Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Dance Together?: Characterizing the Artist-to-Artist Collaborations in K-Pop Challenges on YouTube Shorts

Andrea Chung

chungan@iu.edu

Indiana University

Jisun An

Indiana University

Haewoon Kwak

Indiana University

Research Article

Keywords:

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11405852/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Dance Together?: Characterizing the Artist-to-Artist Collaborations in K-Pop Challenges on YouTube Shorts

Andrea Chung (Corresponding Author), Jisun An, Haewoon Kwak

Luddy School of Informatics, Computing, and Engineering, Indiana University Bloomington, 700 N Woodlawn Avenue, Bloomington, 47408, IN, USA.

Contributing authors: chungan@iu.edu; jisunan@iu.edu; hwkwak@iu.edu;

Abstract

Artist-to-artist challenge collaborations in K-pop have emerged as a prominent promotional strategy where original artists invite guest performers to showcase choreographed highlights of their songs on official channels. Despite their widespread adoption, their dynamics and effectiveness remain underexplored. We address this gap through a data-driven analysis of 6,008 collaboration Shorts spanning 597 songs and 207 artists, integrating YouTube popularity metrics, artist attributes, and daily chart performance on an external music streaming service. Our findings show that these collaborations are highly reciprocal, reflecting a norm of mutual promotion. Partner selection also differs between first-time and recurring collaborations. First-time collaborations are most strongly associated with overlapping promotional schedules and social proximity, whereas recurring collaborations exhibit a more distributed pattern involving guest artist popularity, social proximity, and guest's musical characteristics. However, Granger causality analysis provides limited evidence that collaboration activity precedes improvements in streaming-chart performance. Together, these results suggest that artist-to-artist challenge collaborations are more closely aligned with existing promotional and commercial dynamics than with systematically generating subsequent chart gains, highlighting their strategic role within broader promotional ecosystems.

Keywords: Short-form video, artist-to-artist challenge collaborations, Network analysis, Granger causality

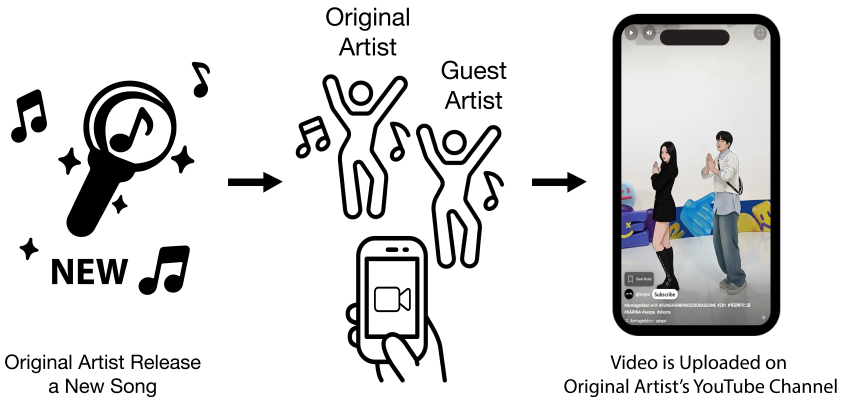


Fig. 1: Overview of a typical artist-to-artist challenge collaboration in K-pop. An original artist invites a guest artist to perform a choreographed excerpt of their new song together. The original challenge screenshot on the far right was replaced with an AI-generated illustration using Nanobanana [2025] to avoid reproducing copyrighted imagery.

1 Introduction

Digital platforms have reshaped the music industry, redefining audience expectations towards visual and interactive experiences [1, 2]. At the forefront of this shift is Korean pop music (K-pop), which has established itself as a global cultural phenomenon. In 2023, 19 out of 20 best-selling albums worldwide were produced by K-pop artists [3]. With a strong focus on choreography, synchronized visuals, and fan-centered promotional culture, K-pop is particularly well-suited for visual-first platforms like YouTube Shorts [1, 4]. Korean entertainment agencies were early adopters of social media for distribution and promotion, enabling artists to reach global audiences through participatory and interactive content strategies [5–7].

Among these strategies, one of the most prominent is the artist-to-artist challenge collaboration, a short-form video in which an artist performs choreographed highlights of a new song with an invited fellow artist (Figure 1). Throughout this paper, we refer to the artist who released the song and initiated the challenge as the *original artist*, and the invited performer as the *guest artist*. The resulting video is uploaded to the original artist's official YouTube channel, where it can reach the fan bases of both artists. Such collaborations have thus emerged as a platform-native promotional strategy that leverages choreography and cross-artist collaboration to maximize visibility [8]. While the broader term 'K-pop challenge' generally includes fan-generated dance covers, we focus exclusively on collaborations between professional artists posted on official channels and refer to this subset as *artist-to-artist challenge collaborations* throughout the paper.

In K-pop, promotional activities for a new release are concentrated in a short and intensive period. Arranging a joint performance within this window is therefore

unlikely to rest only on personal friendship between artists. Instead, it involves deliberate industry networking and strategic coordination between entertainment agencies seeking to maximize promotional impact. Empirical work on collaborations more broadly documents their role in amplifying visibility and improving commercial outcomes [10], and Silva et al. [11] show that incorporating collaboration features improves hit-song prediction. Nevertheless, to the best our knowledge, no quantitative study has isolated artist-to-artist challenge collaborations to answer basic questions about how they have evolved and are structured, how guest artists are selected, and how they relate to streaming chart performance.

In this work, we formulate the following research questions:

- RQ1. What are temporal trends, popularity distribution, and network structure of artist-to-artist challenge collaborations?
- RQ2. What factors shape the selection of guest artists, and how do they differ between first-time and recurring collaborations?
- RQ3. To what extent are these collaborations temporally associated with a song's streaming chart performance?

To address these questions, we conduct a data-driven analysis of 6,008 professional collaboration YouTube Shorts spanning 597 songs and 207 artists, integrating online engagement, artist attributes, and daily chart performance.

Our findings reveal that artist-to-artist challenge collaborations form highly reciprocal networks, with distinct patterns of partner selection between first-time and recurring collaborations. Although collaboration Shorts are associated with higher platform engagement than non-collaboration Shorts, we find limited evidence that collaboration activity consistently precedes changes in streaming-chart performance. These findings highlight how professional relationships and promotional opportunities are intertwined in short-form music promotion, while distinguishing platform visibility from commercial outcomes. By examining collaboration structure, partner selection, and promotional outcomes together, this study extends existing research on music collaboration beyond recorded works to interactions mediated by short-form platforms. As similar strategies increasingly appear in global music markets [12], our findings contribute to a broader understanding of the interplay between professional networks and digital promotion.

2 Literature Review

Social media platforms have long shaped cultural consumption, including how music is discovered, shared, and disseminated. Early studies show that social media data can reveal users' musical tastes and evolving patterns of media use [13]. YouTube, in particular, has been identified as a central site of music consumption and cultural production, amplifying mainstream popularity [14]. Content on such platforms often spreads through cascade-like dynamics, with a relatively small subset of videos receiving disproportionate attention [15]. Building on these insights, the rise of short-form platforms such as YouTube Shorts represents the next stage in the "platformification"

of music consumption through bite-sized, algorithmically distributed content [16]. These platform transformations also reshape how audiences experience music. The convenience and personalization of streaming services increased the role of algorithmically curated content in music discovery and consumption [17, 18]. Short-form platforms extend this logic by emphasizing rapidly consumable and continuously recommended content. As a result, growing prominence of such content may alter patterns of engagement with long-form content, reflecting broader changes in how platforms organize and prioritize user attention [19, 20].

Beyond shaping audience consumption, short-form platforms also altered how artists and the music industry approach promotion. Musicians actively use social media to expand their reach and engagement, while emerging artists can leverage viral trends and platform visibility to build their presence [21]. In the K-pop industry, social media serves as a core distribution channel [22]. Short-form dance challenges integrate choreography and visuals into song promotion, with artists and fans contributing to the circulation of recognizable song segments across platforms [8, 23]. Despite their growing prevalence, the ways short-form videos function as systematic promotional strategies among professional artists remain underexplored.

Collaboration in the music industry is shaped by a range of factors, including creative alignment, social relationships, and industry positioning. Musical proximity, especially genre and stylistic similarity, plays a key role in shaping collaboration patterns. Ortega [24] analyzes cover versions to show influence and visibility in music often follow genre-specific trajectories, with artists more likely to collaborate within stylistically similar genres. More recent research similarly finds that stylistic similarity is associated with the likelihood of collaboration between musicians [25].

Collaboration can also reflect social and strategic considerations beyond stylistic alignment. In digital environments, occasional collaborations often serve strategic or popularity-driven goals, while ongoing partnerships reflect relational history and shared creative norms [26]. Artist status can further shape access to collaboration opportunities. In the K-pop industry, for example, Lee and Gargiulo [27] show that status differences constrain the formation of new collaboration ties among freelance songwriters, with less successful songwriters facing barriers to forming ties with commercially successful peers. Beyond stylistic alignment, collaboration also functions as a strategic tool to expand audience reach and visibility [10, 28, 29]. Silva et al. [11] show that incorporating collaboration features improves hit-song prediction, highlighting the commercial significance of artist interactions.

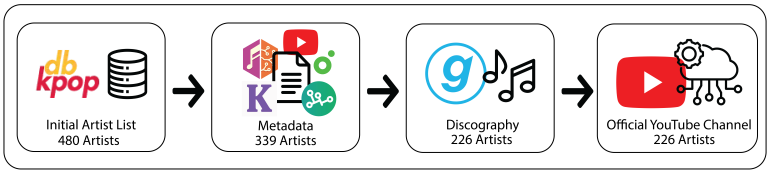
Most existing work, however, has examined formal collaborations embedded in recorded music, such as featuring relationships, co-releases, or other track-based partnerships among professional artists [11, 30, 31]. Artist-to-artist challenge collaborations differ from these arrangements in several ways. They are typically short-lived, tied to the promotion of a particular release, visually oriented, and distributed through short-form social platforms rather than through the recorded track itself. At the same time, they involve deliberate partner selection and repeated interaction among professional artists, making them both promotional content and observable relational ties. By focusing on professional artist-to-artist K-pop challenges within short-form

platforms, this study aims to reveal how informal, visually-driven collaborations influence promotional outcomes, audience engagement, and evolving artist relationships in algorithmically mediated environments.

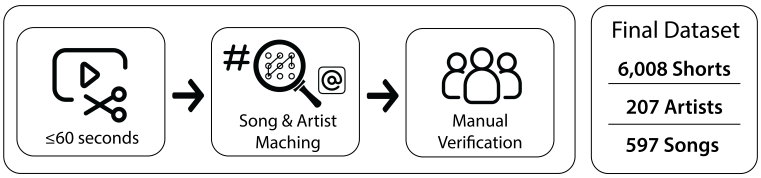
3 Data Collection

3.1 K-pop Challenge Shorts

As different K-pop artists may employ varied strategies for collaborating with their fellow artists due to diverse goals and constraints, it is crucial to analyze a *comprehensive* dataset of K-pop challenge Shorts rather than relying on sampled data. To this end, we collected all available K-pop challenge Shorts spanning from the launch of YouTube Shorts in September 2020 to October 2024 [32]. We focus on YouTube because it is the dominant media platform in Korea, where the average monthly time spent on the platform is 2,640 minutes, significantly longer than that of rival platforms, such as TikTok (633 minutes) [33].



(a) Data Collection



(b) Data Preprocessing and Validation

Fig. 2: Overview of the data collection and preprocessing pipeline for artist-to-artist K-pop challenge collaborations

To ensure the completeness of our dataset, the data collection process was systematically executed in the following four stages (Figure 2a:

First, we consolidated a comprehensive list of 480 K-pop artists, drawing from DBKpop.com[1], one of the most popular K-pop databases tracking the broad spectrum of K-pop artists in the industry. This list includes 320 groups and 160 solo

[1]https://dbkpop.com

artists, totaling 1,976 individuals. Throughout this paper, the term 'artist' is used collectively to refer to both groups and solo artists. Second, we aggregated detailed metadata, including their YouTube channel, full member names (in English and Korean), gender, agency affiliation, and debut dates by using multiple sources, such as Kprofiles.com[2], MusicBrainz[3], Melon[4], and Namuwiki[5]. When inconsistencies were found across sources, we prioritized official records from Melon and MusicBrainz for validity. From the initial list, we identified 339 artists (276 group and 63 solo artists) who had an official YouTube channel. Third, for each collected K-pop artist, we retrieved their full discography from Genie[6], a major Korean music streaming platform and filtered out artists who did not release any album after 2020, when YouTube Shorts was officially launched. After this filtering, we obtained 226 artists (188 groups and 38 soloists). Fourth, we utilized the YouTube API [34] to collect the metadata of all videos uploaded to each artist's channel.

Following data collection, we applied a series of preprocessing steps to identify eligible K-pop challenge Shorts and verify their associated songs and participating artists (Figure 2b. Since the YouTube API does not explicitly provide a label for Shorts, we identified the videos based on duration. Consistent with platform's standard definition during the data collection period [35], we filtered videos with a duration of 60 seconds or less. We then mapped the song titles and artist names using hashtags and mentions contained in the title and description. The first author manually verified all entries, annotating Shorts without explicit mentions of guest artists or songs. To further validate the annotations, a second author independently reviewed 100 randomly selected samples. This additional validation found that fewer than 2% of songs were mislabeled, while all original and guest artists were correctly labeled.

It is also common for only one member of a group to participate in a challenge. Accordingly, we checked group membership and transformed individual members' participation to their group's participation. Furthermore, it is common for K-pop group artists to form smaller 'unit' groups for better visibility of individual members. For example, the parent group NCT (25 members) has multiple unit groups like NCT 127 (9 members) and NCT Dream (7 members). As these unit groups are typically seen as extensions rather than independent acts [36], we treated unit groups as part of their parent group. Accordingly, NCT 127 was treated as NCT rather than as an independent group. We note that, as our analysis focuses on artist-to-artist collaborations, we excluded challenges that did not involve K-pop artists.

As a result, our final dataset consisted of 6,008 K-pop challenge Shorts, featuring 597 distinct songs and involving 207 unique K-pop artists.

[2]https://kprofiles.com/
[3]https://musicbrainz.org/
[4]http://www.melon.com
[5]https://en.namu.wiki
[6]https://www.genie.co.kr/

3.2 Additional Data for Modeling

3.2.1 Popularity of Challenges, Songs, and Artists

We collected data to capture the popularity of K-pop challenges, the songs featured in them, and the participating artists, providing a foundation for analyzing engagement and commercial performance.

For the popularity of K-pop challenges, we collected their view and comment counts on YouTube. Each comment includes a timestamp, allowing us to analyze temporal evolution of user engagement.

For the popularity of songs, we used Circle Chart[7], maintained by the Korea Music Content Industry Association. It is widely recognized as a reliable measure of song performance [37]. Among the several charts provided by Circle Chart, we used the Global K-pop Chart that tracks streaming performance across global platforms. Because this chart has been available only since January 2022, we restricted our analysis to challenges published from January 2022 onward across all three research questions.

For the popularity of artists, we used physical album sales data retrieved from Circle Chart's Retail Album Chart. Similar to Billboard rankings, album sales are widely recognized as a key indicator of commercial success [10].

3.2.2 Musical Characteristics of Songs

To capture the musical profiles of songs featured in K-pop challenges, we manually collected musical attributes from Tunebat [8], which sources its data from Spotify. These include tempo (BPM), energy, danceability, valence, loudness, acousticness, instrumentalness, liveness, speechiness, and track duration.

3.2.3 Previous Engagement Between Artists

Many K-pop challenges occur between artists with pre-existing relationships, often formed through joint appearances on TV shows [38, 39]. Given their tight schedules, such arrangements are both practical and common in the industry. To capture this prior engagement, we collected media-related news articles using Naver News API [40], the largest news portal service in Korea.

We measured prior engagement at two levels. First, we constructed a more specific measure of direct media interaction by retrieving articles that mention both artists together with the keywords "출연" (appearance) and "방송" (TV show). Second, we constructed a broader measure of prior association by retrieving articles that co-mention both artists without requiring the broadcast-related keywords. This broader measure captures general media co-occurrences and may reflect a wider range of public associations between two artists.

3.2.4 Nationality of Artists

K-pop groups are becoming increasingly multinational and diverse in terms of member composition reflecting the industry's global orientation [41]. This diversity can play a

[7] https://circlechart.kr
[8] https://tunebat.com/

role in shaping collaboration patterns in K-pop challenges, where cross-cultural appeal and shared identities may encourage artists to engage with one another. To capture this aspect, we collected the nationality of artists.

4 RQ1. Trends, Popularity, and Network Structure of Artist-to-Artist Challenge Collaborations

In this section, we provide a data-driven analysis of artist-to-artist challenge collaborations to address our first research question: What are the structural and temporal characteristics of artist collaborations in K-pop challenge Shorts?

4.1 Adoption and Temporal Dynamics of K-pop Challenges

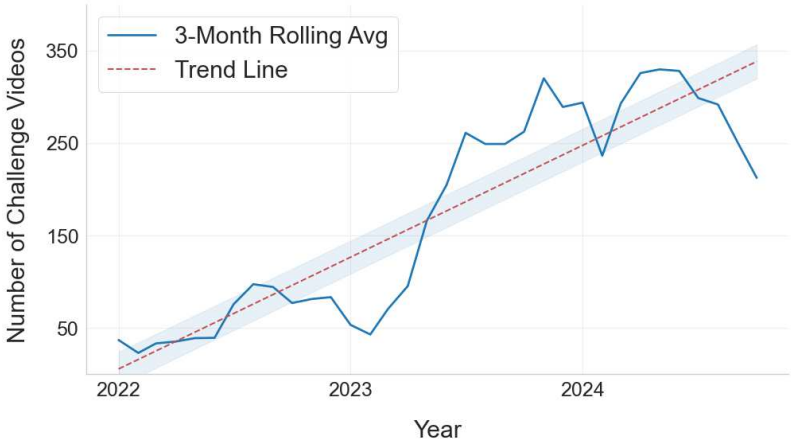
We first examined the growth of K-pop challenges to understand the current practice in the K-pop industry. Figure 3a shows a consistent upward trend with some fluctuations in challenge video uploads. A Mann-Whitney U test [42] confirms that the median number of K-pop challenge videos uploaded monthly increased significantly from 53 to 282 between 2022 and 2024 ($U = 0.0$, $p < 0.001$). This growth is not driven by a single artist or a small number of artists. In Figure 3b, the blue and red lines show substantial increases in artists engaged with K-pop challenges. Specifically, the median number of unique original artists initiating challenges per month increased from 12 to 39 ($U = 0.0$, $p < 0.001$), while guest artist participation showed an even steeper trajectory, rising from 31 to 81 ($U = 0.0$, $p < 0.001$).

More importantly, blue bars in the same figure show that, while a median of 32.0% of artists in top 200 charts [9] used K-pop challenges in 2022, this proportion increased to 64.3% in 2024 ($U = 0.0$, $p < 0.001$). Our findings confirm that K-pop challenges are growing in volume and participation, showing that they are becoming one of the common promotional activities in the industry.

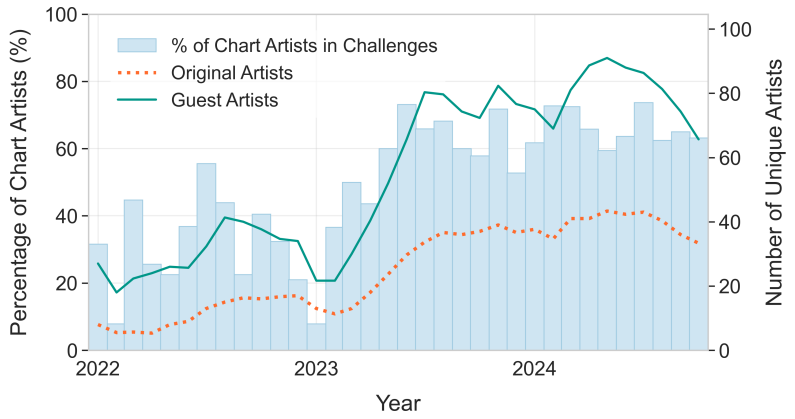
Having established the growing prevalence of K-pop challenges, we next examined *when* these challenges are uploaded relative to the song's release. While this promotional strategy should be well aligned with the song's release, the exact timing and the duration of user engagement remain unclear. We examine both the production and consumption lifecycles of challenge videos to characterize their fast-paced nature.

Figure 4 presents the distribution of challenge video uploads relative to song's release, focusing on the core range of -14 (2 weeks before the song's release) to 100 days, where the majority of activity (98.32%) occurs. The distribution reveals that 58.75% of all challenge videos are uploaded within the first two weeks of a song's release, indicating that these videos are closely aligned with the early promotional window surrounding song release. A closer look at this initial window (first two weeks) shows that upload volume does not peak immediately. Instead, challenge activity intensifies in the second week, with Week 2 uploads (1,819) surpassing Week 1 (1,711). This pattern holds consistently across all years and suggests that challenge activity extends beyond the release week and remains concentrated during the early promotional period. After

[9] Global K-pop Chart provided by Circle Chart.



(a) 3-Months Rolling Average of K-pop challenge collaboration video uploads



(b) Number of Unique Artists Participating over Time

Fig. 3: Trend in K-pop Challenges over Time

the first month, uploads decline sharply, with only 9.59% appearing between days 31 and 100. This tail period comprises 576 videos spanning 192 unique songs.

We further analyze the temporal pattern of user engagement. Using the timestamps of comments, for each challenge, we calculate the time required to accumulate 50% and 95% of its total comments, which we define as *half-life* [43] and *lifetime* of the challenge, respectively. Our analysis reveals that median half-life is 0.7 days, with 66.1% of videos accumulating half of their total engagement within 24 hours. Although challenge uploads are concentrated within the first two weeks after release, engagement with individual videos often persists well beyond this initial promotional window. The median lifetime of a challenge video is 43.3 days, with 60.2% of videos taking more than

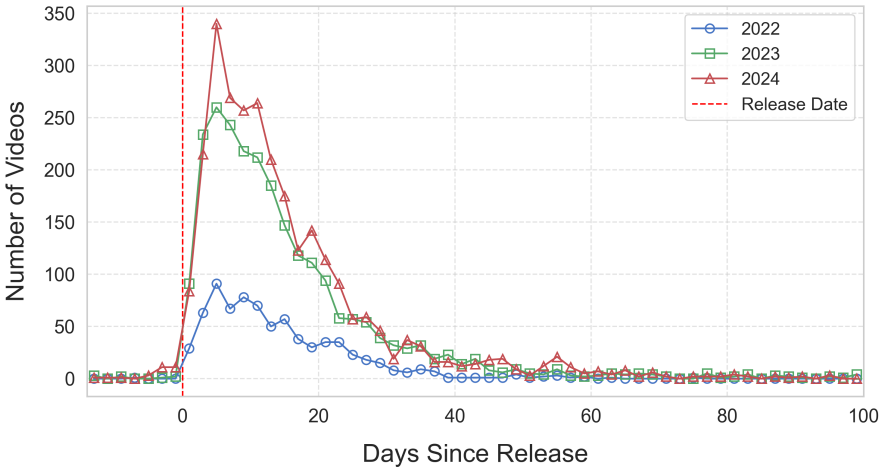


Fig. 4: Distribution of K-pop challenge uploads relative to song release dates.

30 days to accumulate 95% of their observed total comments. Taken together, these results reveal a temporal asymmetry in artist-to-artist challenge collaborations, with production concentrated near a song's release while audience engagement persists over a substantially longer period. This distinction highlights their role as short-window promotional content with a longer engagement lifecycle.

4.2 Skewed Popularity Distribution of K-pop Challenge Videos

It is well known that the popularity of YouTube videos follows 'winner-takes-all' dynamics, where only a few videos capture the majority of engagement [2, 15]. In this section, we examine whether such skewed popularity persists at two distinct scales: the global level across the entire industry and the local level within a single artist.

Figure 5 shows the skewed popularity of challenges, songs, and artists at the global level. It presents the complementary cumulative distribution of aggregate views across three dimensions: videos, songs, and artists. The horizontal axis represents the entities sorted from the most popular to the least popular, with ranks normalized between 0 and 100. It shows a heavy-tailed distribution where a small subset of challenges accounts for a disproportionately large share of total views. At video level, the top 20% of challenges capture 86% of total views, revealing that a small subset drives the majority of engagement. This disparity is more apparent at artist level, where the top 20% of artists account for 91.5% of all views, which highlights the influence of "superstars" in the ecosystem. The trend persists at the song level, where the top 10% of tracks capture 70.9% of the total view counts. These findings indicate that while challenges allow for broad participation, the visibility is disproportionately concentrated among a small subset.

To examine the skewness of the popularity at the local level, we calculated the Gini coefficient, a standard measure of distributional inequality [44], ranging from 0 (perfect equality) to 1 (perfect inequality). Within an individual artist's channel, we

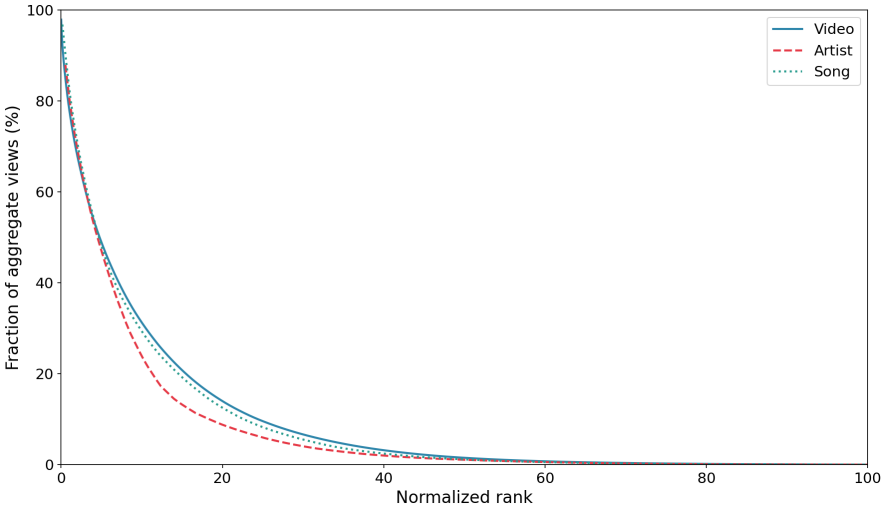


Fig. 5: Skewness of view counts across videos

assess inequality at two levels: (1) across songs (based on aggregated views per song) and (2) across individual challenge Shorts. For the song-level analysis, we included only songs with at least four challenge videos, which account for 75% of all songs.

Within an artist's channel, engagement is highly unequal across individual Shorts (mean Gini = 0.469). To assess whether songs differ in engagement beyond this Short-level variability, we compare the observed song-level Gini (mean = 0.292) with a null distribution obtained by randomly reassigning Shorts to songs within each artist while preserving the number of Shorts per song (1,000 permutations). This baseline is necessary because the Gini coefficient is biased downward when computed over few items and because aggregating Shorts into songs mechanically smooths inequality. The observed song-level Gini is much higher than the permutation baseline (mean = 0.214, 95% null interval = [0.202, 0.227], two-sided $p = 0.002$), indicating greater inequality in engagement across songs than would be expected under random assignment of Shorts to songs.

Taken together, attention to artist-to-artist challenge collaborations is highly concentrated both across the industry and within individual artists' channels, confirming the well-known heavy-tailed popularity of online video [15]. Since challenges featuring the same song can attract very different levels of attention, the choice of collaborator may matter. This raises the question of how artists select their collaborators, which we turn to next.

4.3 Macro and Micro Structural Patterns of Collaboration

While the artist-to-artist challenge collaborations are actively adopted and generate sustained user engagement (§4.1), the popularity of challenges and artists are highly skewed (§4.2). Examining who collaborates with whom allows us to assess whether

these ties are organized differently from random pairing and how this structure evolves over time.

In this section, we analyze how artists collaborate with each other through a directed, weighted network where nodes represent artists and edges denote their co-occurrences in challenge Shorts. We add a directed edge from an original artist to a guest artist when at least one collaboration Short features both artists, and we set the edge weight to the number of such videos between the ordered pair. The resulting network contains 207 nodes and 4,369 distinct directed artist pairs. For visual clarity, Figure 6 displays only the top 1% of edges by weight; all statistics below are computed on the full network.

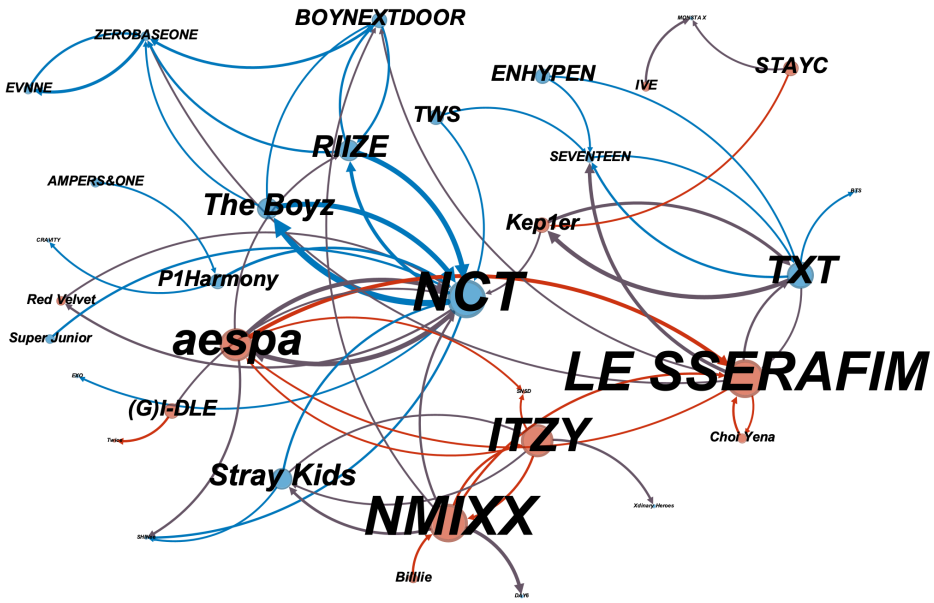


Fig. 6: Artist Collaboration Network visualized using Gephi [45]. For visualization purposes, top 1% of collaboration relationships (by frequency) were selected. The colors are coded to indicate the gender of each artist.

We compute network density on the unweighted projection (an edge exists if its weight is > 0). The standard directed edge density excluding self-loops is $D = |E|/(|V|(|V|-1)) = 4{,}369/(207 \times 206) \approx 0.102$, meaning that 10.2% of all possible ordered artist pairs are observed. Despite this sparsity, the network exhibits high reciprocity: 63.9% of collaborations are bidirectional, suggesting that artists frequently exchange challenge invitations and reflecting a norm of mutual promotion.

Despite this high reciprocity, repeated collaborations between the same ordered pair are uncommon: 77.5% of edges have a weight of 1. Considering that artists release

an average of 3.2 songs yet engage with a median of 16 partners, this low repetition rate suggests a diversification strategy in which artists continually broaden their reach to maximize cross-audience exposure. At the individual level, artists invite a median of 16 unique partners, which is comparable to the median of 17 invitations they receive. Finally, given the highly skewed distribution of album sales, we log-transform sales and find that they correlate more strongly with the number of unique partners from whom artists receive invitations ($r = 0.515$) than with the number they invite ($r = 0.433$). This difference is significant according to Steiger's test [46] for dependent correlations ($z = 2.313$, $p = 0.021$), indicating that commercial success translates into greater demand as a guest collaborator.

We now shift our focus to micro-structural patterns to uncover local dynamics obscured at the macro-level. To this end, we employ network motifs, which are three- or four-node subgraph patterns that appear more frequently in the observed network than in randomized networks [47]. Milo et al. [47] have shown that motif profiles, based on the Z-score of thirteen three-node motifs, serve as signatures of the network functions across social, biological, or scientific systems. For example, an over-representation of fully connected triads compared to a randomized network is considered a signature of social networks, reflecting the high reciprocity of social interaction [48].

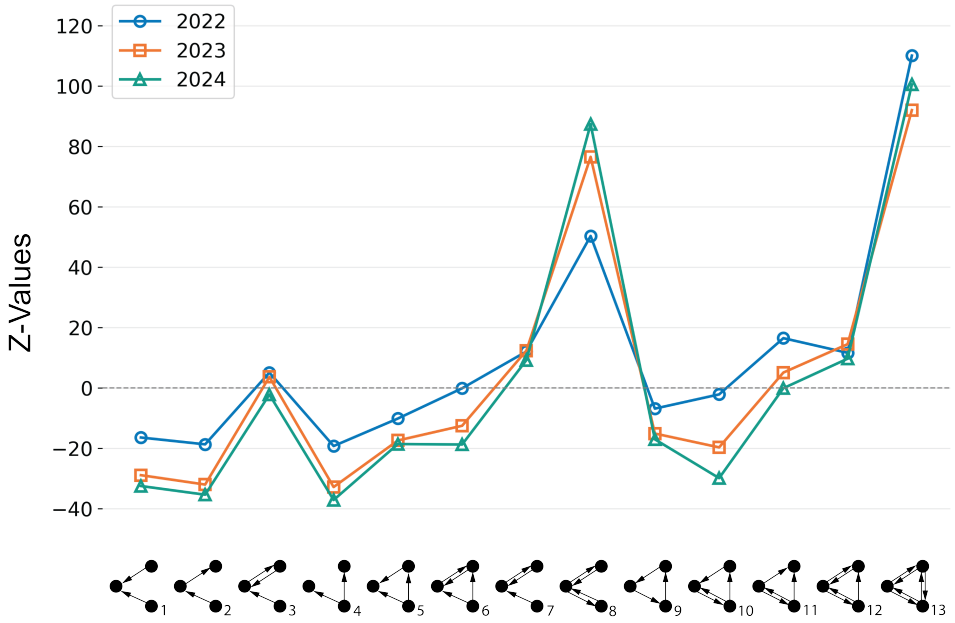


Fig. 7: Z-values of 13 three-node motifs in the artist collaboration network (2022-2024)

We conducted motif analysis using FANMOD [49], implemented with igraph [50], focusing on three-node (triadic) motifs. As a null model, we generated 5,000 random graphs that preserved the degree distribution of the original network. For

each motif, we calculated a Z-value, with positive and negative values indicating over-representation and under-representation, respectively.

Figure 7 shows the motif profile of yearly artist collaboration networks from 2022 to 2024, highlighting its stable structural characteristics. We observe that non-reciprocal structures, such as feed-forward loop (motif 2), fan-in (motif 1), and fan-out (motif 4) are under-represented. In contrast, fully connected triads (motif 13) are over-represented. These findings indicate that the collaboration network shares structural characteristics with social networks, where reciprocal and densely connected motifs are dominant [51]. This observation aligns with our earlier analysis in the previous section, where the network was shown to be sparse overall yet characterized by a high degree of reciprocity.

A longitudinal analysis from 2022 to 2024 shows that the overall shape of the network profile is relatively stable but with some noticeable changes over time. In the same figure, non-reciprocal motifs associated with hierarchical structures, such as fan-in (motif 1), fan-out (motif 4), feed-forward loop (motif 2), and middleman (motif 5) show declining z-values over the years. This change suggests that artist collaborations become increasingly reciprocal and less hierarchical within the K-pop challenge ecosystem.

Conversely, motif 8 becomes more prevalent in recent years, which contrasts with what is typically observed in social networks. This motif has one central node who reciprocally interacts with two other nodes, while those two remain unconnected. This configuration parallels the concept of a structural hole where a central actor maintains social capital by bridging unconnected peers [52].

The longitudinal analysis also reveals that fully connected triads (motif 13) rarely persist across all three years. One plausible explanation is that collaborations in K-pop challenges are situation-driven; artists with overlapping release schedules ephemerally help promote each other rather than sustaining structurally rigid triads over long periods. However, as one may expect, artists under the same agency or label are better able to maintain these fully connected triads. For example, the following fully connected triads were observed across two years: [NCT - RIIZE - Red Velvet] (from SM Entertainment), [ENHYPEN - SEVENTEEN - TXT] (from HYBE Labels), and [ITZY - NMIXX - Stray Kids] (from JYP Entertainment).

In conclusion, our structural analysis reveals that K-pop collaborations are governed by a complex interplay of social reciprocity and strategic necessity. While the network exhibits high reciprocity typical of social circles, it also shows noticeable deviations from standard social relationships. These distinct structural patterns imply that collaboration decisions are not random but driven by specific underlying factors. In the following section, we investigate these determinants by constructing a predictive model for collaboration formation.

5 RQ2. Partner Selection in Artist Collaboration

Our network analysis shows that most collaboration ties are one-off (77.5% of directed edges have a weight of 1), while a smaller subset of pairs collaborate repeatedly. Prior work on online music co-creation suggests that occasional and recurring collaborations

Category	Feature Name	Description
Temporal Proximity	Has recent title track	Whether the guest artist released a title track within two weeks before or after the song featured in the Shorts video (binary)
Social Proximity	Debut date difference	Difference in debut days between original and guest artist
	Same gender	Whether both artists share the same gender (binary)
	Same agency	Whether both artists are from the same agency (binary)
	Past engagement	Level of prior engagement between the artists outside of collaboration videos
	Korean member ratio (Original Artist)	Proportion of members of original artist who are Korean
	Korean member ratio (Guest Artist)	Proportion of members of guest artists who are Korean
	Shared non-korean nationality	Whether both artists share at least one member with the same non-Korean nationality (e.g. 1 if both have a Japanese member, and 0 otherwise).
Music Similarity	Music similarity	Measures the similarity between the song used in the Shorts video (by the original artist) and the guest artist's most recent title track
Original Artist Song & Guest Artist Song Attributes	BPM	Tempo, measured in beats per minute
	Energy	How intense and active the track is (out of 100)
	Danceability	How appropriate the track is for dancing based on overall regularity (out of 100)
	Happiness	How cheerful and positive the track is (out of 100)
	Acousticness	How likely the track is acoustic (out of 100)
	Instrumentalness	How likely the track contains no spoken word vocals (out of 100)
	Liveness	How likely the track was recorded with a live audience (out of 100)
	Speechiness	How present spoken words are in the track (out of 100)
	Loudness	Average decibel amplitude across the track ranging from -60dB to 0dB
Popularity	Average past album sales (Original Artist)	Average past album sales of the original artist, serving as a proxy for the artist's popularity at a given time
	Average past album sales (Guest Artist)	Average past album sales of the guest artist, serving as a proxy for the artist's popularity at a given time
	Ratio of average album sales	Ratio of the past album sales between the original and guest artists

Table 1: Overview of input features used for collaboration prediction model. Features are grouped into six categories: temporal proximity, social proximity, music similarity, original artist song, guest artist song, and popularity. Song-related features (except music similarity) are listed separately for the original and guest artists, with audio feature definitions based on Tunebat [2025].

are driven by different motivations: occasional collaborations tend to serve strategic or popularity-driven goals, whereas recurring ones reflect relational history and shared norms [26]. We therefore build separate prediction models for first-time and recurring collaborations to examine which social, temporal, musical, and popularity-related factors are most informative for partner selection, and how their relative importance differs between the two settings. We use Random Forest [53] because it can capture nonlinear relationships and interactions among heterogeneous features while supporting feature-level interpretation.

To structure our analysis, we organize features into six categories (Table 1): (1) Social Proximity, (2) Original Artist Song, (3) Guest Artist Song, (4) Musical Similarity, (5) Popularity, and (6) Temporal Proximity. **Social proximity** features capture the relational ties between artists. These include demographic alignments (debut date differences, same agency, same gender, proportion of Korean members, and shared non-Korean nationality; see § 3.2.4) and interaction history (past collaboration Shorts and past media engagement). **Temporal Proximity** captures overlap in artists' promotional schedules, defined as whether the guest artist released a song within $\pm$ 2 weeks of the original artist's song. **Song-related features** consist of audio descriptors for both the original artist's challenge song and the guest artist's most recent title

track. **Music similarity** quantifies the tonal proximity between these two tracks (see § 3.2.2). Finally, **Popularity** features include the average past album sales of both artists and their sales ratio reflecting relative commercial standing.

5.1 Experimental Setup

We implement a train-test split that preserves the chronological order of collaborations, ensuring that only past data are used to predict future events. Specifically, we use the earliest 80% of data for training and the most recent 20% for testing. We treat collaborations as directed relationships, such that $A \to B$ and $B \to A$ are considered distinct collaboration events.

Since the dataset includes only positive samples, we construct a balanced dataset by generating an equal number of negative samples based on artist participation frequency as guest artists. Reflecting the skewed activity distribution (§ 4.2), guest artists are grouped into High (top 30%), Medium (next 30%), and Low (bottom 40%) tiers according to their participation frequency. Robustness checks using alternative thresholds and purely random sampling confirm that model performance and feature importance remain consistent.

For both models, negative samples are constructed by anchoring each observed collaboration to the same original artist and upload date, while replacing the observed guest artist with an eligible alternative. By holding the original artist and temporal context fixed, this design creates a more comparable counterfactual pairing and reduces differences arising from artist-specific activity or time-varying conditions. For the first-time collaboration model, the replacement guest artist is sampled from a similar collaborator-activity tier and must not have previously collaborated with the original artist before the corresponding upload date. For the recurring collaboration model, the replacement guest artist is likewise drawn from a similar activity tier, but eligibility is restricted to artists who had previously collaborated with the original artist in the same direction before that date. The negative therefore represents an existing relationship that could have recurred but was not the observed collaboration at that time.

During sample construction, 13 first-time and 46 recurring collaborations were excluded due to feature availability and negative-sampling eligibility constraints, respectively. As a result, we retained 4,508 first-time and 1,441 recurring collaborations as positive samples, each paired with one negative sample through stratified sampling by artist activity to construct balanced 1:1 datasets. We use ROC-AUC as the primary metric for model evaluation. Random Forest hyperparameters were selected separately for the first-time and recurring collaboration models through grid search [55] on the training data. Both models used a maximum tree depth of 5, a minimum of 30 samples to split, and a minimum of 15 samples per leaf. The first-time model used 300 trees, while the recurring model used 200 trees.

5.2 Results: Determinants of First-time vs. Recurring Collaborations

The two models show moderate predictive performance. On the balanced test set, the first-time model achieves ROC-AUC of 0.7549, while the recurring collaboration model achieves ROC-AUC of 0.6711. These results indicate that first-time collaborations exhibit a clearer predictive structure, whereas recurring collaborations are more difficult to distinguish from plausible alternative guest artists.

The feature importance analysis further reveals distinct patterns between first-time and recurring collaborations. For first-time collaborations, the strongest predictor is whether the guest artist has a recent title track released within the same promotional window (has-recent-title-track; importance = 0.3827). This is followed by same-agency affiliation (0.0771), debut-date difference (0.0739), same gender (0.0569), and the ratio of average past album sales between the two artists (0.0436). The SHapley Additive exPlanations (SHAP) analysis further clarifies the direction of these associations. Having a recent title track consistently shifts model predictions toward collaboration, while the absence of a recent release shifts predictions in the opposite direction. Same-agency and same-gender pairs also contribute positively to predicted collaboration likelihood, while smaller debut-date differences are generally associated with more positive predictions. These features suggest that the formation of new collaboration ties is strongly associated with temporal opportunity and relational proximity. Together, overlapping promotional schedules may create practical opportunities for artists to participate in one another's challenges, while similarities in career stage and agency affiliation may further facilitate new partnerships.

Recurring collaborations exhibit a different pattern. The ratio of average album sales between original and guest artists has the highest feature importance (0.0996), followed by the guest artist's average past album sales (0.0989), debut-date difference (0.0653), same-agency affiliation (0.0575), and the guest artist's danceability (0.0564). SHAP values indicate that higher guest-album sales, same-agency and same-gender relationships, smaller debut-date differences, and overlapping release schedules generally shift predictions toward recurring collaboration. Overall, feature importance is more evenly distributed in the recurring model, with no single feature showing the same dominance as temporal proximity in the first-time model.

An ablation study (Table 2) further clarifies the distinct contributions of each feature group. For first-time collaborations, music-related features provide only limited predictive power, with ROC-AUC increasing from 0.0.5137 to 0.5625 after incorporating original- and guest-artist musical attributes. Adding *Social Proximity* further improves ROC-AUC to 0.6136, while *Popularity* provides only a modest additional gain to 0.6293. The largest improvement occurs after incorporating *Temporal Proximity*. This pattern indicates that overlapping promotional schedules provide particularly strong predictive information for distinguishing first-time collaboration partners.

In contrast, recurring collaborations show a more distributed pattern of contribution across feature groups. *Social Proximity* features improve performance. However, *Popularity* does not yield an additional improvement, with ROC-AUC decreasing slightly to 0.6536 from 0.6636, while the full model including *Temporal Proximity* reaches 0.6709 overall. This pattern suggests that recurring collaborations are shaped

by a broader combination of musical and relational characteristics rather than being dominated by a single feature category.

To assess robustness beyond the balanced test set, we additionally evaluate the trained models against the full set of eligible collaboration partners available at each observed collaboration date. Each unique original artist and date in the held-out test period is treated as a ranking query, with all observed guest artists on that date labeled as positives and all other eligible artists as alternatives. Rather than sampling a fixed number of negative pairs, we rank all eligible candidates using the predicted probabilities from the trained models. We evaluate ranking performance using Mean Average Precision (MAP). To account for differences in candidate-set size and the number of observed collaborators, we compare MAP against the expected MAP under a uniformly random ranking calculated separately for each query. For first-time collaborations, the model achieves a MAP of 0.1520, compared with a random-ranking baseline of 0.0342, corresponding to a 4.45× lift over random ranking. For recurring collaborations, MAP is 0.2308 compared with a higher random baseline of 0.1336, corresponding to a 1.73× lift.

These candidate-set results reinforce the distinction between the two prediction tasks. Although both models rank observed collaborators above their respective random-ranking baselines, the first-time model shows a substantially larger improvement over random ranking. In contrast, the recurring model shows more limited separation between observed collaborators and plausible alternative recurring partners. This suggests that first-time collaboration formation is more systematically associated with observable factors such as overlapping promotional schedules and social proximity, whereas the recurrence of an existing partnership may depend more heavily on contextual or unobserved factors not captured by the current feature set.

Feature Group	First-Time	Recurring
MS	0.5137	0.4530
MS+OAS	0.5234	0.5165
MS+OAS+GAS	0.5625	0.6321
MS+OAS+GAS+SP	0.6136	0.6636
MS+OAS+GAS+SP+P	0.6293	0.6536
MS+OAS+GAS+SP+P+TP	0.7549	0.6711

Table 2: Impact of feature groups on ROC-AUC (Ablation Study). Feature groups: SP (Social Proximity), OAS (Original Artist Song), GAS (Guest Artist Song), MS (Music Similarity), and TP (Temporal Proximity), and P (Popularity)

6 RQ3. Collaboration and Musical Success

Professional artist-to-artist challenge collaborations have become a common promotional strategy, yet their association with a song's commercial success remains empirically underexplored. We use Granger causality tests [56] to examine whether professional challenge-collaboration activity in YouTube Shorts temporally precedes

movements in a song's streaming-chart performance. Because Granger tests capture predictive lead–lag relationships rather than mechanistic causality, we interpret statistical significance as evidence of temporal ordering (predictive association), not definitive causal impact.

We model the dynamic relationship between daily chart performance and three challenge-based signals: (1) the number of artist-uploaded collaboration challenge Shorts, (2) engagement on those collaboration Shorts (daily comment volume), and (3) participation by top-tier guest artists (top 20% by cumulative past album sales). We align all time series to Circle Chart days and operationalize chart performance as a bounded *chart score* on $[0, 200]$: rank 1 is mapped to 200, rank 200 is mapped to 1, and days in which a song does not appear in the top-200 chart are assigned 0. This yields complete daily series without missing ranks, but introduces a floor at 0 due to top-N truncation.

For each song, we construct a 60-day time series beginning on the official release date to estimate the temporal relationships with a focus on the song's initial promotional period. Three separate Toda–Yamamoto Granger models are then estimated, pairing chart score with each challenge-based signal and testing both directions. The maximum integration order, m, is determined using Augmented Dickey–Fuller (ADF) [57], with Kwiatkowski–Phillips–Schmidt–Shin (KPSS) [58] tests used as a diagnostic cross-check. The base VAR lag order, p, is selected using Bayesian Information Criterion (BIC) [59]. Residual whiteness is additionally assessed using a Portmanteau test [60] to verify that the selected lag structure adequately captures temporal dependence in the series. An augmented VAR with $p+m$ lags is then estimated, and Granger causality is assessed using a modified Wald test of the joint significance of the predictor's first p lagged terms [61]. We applied the Benjamini–Hochberg [62] procedure to control the false discovery rate (FDR) at 5% within each testing family. As a result, the Granger causality analysis includes 92 songs for collaboration Shorts, 95 for comments, and 84 for top-tier collaboration Shorts.

Across the three challenge-related signals, the results reveal limited evidence that these activities statistically precede and predict chart success. Specifically, daily collaboration Shorts precede ranking changes in only 2 of 92 songs (2.2%), while collaboration Shorts involving top-tier guest artists do so in 1 of 84 songs (1.2%). In the reverse direction, chart performance precedes collaboration Shorts and comments and replies in 2 songs each, and top-tier collaboration Shorts in 2 songs. Overall, for the majority of songs across all signals, no significant Granger causality was detected in either direction. The prevalence of null results highlights a distinction between the intention to generate virality and the outcome of commercial popularity. A similar distinction is reflected in the separation of Viral and Top charts on major streaming platforms such as Spotify, where rapidly circulating content does not necessarily correspond to the most-streamed music. While the act of producing K-pop challenge content represents a deliberate attempt to generate virality, our results suggest that this effort optimizes for visual consumption, which functions independently from and does not automatically convert into the active listening behaviors required for chart

performance. These findings suggest that challenge activity may play a complementary role within broader promotional strategies, and we find little evidence that it drives streaming performance.

To further understand the structural utility of artist-to-artist challenge collaborations despite their limited chart impact, we compare engagement levels between collaboration Shorts and non-collaboration Shorts for the same set of songs. Non-collaboration short videos were filtered using the following process: (1) keeping only Shorts associated with songs that also had collaboration Shorts, (2) limiting videos to those whose durations fell within the empirical distribution of challenge videos with the same music, (3) excluding titles with keywords "with" in both English and Korean, and (4) performing manual checks. In total, there were 5,555 short videos.

	Collaboration Shorts	Non-Collaboration Shorts
Median # View	64,016.5	38,913
Median # Like	7,157.5	4,794
Median # Comment	70	52

Table 3: Median Engagement Metrics per Video Type

Unlike the Granger causality results, collaboration Shorts significantly outperformed non-collaboration Shorts across all engagement metrics (Table 3). Given strongly right-skewed and heteroscedastic outcome distributions, we report medians (IQR) and use two-sided Mann–Whitney U tests with Bonferroni correction across three outcomes. Collaboration Shorts had higher views (64,017 [11,990–413,794] vs. 38,913 [5,659–236,710]; +64.5%; $p_{\text{adj}} = 2.40 \times 10^{-41}$), likes (7,158 [1,518–38,371] vs. 4,794 [952–27,152]; +49.3%; $p_{\text{adj}} = 1.04 \times 10^{-17}$), and comments (70 [27–224] vs. 52 [21–174]; +34.6%; $p_{\text{adj}} = 3.10 \times 10^{-22}$). Mean differences were directionally consistent (views +47.0%, likes +41.7%, comments +32.9%). In other words, the presence of a guest artist significantly boosts the content's reach and engagement within the platform. These findings reveal a distinct structural utility of artist collaborations. While they may not function as a direct lever for streaming conversion, they are still highly effective mechanisms for increasing engagement from audiences who are not typically reachable.

7 Discussion and Conclusion

This study provides a comprehensive, data-driven examination of artist-to-artist challenge collaborations, a unique and understudied form of promotional content. By analyzing structural patterns, modeling collaboration, and examining its temporal relationship with chart performance, we unveil how artists or agencies strategically engage with short-form content within the socially-mediated landscape of digital music promotion.

Our findings reveal three key insights. First, professional artist-to-artist challenge collaborations operate under a mutualistic norm supported by a highly reciprocal artist collaboration network. Second, partner selection behind collaboration formation differs

between first-time and recurring ties. First-time collaborations are most strongly associated with overlapping promotional schedules and social proximity, whereas recurring collaborations show a more distributed pattern in which guest popularity, social proximity, and musical characteristics become relatively more important. Third, evidence that artist-to-artist challenge collaborations systematically precede chart success is limited. This limited effect may be linked to broader factors shaping chart performance such as fandom-led activities, agency-level promotions, and cross-platform virality, which can overshadow the influence of challenge activity.

These findings contribute to a growing understanding of how short-form content mediates visibility in the music industry, showing how artists strategically collaborate within platform-driven spaces. While our dataset exclusively focuses on professional artists to capture industry-level dynamics, we acknowledge that fan-driven virality is another major component of the challenge phenomenon. However, our focus on artist-to-artist collaboration provides a crucial "top-down" perspective on how professional social capital is negotiated, which serves as a necessary foundation for future work understanding how subsequent fan participation is legitimized and encouraged.

Nonetheless, our study has several limitations. First, our analysis focuses exclusively on YouTube Shorts and does not include other short-form platforms such as TikTok. Although TikTok is also a major landscape for challenge dissemination, our study centers on official artist-to-artist collaborations used as promotional content. In practice, artists frequently upload the same collaborative Shorts across multiple platforms, which reduces platform-specific variation in content itself. This choice allowed us to utilize YouTube as a primary data source for this particular collaborative phenomenon. Future work could extend this analysis by incorporating cross-platform audience engagement to capture potential platform-specific differences. Second, the dataset is limited to K-pop artists, excluding collaborations with non-musician celebrities and social media influencers whose participation may affect challenge visibility differently. Third, our use of Naver News articles to infer prior engagement may introduce noise due to the inherent limitations of a keyword-based search, which could be addressed with advanced NLP methods. Moreover, this measure provides only a bounded approximation of artist relationships. Because media-based interactions occur at multiple levels, this approach may miss informal ties. Finally, our models do not account for possible latent variables. Regarding social proximity, our features rely on public interactions, which may overlook hidden relationships formed during shared training periods, school ties, or personal connections that are characteristic of the K-pop industry. Furthermore, we do not capture broader commercial drivers, such as specific agency-level promotional strategies or organized fan-driven activities.

Future work could expand this study by incorporating cross-platform analysis. Additionally, examining the role of fandom communities in shaping collaborative visibility and success would offer deeper insight into how digital ecosystems mediate promotional outcomes. Future research can also develop a recommendation system to identify "optimal" matches for collaboration. Our findings suggest that partner selection reflects a combination of temporal, social, musical, and popularity-related factors, providing a basis for exploring data-driven collaboration recommendations. This modeling of partnership selection can extend beyond K-pop to the broader creative field.

Abbreviations. **K-pop** Korean Pop; **API** Application Programming Interface; **BPM** Beats Per Minute; **ROC-AUC** Receiver Operating Characteristic Area Under the Curve; **SHAP** SHapley Additive exPlanations; **MAP** Mean Average Precision; **CI** Confidence Interval; **MS** Music Similarity; **OAS** Original Artist Song; **GAS** Guest Artist Song; **SP** Social Proximity; **TP** Temporal Proximity; **P** Popularity; **ADF** Augmented Dickey-Fuller; **KPSS** Kwiatkowski-Phillips-Schmidt-Shin; **VAR** Vector Autoregression; **BIC** Bayesian Information Criterion; **FDR** False Discovery Rate; **IQR** Interquartile Range

Supplementary Information

Supplementary information. None

Acknowledgements. Not Applicable

Funding. Not Applicable

Data Availability. The dataset used for the current study is available [here]

Declarations

Conflict of interest/Competing interests. The authors declare that they have no competing interests

Author contribution. A.C. designed research, performed data collection and analysis, and wrote the paper. J.A. designed and supervised research and wrote the paper. H.K. designed and supervised research and wrote the paper.

Ethics approval and consent to participate. This study uses only publicly available YouTube data. Comment data were used solely in aggregate, and no personally identifiable information was collected or analyzed.

Consent for publication. Not Applicable

References

[1] Parc, J., Kim, S.D.: The digital transformation of the korean music industry and the global emergence of k-pop. Sustainability **12**(18), 7790 (2020)

[2] Meier, L.M.: 24. popular music, streaming, and promotional media: Enduring and emerging industrial logics. Making media, 321 (2019)

[3] McIntyre, H.: K-Pop's Global Impact: 19 Of The 20 Bestselling Albums In 2023 Are From South Korea (2024). https://www.forbes.com/sites/hughmcintyre/2024/03/27/k-pops-global-impact-19-of-the-20-bestselling-albums-in-2023-are-from-south-korea/ Accessed 2025-03-27

[4] Jung, S.: K-pop, indonesian fandom, and social media. Transformative Works & Cultures **8** (2011)

[5] Putri, N.W.L.M., Dilia, D., *et al.*: Analyzing social media strategy in korean music industries. PERSPEKTIF **12**(4), 1206–1214 (2023)

[6] Duffy, B.E., Wissinger, E.: Mythologies of creative work in the social media age: Fun, free, and "just being me". International Journal of Communication **11**, 20 (2017)

[7] Jin, D.Y., Ryoo, W.: Critical interpretation of hybrid k-pop: The global-local paradigm of english mixing in lyrics. Popular Music and Society **37**(2), 113–131 (2014)

[8] Liang, J., Miao, H., Sai, R.: TikTok Challenges and Cross-Cultural Success: BLACKPINK's Global Reach Through Short-Form Video in the K-Pop Industry. Interdisciplinary Humanities and Communication Studies **1**(10) (2024) https://doi.org/10.61173/3nyan341

[9] Nanobanana: Nanobanana AI Image Generator. Accessed: 2025-09-15 (2025). https://nanobanana.ai/

[10] Kang, I., Mandulak, M., Szymanski, B.K.: Analyzing and predicting success of professional musicians. Scientific Reports **12**(1), 21838 (2022)

[11] Silva, M.O., Oliveira, G.P., Seufitelli, D.B., Lacerda, A., Moro, M.M.: Collaboration as a driving factor for hit song classification. In: Proceedings of the Brazilian Symposium on Multimedia and the Web, pp. 66–74 (2022)

[12] Herwees, T.: How TikTok Is Rewriting the Rules for Emerging Artists (2025). https://www.rollingstone.com/music/music-news/tiktok-rewriting-rules-emerging-artists-1235319787/ Accessed 2025-09-16

[13] Park, M., Weber, I., Naaman, M., Vieweg, S.: Understanding musical diversity via online social media. In: Proceedings of the International AAAI Conference on Web and Social Media, vol. 9, pp. 308–317 (2015)

[14] Kim, Y.H., Lee, D., Han, N.G., Song, M.: Exploring characteristics of video consuming behaviour in different social media using k-pop videos. Journal of Information Science **40**(6), 806–822 (2014)

[15] Cha, M., Kwak, H., Rodriguez, P., Ahn, Y.-Y., Moon, S.: I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In: Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, pp. 1–14 (2007)

[16] Stone, P.W.: Rise of the microsong: the platformification of music on tiktok as

an artform. INSAM Journal of Contemporary Music, Art and Technology (13), 69–84 (2024)

[17] Barata, M.L., Coelho, P.S.: Music streaming services: understanding the drivers of customer purchase and intention to recommend. Heliyon **7**(8) (2021)

[18] Liang, T.-P., Lai, H.-J., Ku, Y.-C.: Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings. Journal of Management Information Systems **23**(3), 45–70 (2006)

[19] Violot, C., Elmas, T., Bilogrevic, I., Humbert, M.: Shorts vs. regular videos on youtube: a comparative analysis of user engagement and content creation trends. In: Proceedings of the 16th ACM Web Science Conference, pp. 213–223 (2024)

[20] Rajendran, P.T., Creusy, K., Garnes, V.: Shorts on the rise: assessing the effects of youtube shorts on long-form video content. arXiv preprint arXiv:2402.18208 (2024)

[21] Nwagwu, W.E., Akintoye, A.: Influence of social media on the uptake of emerging musicians and entertainment events. Information Development **40**(4), 667–692 (2024)

[22] Kim, J.-h., Jung, S.-h., Roh, J.-s., Choi, H.-j.: Success factors and sustainability of the k-pop industry: A structural equation model and fuzzy set analysis. Sustainability **13**(11), 5927 (2021)

[23] Abidin, C., Lee, J.: K-pop tiktok: Tiktok's expansion into south korea, tiktok stage, and platformed glocalization. Media International Australia **188**(1), 86–111 (2023)

[24] Ortega, J.L.: Cover versions as an impact indicator in popular music: A quantitative network analysis. Plos one **16**(4), 0250212 (2021)

[25] Leppard, T.R., Davis, A.P.: The ties that rhyme: Duality in symbolic and structural networks of grime music. The British Journal of Sociology (2026)

[26] Calefato, F., Iaffaldano, G., Trisolini, L., Lanubile, F.: An in-depth analysis of occasional and recurring collaborations in online music co-creation. ACM Transactions on Social Computing (TSC) **4**(4), 1–40 (2022)

[27] Lee, Y.G., Gargiulo, M.: Escaping the survival trap: Network transition among early-career freelance songwriters. Administrative Science Quarterly **67**(2), 339–377 (2022)

[28] Oliveira, G.P., Santos, M., Seufitelli, D.B., Lacerda, A., Moro, M.M.: Detecting collaboration profiles in success-based music genre networks. In: ISMIR, pp. 726–732 (2020)

[29] Deshmane, A., Martínez-de-Albéniz, V.: Come together, right now? an empirical study of collaborations in the music industry. Management Science **69**(12), 7217–7235 (2023)

[30] Silva, M.O., Moro, M.M.: Causality analysis between collaboration profiles and musical success. In: Proceedings of the 25th Brazillian Symposium on Multimedia and the Web, pp. 369–376 (2019)

[31] Ordanini, A., Nunes, J.C., Nanni, A.: The featuring phenomenon in music: how combining artists of different genres increases a song's popularity. Marketing Letters **29**(4), 485–499 (2018)

[32] Sherman, T.: One year of YouTube Shorts — what we've learned so far (2021). https://blog.youtube/news-and-events/one-year-of-youtube-shorts-what-weve-learned-so-far/ Accessed 2025-03-27

[33] Yoon-seo, L.: YouTube dominates Korea's media use, clocking nearly 44 hours a month. The Korea Herald (2026). https://www.koreaherald.com/article/10845629

[34] Google: YouTube Data API v3 (2025). https://developers.google.com/youtube/v3 Accessed 2025-09-15

[35] Google: Understand three-minute YouTube Shorts - YouTube Help (2025). https://support.google.com/youtube/answer/15424877?hl=en Accessed 2025-12-16

[36] Herman, T.: Examining NCT and Its Various K-Pop Units (2017). https://www.billboard.com/music/music-news/nct-k-pop-group-guide-explainer-7647980/ Accessed 2025-09-09

[37] Im, H., Song, H., Jung, J.: A survival analysis of songs on digital music platform. Telematics and Informatics **35**(6), 1675–1686 (2018)

[38] MKSports: '아는형님' 선미, 지코와 함께 '아무노래 챌린지'[MK★TV컷] (2020). https://www.mksports.co.kr/news/entertain/9426009 Accessed 2025-09-13

[39] topstarnews: '지구오락실3' 이은지 · 안유진 · 미미 · 이영지, 열정 가득한 'I AM' 챌린지 [종합] (2025). https://www.topstarnews.net/news/articleView.html?idxno=15655380

[40] Naver: 검색 > 뉴스 - Search API (2016). https://developers.naver.com/docs/serviceapi/search/news/news.md#%EB%89%B4%EC%8A%A4 Accessed 2025-03-28

[41] Jung-youn, L.: Multinational groups look to expand global scope of K-pop. Section: The Korea Herald (2024). https://www.koreaherald.com/article/3298324 Accessed 2025-09-13

[42] Mann, H.B., Whitney, D.R.: On a test of whether one of two random variables is stochastically larger than the other. The annals of mathematical statistics, 50–60 (1947)

[43] Graffius, S.M.: Lifespan (half-life) of social media posts: Update for 2022. Social Media Analytics Blog. December **28**, 2022 (2022)

[44] Dagum, C.: The generation and distribution of income, the lorenz curve and the gini ratio. Economie appliquée **33**(2), 327–367 (1980)

[45] Bastian, M., Heymann, S., Jacomy, M.: Gephi: an open source software for exploring and manipulating networks. In: Proceedings of the International AAAI Conference on Web and Social Media, vol. 3, pp. 361–362 (2009)

[46] Steiger, J.H.: Tests for comparing elements of a correlation matrix. Psychological bulletin **87**(2), 245 (1980)

[47] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U.: Network motifs: simple building blocks of complex networks. Science **298**(5594), 824–827 (2002)

[48] Stone, L., Simberloff, D., Artzy-Randrup, Y.: Network motifs and their origins. PLoS computational biology **15**(4), 1006749 (2019)

[49] Wernicke, S., Rasche, F.: Fanmod: a tool for fast network motif detection. Bioinformatics **22**(9), 1152–1153 (2006)

[50] Csárdi, G., Nepusz, T., Horvát, S., Traag, V., Zanini, F., Noom, D.: Igraph. https://doi.org/10.5281/zenodo.17014224

[51] Adler, M., Medzhitov, R.: Emergence of dynamic properties in network hypermotifs. Proceedings of the National Academy of Sciences **119**(32), 2204967119 (2022)

[52] Burt, R.S.: Structural holes. In: Social Stratification, pp. 659–663. Routledge, New York (2018)

[53] Breiman, L.: Random forests. Machine Learning **45**(1), 5–32 (2001) https://doi.org/10.1023/A:1010933404324

[54] Tunebat: Key & BPM of any song - Music Database by Tunebat (2025). https://tunebat.com/ Accessed 2025-03-20

[55] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011)

[56] Granger, C.W.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica: journal of the Econometric Society, 424–438 (1969)

[57] Dickey, D.A., Fuller, W.A.: Distribution of the estimators for autoregressive time series with a unit root. Journal of the American statistical association **74**(366a), 427–431 (1979)

[58] Kwiatkowski, D., Phillips, P.C., Schmidt, P., Shin, Y.: Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of econometrics **54**(1-3), 159–178 (1992)

[59] Schwarz, G.: Estimating the dimension of a model. The annals of statistics, 461–464 (1978)

[60] Lütkepohl, H.: New Introduction to Multiple Time Series Analysis. Springer, Berlin (2005). https://doi.org/10.1007/978-3-540-27752-1

[61] Toda, H.Y., Yamamoto, T.: Statistical inference in vector autoregressions with possibly integrated processes. Journal of econometrics **66**(1-2), 225–250 (1995)

[62] Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) **57**(1), 289–300 (1995)