and movies (Nam, Joshi, & Kannan, 2017). These user-generated tags are an expression of the user's preference (Carlson, Saari, Burger, & Toiviainen,

2017; Zheng & Li, 2011) and should provide valuable information for recommendation agents (Agrawal, Roy, & Mitra, 2021; Dong, Li, Ma, & Liu, 2022).

This research proposes an intuitive method of concisely inferring unobserved (latent) preference by leveraging rich information available in social tags that are already widely available on various digital media platforms. In addition, we propose a better metric and algorithm for RAs using UGC and empirically show that our proposed approach outperforms other popular RA algorithms that rely on professionally generated content (PGC).

DATA, METRICS, AND EMPIRICAL TESTING

As of December 2011, there were 76,249 communities on Last.fm, but most were dormant. Since active communities tend to have more users, we randomly sampled 300 communities from 1,413 communities with more than 1,000 members. In the sampled communities, there were 25,705 users who played one song per day on average from 2011 to 2013 and lived in the United States. We randomly sampled 1,000 users from these 25,705 users.

We construct novel metrics to identify users' preferences using the rich social tag data and individual users' consumption behavior (music library) on Last.fm. Using Aitchison distance, we further calibrate the fit between users' preferences and the product (artist of the song). We compute the fit of newly released songs with these users' preferences. We identified their preferences using social tags on the songs that they listened to. We also quantified the characteristics of the artists by extracting social tags on songs they had previously released. By calculating the distance between each user's preferences and the artist's characteristics, we estimate the degree of fit between a user and an artist. This approach allows us to make the best recommendations for the users. We empirically demonstrate that our proposed metric outperforms the other two alternative metrics.

To compare our new RA algorithm, we consider two alternative approaches to building the metric of preference: (1) collaborative filtering (CF), which is the most widely adopted algorithm for RA (Gong, 2009; Herlocker, Konstan, Terveen, & Riedl, 2004; Ricci, Rokach, & Shapira, 2011) and (2) identifying users' preferences, e.g., tempo, danceability, and speechiness, through professionally generated content (PGC) of songs from Spotify.

Identifying Preferences and Calculating Fit

Social tags (or bookmarking) are user-generated content that aggregates product categorization criteria created and used by users. Popular social tags on products provide insights into how customers classify multiple products in their minds. We also believe that social tags created by customers provide useful insights into their preferences, and we empirically test our approach by identifying and quantifying individual preferences from extensive behavioral data using popular social tags.

We measure users' music preferences by analyzing their play histories in 2011 (i.e., prior to the test period). Since preference is a precursor of product choice and consumption, we use an observed consumption behavior (i.e., music play history) as a proxy to measure preferences. This approach is consistent with the way that users of Last.fm infer their peers' preferences.

We use social tags from the users' play history to quantify user preferences. Social tags are user-generated content that classifies products (Nam et al., 2017). Given the popularity of social tag usage in digital media, extant research has used social tags to better understand brand perception (Culotta & Cutler, 2016; Nam et al., 2017; Ringel & Skiera, 2016). However, this study uses social tags in a reverse but novel way. We use social tags not to understand a brand, but to understand consumers. The frequency of social tags in a user's playlist contains information about their preferences, which means that social tags of songs that the user likes will appear more often in their play history.

We can access users' complete music play histories on their personal web pages on Last.fm (Figure 1). Also, the top tags on an artist (Figure 2A) or songs (Figure 2B) are available on their web pages.

FIGURE 1
EXAMPLE OF A USER'S PLAY HISTORY DATA FROM A PERSONAL WEB PAGE

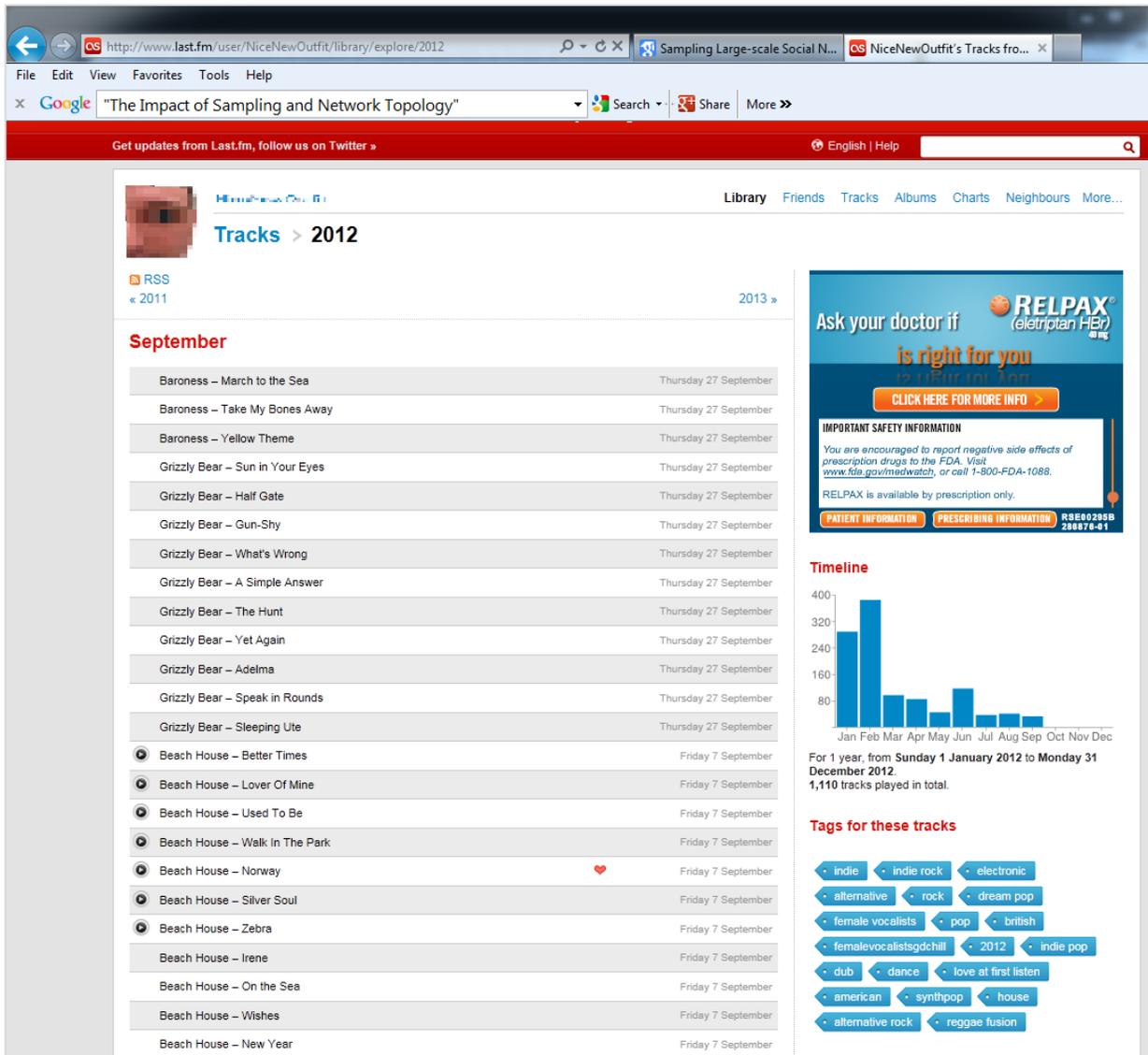


FIGURE 2
A. SOCIAL TAGS ON AN ARTIST

B. SOCIAL TAGS ON A SONG

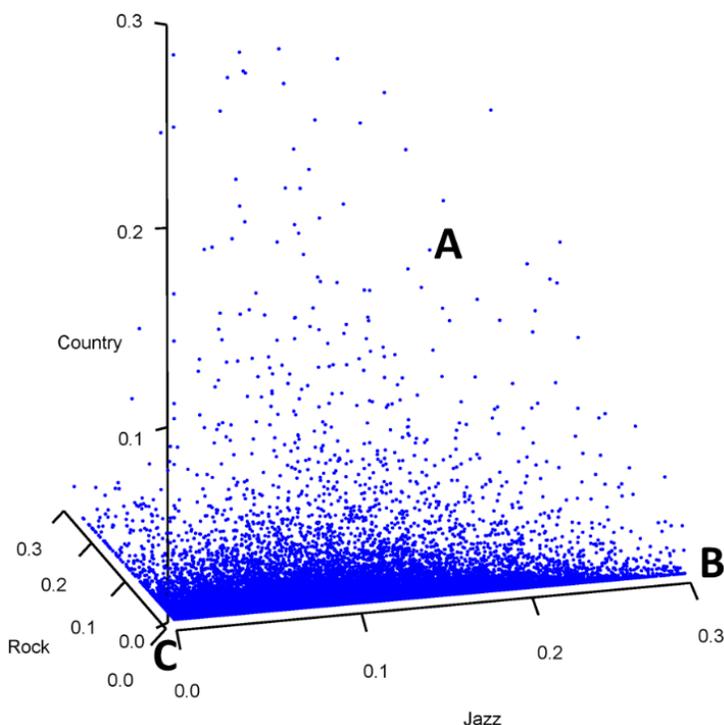
Since Last.fm provides rich and detailed information about users and artists, many studies have used its data to investigate individual music consumption behavior (Bapna & Umyarov, 2015; Datta, Knox, & Bronnenberg, 2017; Garg, Smith, & Telang, 2011). Consistent with past studies that examined music preference by genre (Greenberg, Baron-Cohen, Stillwell, Kosinski, & Rentfrow, 2015; Rentfrow & Gosling, 2003), we focus on 19 popular social tags on genres found on www.last.fm/music such as “electronic,” “rock,” and “hip hop.”

Following the metrics proposed by Pyo, Lee, and Park (2022), we identified a user’s preference from their playlist. First, we computed the proportion of genre tag g in the user i ’s past playlist as follows:

$$c_{ig} = \frac{freq_{ig}+1}{\sum_{g \in \{1,2,..G\}}(freq_{ig}+1)} \quad (1)$$

where $freq_{ig}$ denotes the frequency that genre tag g appears across all songs that user i played in the given time period and C_i is a vector with 19 genre entries. The high value of $freq_{ig}$ means a user listened to genre g more frequently, i.e., g is their preferred genre. We call C_i the preference coordinates for user i . Figure 3 (we call it Preference Map) visualizes C_i of users in our data along with three genre tags (i.e., rock, jazz, country). Our data shows that musical taste looks reasonably well dispersed. For instance, user A has a balanced preference across rock, jazz, and country, but user B's preference is highly skewed towards jazz. User C has almost zero on all three coordinates, so they majorly listen to music other than rock, jazz, and country.

**FIGURE 3
PREFERENCE MAP**



Note: C_i on all 3 axes are truncated at 0.3 for display purpose.

Using tags from songs previously released by artist k , we calculate coordinates for the artist's music style across the same 19 genres (A_k) used to construct preference coordinates of users (C_i). Intuitively, artist k 's music styles fit user i more (less) if A_k and C_i are closer (farther) to each other in the 19-dimensional preference space. To compute the distance between the artist and each user, we use Aitchison Distance (Aitchison, 1982), which is known to suit compositional data the best, where the sum of all coordinates equals 1 (Martín-Fernández et al., 1998; Otero, Tolosana-Delgado, Soler, Pawlowsky-Glahn, & Canals, 2005). Fit between user i and artist k is defined as follows (Pyo et al., 2022):

$$fit_{ik} = \sqrt{\sum_g \left(\ln \left(\frac{c_{ig}}{g(C_i)} \right) - \ln \left(\frac{a_{kg}}{g(A_k)} \right) \right)^2} \quad (2)$$

where geometric means that $g(C_i) = \sqrt[19]{c_{i1}c_{i2}\dots c_{iG}}$ and $g(A_k) = \sqrt[19]{a_{k1}a_{k2}\dots a_{kG}}$.

Measuring Preference Fit by CF

From 1,000 focal users' and the other 40,233 users' play history data in 2011, we measure the level of fit of the artist k 's song to the user i (fit_CF_{ik}) using the algorithm of collaborative filtering (Adomavicius & Tuzhilin, 2005; Ahn, 2008; Ricci et al., 2011).

We first define I_{jk} as an indicator that takes the value 1 if the artist k 's song appears in user j 's music library in 2011 and 0 otherwise. Then, we calculate the similarity between artist k_1 and k_2 ($sim_{k_1k_2}$) as follows (Pyo et al., 2022):

$$sim_{k_1k_2} = \frac{\sum_{j=1}^J I_{jk_1} I_{jk_2}}{\sqrt{\sum_{j=1}^J I_{jk_1}^2} \sqrt{\sum_{j=1}^J I_{jk_2}^2}}, \text{ where } J \text{ equals } 41,233. \quad (3)$$

For the matrix operation of equation 3, we create matrix Q , in which $[j,k]$ entry will be 0 if user j has not listened to artist k 's music and 1 otherwise. This matrix Q has 41,233 rows (users) and 247 columns (artists). If Q_j is set to be the j -th row of matrix Q , then the vector's length will be 247 and will contain information about the 247 artists that user j has listened to. Finally, if Q_k is set to be the k -th column of matrix Q , then the vector has information about those users that have listened to artist k 's music, and its length will be 41,233.

Consequently, we constructed Cos as a matrix that summarizes the similarity of listenership across the 247 artists. Cos is a symmetric matrix of the order 247×247 , where the $[k_1, k_2]$ entry is the cosine similarity between Q_{k_1} and Q_{k_2} , and takes a larger value when the overlap between listeners of artists k_1 and k_2 is bigger. Since the Cosine similarity between identical vectors is 1, the diagonals of Cos are also 1. If Cos_k is the k -th column of the matrix Cos , then the vector will contain information about how the artist k 's listeners are similar to the listeners of the other 246 artists.

Next, we create a metric that captures how likely it is that artist k 's music is recommended to the user i (fit_CF_{ik}). We operationalize this fit as the inner product between Cos_k and Q_i divided by the sum of vector elements of Q_i . If user i has listened to the songs of other artists whose listeners have also listened to many songs of artist k then this inner product (i.e., fit_CF_{ik}) should be a high number.

Measuring Fit by Alternative Source (PGC)

We construct preference similarity analogously to our approach using PGC as an alternative approach. Although such information is often unavailable in categories such as movies, books, or food, we found audio feature data on music tracks consisting of 9 dimensions (e.g., tempo, danceability, and speechiness) from Spotify Web API. Using 9-dimensional data from Spotify, we calculate fit_PGC_{ik} . The procedures are the same as fit_PGC_{ik} except that nine-dimensional audio feature data is used instead of 19 genre-based dimensions in our tag data.

Comparison of Hits From Three Approaches

As shown in Figure 4, the distribution of the proposed measures for the fit between song k and individual i (fit_{ik}) is close to normal, while others have many measures from CF (fit_CF_{ik}) returned to 0 or the distribution of measures from PGC (fit_PGC_{ik}) are more skewed to the right. The correlation matrix in Table 1 also indicates that the correlation between fit_{ik} and fit_CF_{ik} is minimal (-.04), which suggests that the proposed metrics capture different preference structures among individuals from those by CF.

FIGURE 4
HISTOGRAMS OF FITS BY THREE METRICS

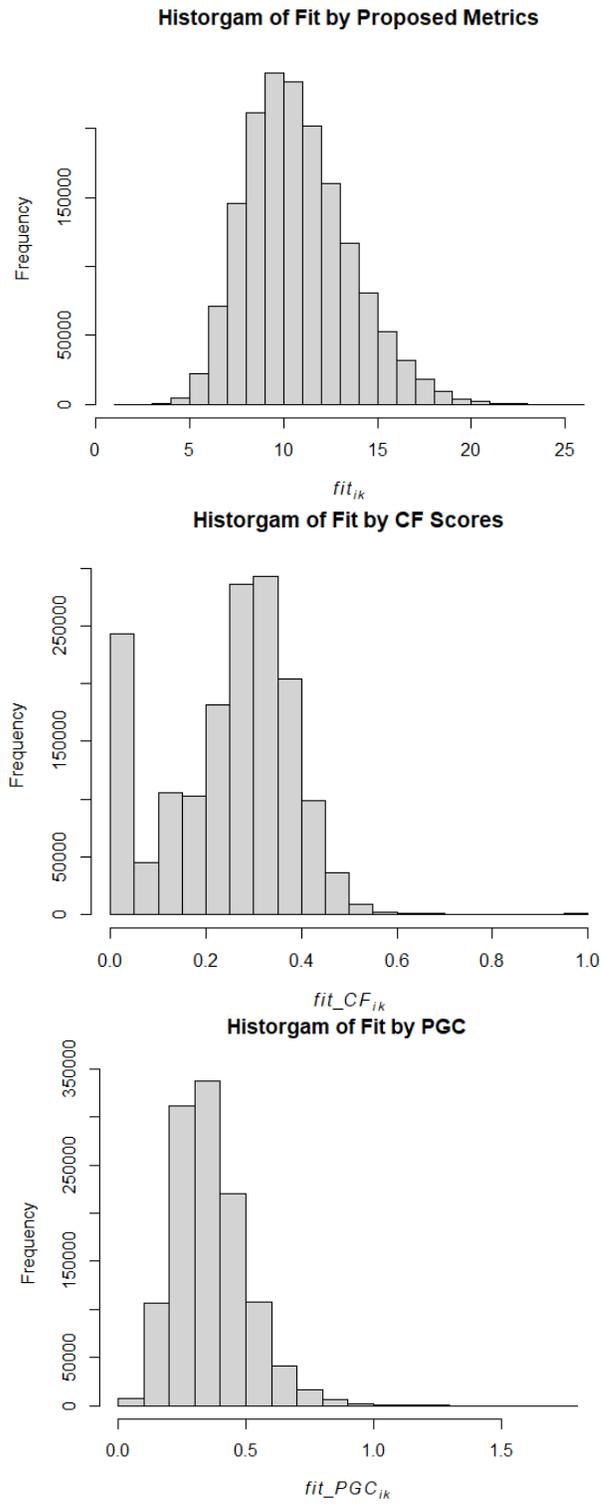


TABLE 1
DESCRIPTIVE STATISTICS

	Mean	SD	Median	Min	Max	Correlation Matrix		
						<i>fit (Proposed)</i>	<i>fit_CF</i>	<i>fit_PGC</i>
<i>User Stats</i>	3.13	3.93	2	0	21			
<i>Song Stats</i>	9.78	5.56	10.4	1.9	25.8			
<i>fit (Proposed)</i>	10.74	2.71	10.45	1.08	25.69	1		
<i>fit_CF</i>	0.25	0.13	0.27	0	1	-.04	1	
<i>fit_PGC</i>	0.36	0.14	0.34	0.03	1.77	.19	-.28	1

We compare the performance of recommendations across the three approaches by analyzing 1,000 randomly selected users' music-listening behavior for newly released songs. We selected the same 32 songs as in Pyo et al. (2022). These 32 songs were released between April and October 2012 and were amongst the top 50 most listened songs on Last.fm as of August 2013. Users are selected to recommend these 32 songs using three approaches, and we define an adoption if a user plays the song at least once during our data collection periods (2012 and 2013).

For the proposed approach (fit_{ik}) and PGC (fit_PGC_{ik}), we ordered the 1,000 selected users from the smallest to the largest value. The user i with the smallest fit value is selected for the given number of users to recommend song k , while the users with the highest Sim_{ik} are selected for CF.

It is well established that targeting innovators, i.e., the customer who adopts the new product early in its lifecycle, is critical to the success of new product campaigns (Mahajan & Muller, 1998; Robertson, 1967). Therefore, we empirically tested our approach by evaluating how many actual adopters it correctly identified during the early weeks of the new song release compared to the two other approaches. Table 2 presents a comparison of hit counts among 1,000 users for the 32 songs during the first 12 weeks after the song's release; the hit count is calculated as the number of actual adopters amongst the selected users to recommend the song k (hit) using the three approaches to recommend the song. Each recommendation approach selected the same number of users (out of 1,000 users) for recommendation, and the number of users to recommend the song k equals the number of actual adopters of the song k among 1,000 selected users. For example, for the song "I'll Be Alright," the three recommendation approaches chose 107 users to recommend this song. There are 1,362 adoptions for 32 selected songs from 1,000 users. The total hit is 210 for the proposed approach, 175 for CF, and 138 for PGC. Our proposed RA using social tags achieved 20% and 52% higher hit ratios than CF and PGC, respectively. Comparison using hit analysis demonstrates that our proposed algorithm performs best in every aspect. The first 8(4) weeks of hits are, 102(34) for the proposed approach, 92(31) for CF, and 69(26) for PGC from total adoptions of 891(425). Even for the entire data collection period, the proposed approach outperformed CF and PGC, with 770 actual adoptions identified compared to 697 for CF and 520 for PGC.

TABLE 2
COMPARISON OF THREE APPROACHES FOR HIT RATIO IN THE FIRST 12 WEEKS

Song Title	Hit Count			Actual
	Proposed	CF	PGC	
I'll Be Alright	24	12	6	107
Take A Walk	25	11	6	105
Madness	10	8	5	87
Ride	11	10	15	86
Die Young	23	25	0	81
Lover of the Light	19	6	19	80
Anything Could Happen	12	4	14	68
All Your Gold	8	9	5	68
We Are Never Ever Getting Back Together	16	15	16	67
Runaways	7	5	4	63
Sleep Alone	8	7	1	52
Octopus	3	2	1	49
Locked Out of Heaven	6	7	10	47
One More Night	5	4	8	40
Something Good	2	4	0	40
Your Body	10	14	2	40
Wildest Moments	5	5	0	36
Live While We're Young	4	6	0	32
As Long as You Love Me	2	2	4	25
Laura	1	1	1	25
Pound the Alarm	0	4	7	23
I Cry	0	2	3	21
Crew Love	2	1	3	21
Beauty and a Beat	1	2	2	18
Blow Me (One Last Kiss)	2	1	4	18
Goldie	1	3	0	16
Home	0	2	0	13
Whistle	2	0	1	12
Don't Wake Me Up	1	1	0	8
Girl on Fire	0	1	1	7
Don't You Worry Child	0	1	0	4
This Is Love	0	0	0	3
Total Adopter	210	175	138	1362

Bold: highest hit count

CONCLUSION

This research proposes a method on how to quantify preferences by utilizing user-generated social tags. An approach can then be used to make customized recommendations appropriately tailored to each user's

preferences. Since the evaluation of products such as experience goods is highly subjective, marketers must measure and understand customer preferences well. This research proposes an intuitive approach of concisely recovering unobserved preferences using rich information in social tags, which are widely available across various digital media platforms.

We empirically demonstrate that our proposed method outperforms the most common RAs used in e-commerce, particularly in identifying innovators—users who play critical roles in the success of new products. We also note that CF cannot determine an individual's preference, i.e., whether they like pop, rock, or hip-hop, because it simply counts how many customers tried the product together to measure similarity among products. Therefore, its application, besides a recommendation algorithm, is minimal.

Furthermore, CF cannot generate a recommendation for a new artist because it cannot calculate similarity with other artists. CF requires many users to play the new artists' songs to measure the similarity between existing artists and new artists. Therefore, the collaborative filtering approach is limited when the observed consumption is rare (Ghazarian & Nematbakhsh, 2015; Koochi & Kiani, 2020; Sachan & Richhariya, 2013; Yildirim & Krishnamoorthy, 2008). Also, for PGC, the recommendation system must wait until information for the new artist is available from professionals. In contrast to all of these approaches, our proposed approach can quickly identify similarities as long as users who tried the new artist's songs assign tags to them.

Customer preferences are generally identified through surveys or experiments, especially using conjoint analyses such as multivariate regression or discrete choice models. Also, measures of revealed preference (or stated preference) require a pre-determined set of product attributes generally decided upon by a researcher (not customers). Poorly designed experiments could lead to the misidentification of customer preferences. How to define attributes in high-dimensional products such as entertainment or fashion products is also not clearly suggested in current literature. Additionally, findings from experimental settings often lack external validity. In this study, we developed a new metric to identify customers' preferences by utilizing big data on actual customers' past behaviors to address these issues.

Although it is a very popular practice on multiple digital media platforms, social tags have been rarely used in the marketing literature except to identify brand/product attributes (Culotta & Cutler, 2016; Nam et al., 2017; Ringel & Skiera, 2016). Utilizing social tags, which are widely used in digital media, we propose a new method to identify key product attributes (popular social tags) as they exist in customers' minds and use them to quantify individual customer preferences. Our preference metrics also allow us to compare customer preferences and map each customer in a product attributes space, which we named the preference map.

We argue that our preference map - Figure 3 - would serve as a valuable tool for market segmentation since it shows the relative location of customers in terms of their preferences across multiple product attributes. Marketers can use this Preference Map to identify and target customers based on varying degrees of preference for a certain genre, i.e., rock music fans or fans with preferences across multiple genres (e.g., pop and hip-hop). We believe practitioners can easily apply our proposed method to better understand, analyze, and visualize customer preference structures.

Although our study's context is music, our approach should be highly relevant to other product categories where social tag data is widely available. Future research exploring other product categories, such as movies, books, etc., could benefit from adopting our approach.

REFERENCES

- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734–749.
- Agrawal, S., Roy, D., & Mitra, M. (2021). Tag embedding based personalized point of interest recommendation system. *Information Processing & Management*, 58(6), 102690.
- Ahn, H.J. (2008). A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem. *Information Sciences*, 178(1), 37–51.
- Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(2), 139–177.
- Bapna, R., & Umyarov, A. (2015). Do your online friends make you pay? A randomized field experiment on peer influence in online social networks. *Management Science*, 61(8), 1902–1920.
- Carlson, E., Saari, P., Burger, B., & Toiviainen, P. (2017). Personality and musical preference using social-tagging in excerpt-selection. *Psychomusicology: Music, Mind, and Brain*, 27(3), 203.
- Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., & He, X. (2023). Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems*, 41(3), 1–39.
- Chen, Z., & Wang, S. (2022). A review on matrix completion for recommender systems. *Knowledge and Information Systems*, 64(1), 1–34.
- Culotta, A., & Cutler, J. (2016). Mining brand perceptions from twitter social networks. *Marketing Science*, 35(3), 343–362.
- Datta, H., Knox, G., & Bronnenberg, B.J. (2017). Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. *Marketing Science*, 37(1), 5–21.
- Dong, J., Li, G., Ma, W., & Liu, J. (2022). Personalized recommendation system based on social tags in the era of Internet of Things. *Journal of Intelligent Systems*, 31(1), 681–689.
- Ekstrand, M.D., Riedl, J.T., & Konstan, J.A. (2011). Collaborative filtering recommender systems. *Foundations and Trends® in Human-Computer Interaction*, 4(2), 81–173.
- Garg, R., Smith, M.D., & Telang, R. (2011). Measuring information diffusion in an online community. *Journal of Management Information Systems*, 28(2), 11–38.
- Ghazarian, S., & Nematbakhsh, M.A. (2015). Enhancing memory-based collaborative filtering for group recommender systems. *Expert Systems with Applications*, 42(7), 3801–3812.
- Gong, S. (2009). Employing User Attribute and Item Attribute to Enhance the Collaborative Filtering Recommendation. *JSW*, 4(8), 883–890.
- Greenberg, D.M., Baron-Cohen, S., Stillwell, D.J., Kosinski, M., & Rentfrow, P.J. (2015). Musical preferences are linked to cognitive styles. *PLoS One*, 10(7), e0131151.
- Herlocker, J.L., Konstan, J.A., Terveen, L.G., & Riedl, J.T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1), 5–53.
- Hosanagar, K., Fleder, D., Lee, D., & Buja, A. (2014). Will the global village fracture into tribes? Recommender systems and their effects on consumer fragmentation. *Management Science*, 60(4), 805–823.
- Jannach, D., Resnick, P., Tuzhilin, A., & Zanker, M. (2016). Recommender systems—beyond matrix completion. *Communications of the ACM*, 59(11), 94–102.
- Koohi, H., & Kiani, K. (2020). Two new collaborative filtering approaches to solve the sparsity problem. *Cluster Computing*, 1–13.
- Liebman, E., Saar-Tsechansky, M., & Stone, P. (2014). Dj-mc: A reinforcement-learning agent for music playlist recommendation. *arXiv preprint arXiv:1401.1880*.
- Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: A survey. *Decision Support Systems*, 74, 12–32.
- Mahajan, V., & Muller, E. (1998). When is it worthwhile targeting the majority instead of the innovators in a new product launch? *Journal of Marketing Research*, 35(4), 488–495.

- Martín-Fernández, J., Barceló-Vidal, C., Pawlowsky-Glahn, V., Buccianti, A., Nardi, G., & Potenza, R. (1998). Measures of difference for compositional data and hierarchical clustering methods. Paper presented at the *Proceedings of IAMG*.
- Najafabadi, M.K., Mohamed, A., & Onn, C.W. (2019). An impact of time and item influencer in collaborative filtering recommendations using graph-based model. *Information Processing & Management*, 56(3), 526–540.
- Nam, H., Joshi, Y.V., & Kannan, P. (2017). Harvesting brand information from social tags. *Journal of Marketing*, 81(4), 88–108.
- Natarajan, S., Vairavasundaram, S., Natarajan, S., & Gandomi, A.H. (2020). Resolving data sparsity and cold start problem in collaborative filtering recommender system using linked open data. *Expert Systems with Applications*, 149, 113248.
- Otero, N., Tolosana-Delgado, R., Soler, A., Pawlowsky-Glahn, V., & Canals, A. (2005). Relative vs. absolute statistical analysis of compositions: A comparative study of surface waters of a Mediterranean river. *Water Research*, 39(7), 1404–1414.
- Pathak, B., Garfinkel, R., Gopal, R.D., Venkatesan, R., & Yin, F. (2010). Empirical analysis of the impact of recommender systems on sales. *Journal of Management Information Systems*, 27(2), 159–188.
- Pyo, T.-H., Lee, J.Y., & Park, H.M. (2022). The Effects of Consumer Preference and Peer Influence on Trial of an Experience Good. *Journal of Marketing Research*, 59(6), 1161–1178.
- Rentfrow, P.J., & Gosling, S.D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84(6), 1236.
- Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. *Recommender Systems Handbook* (pp. 1–35). Springer.
- Ringel, D.M., & Skiera, B. (2016). Visualizing asymmetric competition among more than 1,000 products using big search data. *Marketing Science*, 35(3), 511–534.
- Robertson, T.S. (1967). Consumer innovators: the key to new product success. *California Management Review*, 10(2), 23–30.
- Sachan, A., & Richhariya, V. (2013). Reduction of data sparsity in collaborative filtering based on fuzzy inference rules. *International Journal of Advanced Computer Research*, 3(2), 101.
- Thorat, P.B., Goudar, R.M., & Barve, S. (2015). Survey on collaborative filtering, content-based filtering and hybrid recommendation system. *International Journal of Computer Applications*, 110(4), 31–36.
- Vaughan, P. (2017). *Half of Your Customers Spend 75% of Time Conducting Research Online*.
- Wilson, J. (2022). The age of digital; music executive reacts to the impact of digitalization in the music industry.
- Wu, S., Sun, F., Zhang, W., Xie, X., & Cui, B. (2022). Graph neural networks in recommender systems: A survey. *ACM Computing Surveys*, 55(5), 1–37.
- Yildirim, H., & Krishnamoorthy, M.S. (2008). A random walk method for alleviating the sparsity problem in collaborative filtering. Paper presented at the *Proceedings of the 2008 ACM Conference on Recommender Systems*.
- Zheng, N., & Li, Q. (2011). A recommender system based on tag and time information for social tagging systems. *Expert Systems with Applications*, 38(4), 4575–4587.
- Zhou, C., Leng, M., Liu, Z., Cui, X., & Yu, J. (2022). The impact of recommender systems and pricing strategies on brand competition and consumer search. *Electronic Commerce Research and Applications*, 53, 101144.