

Article

Multi-Source Data Fusion for Anime Pilgrimage Recommendation: Integrating Accessibility, Seasonality, and Popularity [†]

Yusong Zhou ^{*} and Yuanyuan Wang ^{*} 

Graduate School of Sciences and Technology for Innovation, Yamaguchi University,
Yamaguchi Ube 755-8611, Japan

^{*} Correspondence: e099vgv@yamaguchi-u.ac.jp (Y.Z.); y.wang@yamaguchi-u.ac.jp (Y.W.)

[†] This paper is an extended version of our paper published in Zhou, Y.; Wang, Y. Tourist Spot Recommendation for Anime Pilgrimage Considering Anime Features and Transportation Accessibility. In Proceedings of the 2025 14th IEEE Global Conference on Consumer Electronics (GCCE), Osaka, Japan, 23–26 September 2025; pp. 557–561.

Abstract

Anime pilgrimage refers to the act of fans visiting real-world locations featured in anime works, offering visual familiarity alongside cultural depth. However, existing studies on anime tourism provide limited computational support for selecting pilgrimage sites based on contextual and experiential factors. This study proposes an intelligent recommendation framework based on multi-source data fusion that integrates three key elements: transportation accessibility, seasonal alignment between the current environment and the anime's depicted scene, and a Cross-Platform Popularity Index (CPPI) derived from major global platforms. We evaluate each pilgrimage location using route-based accessibility analysis, season-scene discrepancy scoring, and robustly normalized popularity metrics. These factors are combined into a weighted Multi-Criteria Decision Making (MCDM) model to generate context-aware recommendations. To rigorously validate the proposed approach, a user study was conducted using a ranking task involving popular destinations in Tokyo. Participants were presented with travel conditions, spatial relationships, and popularity scores and then asked to rank their preferences. We used standard ranking-based metrics to compare system-generated rankings with participant choices. Furthermore, we conducted an ablation study to quantify the individual contribution of accessibility, seasonality, and popularity. The results demonstrate strong alignment between the model and user preferences, confirming that incorporating these three dimensions significantly enhances the reliability and satisfaction of real-world anime pilgrimage planning.

Keywords: anime pilgrimage; context-aware recommendation; intelligent data analysis; multi-source data aggregation; transportation accessibility; seasonal alignment



Academic Editor: George Angelos
Papadopoulos

Received: 22 December 2025

Revised: 16 January 2026

Accepted: 16 January 2026

Published: 18 January 2026

Copyright: © 2026 by the authors.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

1.1. Background: From Visual Arts to Spatial Experiences

Anime pilgrimage refers to the act of fans visiting real-world locations depicted in anime, manga, or games. Historically, the connection between animation and real-world geography has undergone significant evolution. As early as 1974, the production of *Heidi, Girl of the Alps* introduced overseas location scouting to establish “spatial realism” in animation [1]. This tradition evolved through works like *Ultimate Superman R* and *Tenchi*

Muyo!, eventually leading to the modern phenomenon where background art serves as a direct map for tourism [2].

Today, this behavior has evolved into a significant form of “content tourism.” It is no longer just about visiting a site but engaging with the “symbolic and cultural value” of the location [3]. For fans, the physical site acts as a bridge between the virtual narrative and reality, requiring a high degree of contextual alignment between the anime scene and the actual environment.

1.2. Social Impact and Economic Motivation

The economic and social implications of anime pilgrimage are profound, justifying the need for technological support. According to the Association of Japanese Animations, the industry market size reached a record 3.34 trillion yen in 2023 [4]. Furthermore, inbound tourism data indicate that approximately 1.41 million foreign visitors engage in anime pilgrimage, with a potential economic impact estimated at 480 billion yen [5,6].

A prime example of sustainable revitalization is Toyosato Town in Shiga Prefecture, the model location for *K-On!*. By transforming the former school building into a community hub, the town attracted approximately 300,000 visitors over five years and fostered a long-term relationship with fans [7]. However, maintaining such engagement requires continuous optimization of the visitor experience. As the number of tourists increases, issues such as overcrowding, seasonal mismatches, and lack of transportation information become barriers to satisfaction. This highlights the urgent need for intelligent systems that can effectively guide tourists.

1.3. Computational Challenges and Contribution

Despite the market’s growth, planning an anime pilgrimage remains a complex task. Unlike general sightseeing, it requires specific knowledge of “anime context,” such as the exact season in which a scene took place or the transportation route used by the characters. Existing tourism recommendation systems typically rely on single-source metrics, such as user review ratings or simple collaborative filtering. While effective for general points of interest, these approaches often fail to capture the multi-dimensional nature of pilgrimage tourism. For instance, a location might have high popularity on social media but be virtually inaccessible due to poor public transportation. Similarly, a site may be geographically accessible, yet lose its value if visited at the wrong time of year. For example, it would not make sense to visit a location famous for its winter scenes during the summer.

The core challenge lies in the mathematical handling of the trade-off between objective and subjective indicators. Transportation accessibility and seasonal alignment are objective constraints derived from geospatial and temporal data. In contrast, anime popularity is a subjective and dynamic metric driven by online fan communities. Previous studies have rarely addressed how to integrate these conflicting data types into a unified scoring model. To address this gap, we propose an intelligent recommendation framework that synthesizes three distinct data sources: route-based geospatial analysis for accessibility, image-based metadata for seasonal alignment, and cross-platform social data for popularity.

This paper is an extended version of our preliminary work presented at [8]. Although the previous study demonstrated the feasibility of the basic framework with a limited dataset, it relied on a single source for popularity analysis. Additionally, it did not rigorously validate the individual scoring components. In this study, we significantly advance the framework through the following contributions:

1. **Refining the Popularity Metric:** We introduce a robust multi-source aggregation method that combines data from global platforms to reduce bias and capture a more objective measure of popularity.

2. **Conducting an Ablation Study:** We perform a comprehensive ablation study to quantify the individual contribution of accessibility, seasonality, and popularity. It provides empirical evidence for the necessity of integrating these three factors.
3. **Expanding Experimental Validation:** We validate the model through an expanded user study involving a broader participant group compared to the preliminary work. It allows for a more reliable assessment of the alignment between system-generated rankings and user preferences.

The remainder of this paper is organized as follows: Section 2 reviews related work in anime tourism and data-driven recommendation systems. Section 3 outlines the proposed methodology, which includes the mathematical modeling of the three scoring factors. Section 4 presents the experimental evaluation and the quantitative results of the ablation study. Section 5 interprets these findings from a behavioral perspective and discusses the validity of the study. Finally, Section 6 concludes the paper and outlines future research directions.

2. Related Work

Recommendation systems have evolved from simple collaborative filtering to complex, intelligent data analysis frameworks. In the context of tourism and entertainment, the challenge lies in predicting user preferences and optimizing the entire experience through the processing of heterogeneous data. This data encompasses semantic content, user sentiment, and physical environmental constraints. This section reviews recent advancements in data-driven recommendations, highlighting the shift toward multi-criteria and context-aware methodologies.

2.1. Data Representation in Anime Recommendation

Modern anime recommendation engines rely heavily on advanced data representation techniques to overcome sparsity issues. Soni et al. [9] developed *RikoNet*, which tackles the high dimensionality of user-item interaction matrices. Instead of traditional matrix factorization, they utilized deep autoencoders to compress sparse interaction data into dense latent vectors, successfully capturing non-linear relationships between users and anime titles.

Similarly, focusing on the semantic structure of data, Javaji and Sarode [10] employed Graph Neural Networks (GNNs) to model the connectivity between users and content. By generating node embeddings that fuse structural graph data with semantic features extracted from synopses via BERT, their system demonstrates how natural language processing can enhance link prediction accuracy in cold-start scenarios.

2.2. General Tourism Recommendation and POI Mining

While this study specifically targets anime pilgrimage, it builds upon established frameworks in general tourism recommendation and Point-of-Interest (POI) mining. Early influential work in Location-Based Services (LBS) explored collaborative filtering approaches to recommend locations based on user check-in history and geographical influence [11,12].

Subsequent studies have integrated context-aware features to enhance recommendation precision. For instance, Yuan et al. [13] introduced time-aware POI recommendation models that account for temporal check-in patterns, acknowledging that user preferences fluctuate across different times of the day or seasons. Similarly, to address the data sparsity problem inherent in tourism recommendation, researchers have proposed geographical modeling and matrix factorization techniques [14], which effectively capture latent spatial user patterns.

Our proposed system adapts these foundational concepts—specifically spatial and temporal feasibility modeling—into the domain of content-induced tourism. Unlike generic POI recommenders that rely heavily on historical check-in data, our approach addresses the unique “content-orientation” of anime pilgrims by synthesizing IP-based relevance with logistical constraints inspired by traditional LBS frameworks.

2.3. Quantifying Sentiment for Tourism and Pilgrimage

Extracting quantitative signals from unstructured textual data is a key trend in intelligent tourism analysis. Alnogaithan et al. [15] proposed a method to convert qualitative user reviews into numerical sentiment scores. By aggregating these scores, they constructed a preference model that prioritizes emotional satisfaction, proving that mining subjective text data is essential for personalized travel suggestions.

In the specific domain of anime pilgrimage, Ma and Wang [16] applied sentiment analysis to bridge the gap between digital content and physical locations. They processed review data from both anime spots and nearby accommodations, integrating these sentiment metrics into a unified recommendation logic. This approach highlights the effectiveness of using “emotional data” as a weighting factor to align recommendations with user expectations.

2.4. Multi-Source Integration in Context-Aware Systems

To support real-world decision-making, recommendation systems must process dynamic environmental data. Kuang et al. [17] introduced a framework that fuses static POI data with dynamic variables such as real-time weather conditions and transportation schedules. Their work emphasizes the computational necessity of filtering candidates based on physical constraints before ranking them by preference.

Building on data fusion, Sha et al. [18] recently proposed a smart tourism platform that integrates heterogeneous data sources, including social media trends and geographic information, to optimize resource allocation. Their work validates that synthesizing multi-source data is crucial for capturing the complexity of modern tourism. Furthermore, addressing the temporal dimension, Sudarshan and Bakyalakshmi [19] highlighted the necessity of “Seasonal Alignment” in recommendation systems, showing that static POI data is insufficient for satisfying tourists’ dynamic needs. In the specific context of anime tourism, Zheng et al. [20] analyzed the behavior of pilgrims traveling to Japan, revealing that fans seek a “homologous emotion,” which is a deep alignment between anime scenes and the physical environment. This finding strongly supports our hypothesis that factors like visual seasonality and specific location context are as important as popularity.

2.5. Tourism Recommendation as Multi-Criteria Decision Making (MCDM)

From an analytical perspective, travel planning is inherently a Multi-Criteria Decision Making (MCDM) problem. Tourists must trade off conflicting objectives, such as maximizing site popularity while minimizing travel time and cost. Recent literature in intelligent data analysis suggests that treating recommendations as an optimization task using weighted scoring models allows for transparent results compared to black-box machine learning approaches. For instance, Putra and Priatdana [21] applied the TOPSIS method to rank tourist destinations based on multiple attributes, such as accessibility and cost.

Similarly, Buasri and Sangpradid [22] proposed a fuzzy TOPSIS framework that integrates subjective tourist preferences with objective destination attributes. This framework demonstrates that explicitly modeling data uncertainty and trade-offs greatly improves the reliability of recommendations.

2.6. Research Gap and Approach

Despite these advancements, a gap remains in effectively bridging the “virtual” and “physical” worlds for niche tourism, such as anime pilgrimages. Most existing studies focus either on content similarity or subjective sentiment. Relatively few studies have attempted to mathematically integrate *physical constraints* (transportation networks), *temporal alignment* (seasonality), and *social metrics* (popularity) into a single optimized framework.

Furthermore, rigorous validation of such multi-factor models is notably absent in this domain. Most tourism recommendation studies rely solely on user satisfaction surveys without quantifying the contribution of individual data sources. This study addresses these gaps by proposing a context-aware scoring model. We distinguish our work by conducting a comprehensive ablation analysis, a method standard in intelligent data analysis, to empirically prove the value of each integrated dimension.

3. Design of Recommendation Method

This section presents a multidimensional recommendation framework designed for anime pilgrimage scenarios. To address the complexity of tourism decision-making, we adopt a decision-level multi-source data fusion strategy. Specifically, the proposed system integrates heterogeneous data streams, including geospatial trajectory data from navigation APIs, cross-platform social sentiment from global anime communities, and temporal environmental context, into a unified weighted scoring model.

As illustrated in Figure 1, this architecture fuses three core dimensions: transportation accessibility, seasonal alignment, and anime popularity. By synthesizing these diverse signals, the model generates context-aware recommendations that balance subjective user preferences with objective real-world travel constraints.

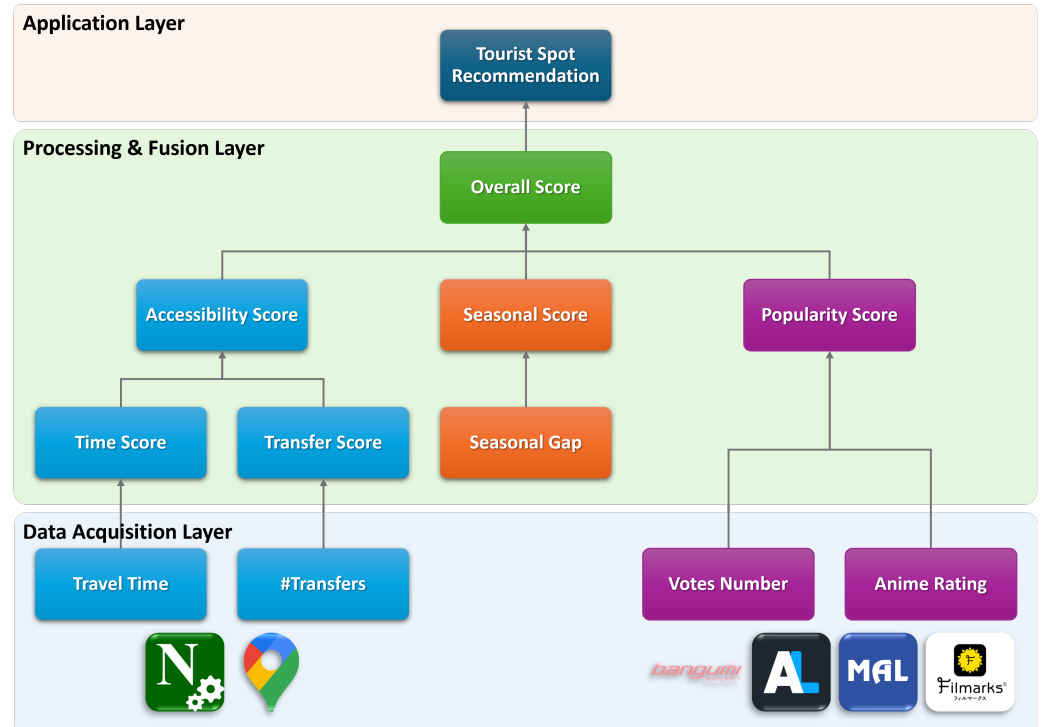


Figure 1. The proposed multi-source data fusion framework architecture. The system consists of three layers: the Data Acquisition Layer for collecting heterogeneous information, the Processing & Fusion Layer for calculating the Cross-Platform Popularity Index (CPPI) and context-aware scores, and the Application Layer for visualizing recommendations to the user.

The system is built on multi-source data fusion and hierarchical quantitative modeling. The recommendation process is divided into three primary feature modules:

1. **Accessibility Score Module:** This module consists of two sub-indicators, *Travel Time* and *Number of Transfers*. These indicators are quantified using a time-score model and a transfer-score model to evaluate the logistical convenience of traveling to each pilgrimage spot.
2. **Seasonal Score Module:** This module is constructed based on the *Seasonal Gap*, which measures the degree of alignment between the current real-world season and the season depicted in the anime scene.
3. **Popularity Score Module:** This module introduces the Cross-Platform Popularity Index (CPPI). It integrates vote counts and aggregated ratings from multiple global platforms to produce a comprehensive indicator of anime popularity.

The system employs a weighted scoring model to calculate an overall score based on these three modules. The following sections provide detailed explanations of the quantification logic for each module and the optimization strategy for the integration weights.

To ensure system efficiency and minimize latency during user interactions, the proposed framework operates on a hybrid data processing model. Instead of invoking real-time API calls for every user request, which could result in significant latency due to network overhead and rate limits, the system uses an offline batch processing strategy. Popularity metrics, specifically the Combined Popularity and Preference Index (CPPI), are collected weekly through Python (version 3.11.5) scripts that interact with platform APIs, such as the Jikan API for MyAnimeList, since aggregate anime ratings tend to remain stable over short periods. Simultaneously, Accessibility and Seasonality data are pre-calculated and stored in a local relational database. By decoupling data collection from user inference, the recommendation engine can query the pre-indexed database to generate rankings in milliseconds ($O(N)$ complexity), ensuring a responsive user experience suitable for real-world web applications.

3.1. Transportation Accessibility Scoring

To evaluate the ease of reaching a pilgrimage site, we compute an *Accessibility Score* using parameters extracted from public transportation route data.

3.1.1. Route Data Collection

We utilize the NAVITIME Route API and Google Maps Geocoding API to extract optimal routes between a user-defined starting point and anime-related locations. The NAVITIME Route API provides door-to-door travel routes within Japan, including detailed information such as optimal station exits and transfer convenience [23]. In instances where an anime location lacks a precise address, the Google Maps Geocoding API converts the location name into latitude and longitude coordinates, which are then processed by the routing engine [24].

3.1.2. Definition of Accessibility Indicators

Two scores are defined to quantify accessibility:

- **Time Score:**

$$S_{\text{time}} = \max\left(0, 10 - \frac{T}{30}\right) \quad (1)$$

where T represents the required travel time in minutes. A score of 10 corresponds to 0 min, and the score decreases as travel time increases. S_{time} is floored at 0 to prevent negative scoring for extreme travel times. This ensures that any travel time exceeding 300 min (5 h) results in a score of zero, rather than a penalty, thereby preserving the non-negativity required for the subsequent weighted aggregation. In our dataset, all candidate routes are within 180 min, but this formulation ensures robustness for generalized cases.

- **Transfer Score:**

$$S_{\text{transfer}} = 10 - 2N \quad (2)$$

where N is the number of transfers. The transfer penalty term is designed to capture not only additional waiting time but also the generalized disutility associated with transfer-related inconvenience, including walking effort, uncertainty, and cognitive load. Prior studies in travel behavior have consistently shown that transfers impose a disproportionate penalty compared to in-vehicle travel time, often equivalent to a substantial amount of perceived travel time rather than a simple linear addition of minutes [25,26]. Empirical evidence further suggests that this disutility is exacerbated by psychological and environmental factors, such as station complexity and walking conditions [27]. Consequently, in our model, the transfer penalty is heuristically set to 2.0. This coefficient was selected to maintain scale consistency with the time-based score while serving as a coarse proxy for the elevated cognitive burden and uncertainty faced by pilgrims navigating unfamiliar transit systems.

3.1.3. Accessibility Score Calculation

Finally, the **Accessibility Score** ($Score_{\text{acc}}$) is calculated as the average of the time and transfer metrics using Equations (1) and (2):

$$Score_{\text{acc}} = \frac{S_{\text{time}} + S_{\text{transfer}}}{2} \quad (3)$$

3.2. Seasonal Alignment Scoring

To enhance the immersive experience, we assess the “Seasonal Alignment,” which quantifies how well the current visiting season matches the season depicted in the anime scene.

3.2.1. Seasonal Definitions and Sequence

To mathematically handle seasonal transitions, we divide the year into four seasons and assign integer codes representing their chronological sequence:

- Spring (March–May): 0
- Summer (June–August): 1
- Autumn (September–November): 2
- Winter (December–February): 3

The seasons follow a strict cyclic progression from Spring (0) to Summer (1), Autumn (2), Winter (3), and back to Spring (0). This directional flow is critical for calculating the temporal wait time between the current season and the target anime season. Figure 2 illustrates this cyclic transition.

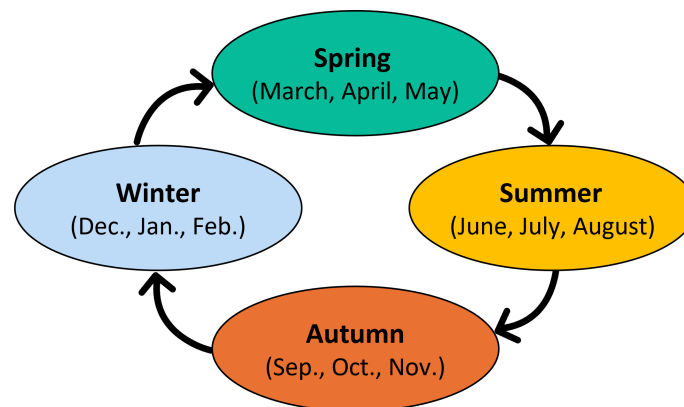


Figure 2. Diagram of seasonal progression and corresponding months.

3.2.2. Seasonal Score Calculation

The seasonal gap ΔS is defined as the forward temporal distance from the current season (S_c) to the anime's season (S_a). This metric quantifies the “waiting time” required to reach the target season within the annual cycle. The calculation is performed as follows:

$$\Delta S = \begin{cases} |S_c - S_a| & \text{if } S_c \leq S_a \\ 4 - |S_c - S_a| & \text{if } S_c > S_a \end{cases} \quad (4)$$

This formulation reflects the asymmetric nature of seasonal anticipation. For example, if the current season is Winter ($S_c = 3$) and the anime depicts Spring ($S_a = 0$), the gap is $\Delta S = 1$, reflecting that Spring is imminent. Conversely, if the current season is Spring ($S_c = 0$) and the anime depicts Winter ($S_a = 3$), the gap is $\Delta S = 3$, indicating that one must wait through Summer and Autumn to reach Winter.

Based on ΔS , the **Seasonal Score** ($Score_{sea}$) is assigned to prioritize seasons that are currently accessible or arriving soon (Table 1).

Table 1. Seasonal alignment scoring based on forward temporal distance.

Seasonal Gap ΔS	Description	Seasonal Score
0	Exact Match (e.g., Spring $\rightarrow$ Spring)	1.00
1	Upcoming Season (e.g., Winter $\rightarrow$ Spring)	0.75
2	Two-Season Wait (e.g., Spring $\rightarrow$ Autumn)	0.50
3	Distant Season (e.g., Spring $\rightarrow$ Winter)	0.25

If the anime season is undefined, a default neutral score of 0.625 is applied. The scoring intervals (1.00, 0.75, 0.50, 0.25) are designed as a quartile-based linear decay representing the temporal proximity of the target season, where each additional season of waiting introduces a uniform accessibility cost. We intentionally enforce a non-zero lower bound (0.25) rather than assigning a normalized zero score, as an anime location retains baseline pilgrimage relevance even when the depicted season is temporally distant. Unlike symmetric visual similarity models, this asymmetric scoring assigns higher values to an “upcoming” season (Gap 1) than to a “just passed” season (Gap 3). This design choice explicitly prioritizes forward temporal reachability for pilgrimage planning over simple visual matching.

3.3. Anime Popularity Scoring: Cross-Platform Popularity Index

A significant improvement in this study is the introduction of the Cross-Platform Popularity Index (CPPI). Unlike preliminary works that relied on single-source data, CPPI

aggregates voting data from four major global platforms (Bangumi, AniList, MyAnimeList, and Filmarks) to mitigate bias and capture global trends.

3.3.1. Data Sources and Characteristics

To construct a globally representative popularity index, we collected data from four major platforms, each representing distinct demographic and cultural user bases.

- **Bangumi (Chinese Demographic):** Established in 2008, Bangumi is a comprehensive ACGN database targeting the Chinese-speaking audience. It is characterized by high data openness and a user culture emphasizing critical evaluation. With approximately 760,000 users and over 7.7 million rating entries, it employs a 1–10 scoring system. Data from Bangumi is crucial for capturing the preferences of East Asian fans outside Japan [28,29].
- **AniList (Western Demographic):** AniList is a modern platform popular among younger audiences in Europe and North America. It features robust social networking functions and a developer-friendly API. The platform hosts metadata for approximately 19,628 titles and facilitates active discourse. Including AniList allows our model to reflect the trends of the emerging generation of international anime consumers [30].
- **MyAnimeList (Global Standard):** Established in 2004, MyAnimeList (MAL) is the world's largest English-language anime database. As of 2025, it boasts over 18 million registered users from more than 240 regions, with 99% residing outside Japan. Its massive catalogue (approx. 26,417 titles) and user base provide the statistical backbone for global popularity trends. Like Bangumi, it uses a 1–10 rating scale [31].
- **Filmarks (Domestic Japanese Demographic):** Filmarks is Japan's largest review platform for films and anime, operated by Tsumiki Inc. It contains over 7.5 million reviews for 6495 anime titles. Unlike the other three platforms, Filmarks uses a 5-point rating system and reflects the domestic reception of works within Japan. This local perspective is essential for balancing international hype with local cultural context [32].

Since Filmarks uses a 5-point scale while others use a 10-point scale, raw ratings from Filmarks are multiplied by 2 during the normalization process to ensure scale consistency.

3.3.2. Score Normalization

For each anime title, we utilize the vote counts and rating scores collected from the platforms. Since the raw values vary significantly in scale, we apply Min-Max normalization to map them into the range $[0, 1]$.

Let N_i denotes the total number of votes and S_i denotes the aggregated rating score for anime i . The normalized values $\tilde{N}_i$ and $\tilde{S}_i$ are computed as

$$\tilde{N}_i = \frac{N_i - N_{\min}}{N_{\max} - N_{\min}} \quad (5)$$

$$\tilde{S}_i = \frac{S_i - S_{\min}}{S_{\max} - S_{\min}} \quad (6)$$

where $N_{\min}$ and $N_{\max}$ in Equation (5) and $S_{\min}$ and $S_{\max}$ in Equation (6) represent the minimum and maximum values in the dataset, respectively. This linear transformation ensures that the popularity metrics are comparable with the accessibility and seasonality scores.

3.3.3. Popularity Score Calculation

The final **Popularity Score** ($Score_{\text{pop}}$) for anime i combines the volume of engagement (votes) and the quality of reception (ratings):

$$Score_{\text{pop}} = 0.7 \cdot \tilde{N}_i + 0.3 \cdot \tilde{S}_i \quad (7)$$

where $\tilde{N}_i$ denotes the normalized *vote count score* by Equation (5) and $\tilde{S}_i$ denotes the normalized *anime rating score* by Equation (6). We prioritize vote volume (0.7) because it is a better indicator of “cultural buzz” and the likelihood that a location will become a pilgrimage site than critical acclaim alone.

3.4. Overall Recommendation Scoring

Following the independent quantification of each module, the *Accessibility Score*, *Seasonal Score*, and *Popularity Score* must be integrated into a single composite indicator. In this study, we employ a Linear Weighted Combination (LWC) approach. This method is selected for its compensatory nature, allowing a location with lower accessibility to be recommended if it has exceptionally high popularity or ideal seasonal alignment. This characteristic is essential for “pilgrimage” behavior, where emotional motivation often takes precedence over logistical challenges.

3.4.1. Score Normalization

Since the raw outputs of the three modules possess different scales and physical units, normalization is a prerequisite for aggregation.

- **Accessibility Score ($Score_{acc}$):** Derived from travel time and transfers, already scaled to $[0, 10]$. We normalize it to $[0, 1]$ by dividing by 10.
- **Seasonal Score ($Score_{sea}$):** As defined in Table 1, the values inherently lie within $[0.25, 1.0]$. No further transformation is applied to preserve the penalty for seasonal mismatch.
- **Popularity Score ($Score_{pop}$):** As calculated in Section 3.3, this score is already normalized to $[0, 1]$ via Min-Max scaling.

3.4.2. Weighted Aggregation Model

Let the normalized feature vector for a candidate location i be $\mathbf{v}_i = [Score_{acc}, Score_{sea}, Score_{pop}]$. The final **Overall Recommendation Score** ($Score_{final}$) is computed as the dot product of the feature vector and a weight vector $\mathbf{w} = [\alpha, \beta, \gamma]$:

$$Score_{final} = \alpha \cdot Score_{acc} + \beta \cdot Score_{sea} + \gamma \cdot Score_{pop} \quad (8)$$

subject to the constraint:

$$\alpha + \beta + \gamma = 1.0, \quad \alpha, \beta, \gamma \geq 0$$

3.4.3. Parameter Determination

The determination of weight parameters is critical for balancing the trade-offs between heterogeneous features. In this study, the weights were determined to reflect the hierarchical nature of anime pilgrimage: prioritizing user interest (Popularity) as the primary motivator, while ensuring logistical feasibility (Accessibility) and treating seasonality as a supplementary experiential enhancer.

Accordingly, the weights are configured as follows:

$$\alpha = 0.3, \quad \beta = 0.2, \quad \gamma = 0.5$$

The rationale for this configuration is based on the distinct roles of each factor:

- $\gamma = 0.5$ (Popularity–Primary Motivator): Assigned the highest weight. It reflects the premise that the fundamental motivation for pilgrimage is the user’s affinity for the anime work itself. If the anime is not interesting (low popularity), the location holds no value as a pilgrimage site, regardless of its accessibility.

- $\alpha = 0.3$ (Accessibility–Logistical Constraint): Assigned a moderate weight. While fans are willing to travel, extremely poor accessibility acts as a physical barrier. This weight prioritizes accessible locations over those that are difficult to reach.
- $\beta = 0.2$ (Seasonality–Experiential Enhancer): Assigned the lowest weight. Seasonal alignment enhances the immersive experience, such as seeing snow scenes in winter, but it functions as a “soft” constraint. Fans often visit a site out of season solely to see the location, making it less critical than the other two factors.

Finally, to ensure this configuration is data-driven rather than purely heuristic, we conduct a detailed sensitivity analysis in Section 4.5. The results confirm that this specific setting ($\gamma = 0.5$) represents the optimal equilibrium point, effectively balancing ranking accuracy with logistical feasibility.

4. Experimental Evaluation

In this section, we evaluate the proposed anime pilgrimage recommendation system through a rigorous user study. The evaluation focuses on two key aspects: (1) the alignment of the system’s ranking with human preferences compared to single-metric baselines, and (2) the contribution of each scoring component validated through an ablation study.

4.1. Experimental Setup

We conducted a user study with 20 university students (12 Chinese and 8 Japanese nationals) who have a demonstrated interest in anime culture. To assess potential demographic bias, we conducted an exploratory comparison of preference ranking patterns between the two nationality groups. No salient or systematic differences were observed in the overall ranking tendencies, suggesting that participants’ shared interest in the anime IP may have played a more dominant role than nationality in shaping preferences. Given the limited subgroup sizes, no formal statistical tests were conducted for cross-national differences. To simulate a realistic pilgrimage planning scenario, we designated Tokyo Station as the starting point for all travel routes.

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. Although formal ethics committee approval was not necessary for this non-clinical, questionnaire-based survey, informed consent was obtained from all participants before the study began. Participants were fully informed about the purpose of the survey, and their anonymity and confidentiality were strictly assured. Participation was entirely voluntary.

We selected 13 representative anime titles set in Tokyo and its vicinity based on the top 200 most-streamed anime rankings (2012–2024) provided by Docomo [33]. Participants were presented with the following detailed information for each title to aid their decision-making:

- **Location Context:** The specific real-world location associated with the anime title.
- **Accessibility Context:** Estimated travel time and public transportation route details starting from Tokyo Station.
- **Content Context:** A brief synopsis (or link to Wikipedia) and aggregated popularity metrics, including rating scores and vote counts from the four major platforms defined in Section 3.3.
- **Spatial Context:** A map illustrating the relative geographical positions of the pilgrimage sites (see Figure 3).

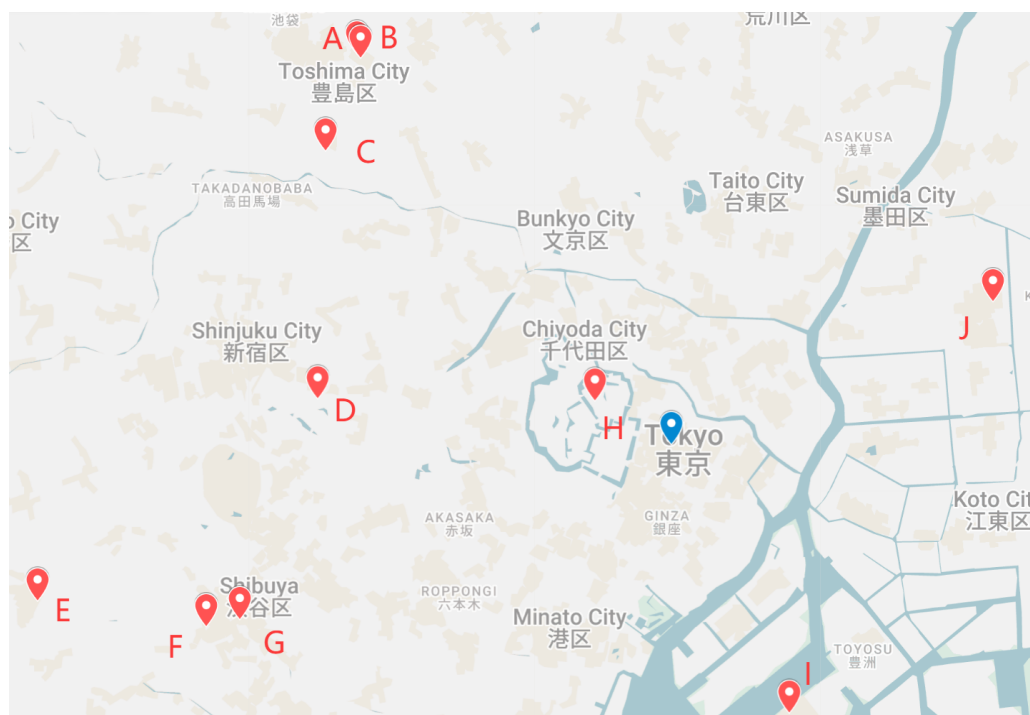


Figure 3. Relative locations of anime pilgrimage sites from Tokyo Station.

Participants reviewed these materials and were instructed to rank their top 10 preferred sites based on their personal interest and travel feasibility. To evaluate alignment, we compared these subjective “ground-truth” rankings with the system-generated rankings.

4.2. Anime Titles and Pilgrimage Sites

The 13 anime titles and their corresponding real-world pilgrimage sites used in the evaluation are listed in Table 2.

Table 2. Anime titles and corresponding pilgrimage sites used in the experiment.

ID	Anime Title	Pilgrimage Site
A	<i>BanG Dream! It's MyGO!!!!</i>	Staircase in front of Sunshine City
B	<i>My Dress-Up Darling</i>	Sunshine Plaza, Sunshine Garden
C	<i>Saekano: How to Raise a Boring Girlfriend</i>	Nozoki Hill
D	<i>The Garden of Words</i>	Shinjuku Gyoen National Garden (Japanese Garden)
E	<i>Bocchi the Rock!</i>	Shimokitazawa Shelter
F	<i>Ya Boy Kongming!</i>	ATOM TOKYO
G	<i>Jujutsu Kaisen</i>	Hachiko Statue
H	<i>Sword Art Online II</i>	Imperial Palace (Three Benches, Inner Garden)
I	<i>Oshi no Ko (Season 2)</i>	IHI Stage Around Tokyo
J	<i>Lycoris Recoil</i>	Kinshi Park
K	<i>Slam Dunk</i>	Kamakura High School Railroad Crossing
L	<i>Rascal Does Not Dream of Bunny Girl Senpai</i>	Shichirigahama Beach
M	<i>Miss Kobayashi's Dragon Maid</i>	East Exit of Koshigaya Station

4.3. Performance Comparison with Baselines

To validate the effectiveness of the proposed integrated framework, we compared our method with two single-metric baselines that represent distinct user behavioral heuristics.

This comparison aims to demonstrate how a multi-factor approach manages the trade-offs between conflicting user needs.

- **Baseline 1 (Accessibility Only):** Rankings generated solely based on transportation convenience ($\alpha = 1.0, \beta = 0, \gamma = 0$). This simulates a user prioritizing logistical ease, choosing the nearest locations regardless of the anime content.
- **Baseline 2 (Popularity Only):** Rankings generated solely based on the Cross-Platform Popularity Index ($\gamma = 1.0, \alpha = 0, \beta = 0$). This simulates a user deciding where to go primarily based on the fame of the anime work, ignoring distance and season.
- **Proposed Method (Integrated):** The weighted scoring model integrating Accessibility, Seasonality, and Popularity ($\alpha = 0.3, \beta = 0.2, \gamma = 0.5$).

The performance of each method was evaluated using Mean Reciprocal Rank (MRR), Mean Average Precision (MAP), and Normalized Discounted Cumulative Gain at rank 5 (NDCG@5). Table 3 presents the comparative results.

Table 3. Performance comparison between baselines and the proposed method. The asterisk (*) indicates a statistically significant difference ($p < 0.01$) compared to the strongest Baseline 2 (Popularity Only) based on the Wilcoxon Signed-Rank Test. Bold values indicate the best performance for each metric.

Method	MRR	MAP	NDCG@5
Baseline 1 (Accessibility Only)	0.237	0.824	0.827
Baseline 2 (Popularity Only)	0.424	0.834	0.794
Proposed Method ($Score_{final}$)	0.397	0.829	0.814 *

The results in Table 3 reveal a critical trade-off between logistical feasibility and content interest, illustrating the strength of the proposed integrated approach as a balanced solution.

Baseline 2 (Popularity Only) achieved the highest MRR (0.424) and MAP (0.834). It confirms that the primary motivation for pilgrimage is indeed the content; users' "number one choice" is almost always a popular site. However, its NDCG@5 dropped to 0.794, the lowest among the three. It suggests that relying solely on popularity may lead users to consider locations that are logistically impractical for a top-5 itinerary, potentially reducing satisfaction in real-world planning.

Conversely, Baseline 1 (Accessibility Only) achieved the highest NDCG@5 (0.827) but a critically low MRR (0.237). It suggests that while users value convenience when filling out their top-5 list, prioritizing distance alone fails to capture the users' favourite destinations (the "top 1"), resulting in a recommendation list that is "easy to go to" but lacks "emotional value".

The Proposed Method successfully bridges this gap. It maintains a high MRR (0.397) and MAP (0.829), comparable to the Popularity baseline, while achieving a strong NDCG@5 (0.814), significantly better than the Popularity baseline. Unlike single-metric approaches that have severe limitations, such as low NDCG for popularity and low MRR for accessibility, the integrated model offers a strong and practical compromise. It ensures that recommendations are not only emotionally appealing but also logistically feasible.

Regarding **Statistical Significance and Practical Implications**, we first conducted a Wilcoxon Signed-Rank Test ($N = 20$) to ensure that the performance improvements shown in Table 3 are not due to random variance. The analysis confirms that the Proposed Method significantly outperforms the strongest content competitor, Baseline 2 (Popularity Only), with a p -value of ≈ 0.0094 ($p < 0.01$).

Given the limited sample size ($N = 20$), the statistical tests in this study are intended to provide supportive evidence rather than definitive proof of population-level effects. The re-

ported p -values should therefore be interpreted with caution, particularly for comparisons near conventional significance thresholds.

Beyond statistical evidence, we also consider the practical value of these metrics. While a marginal improvement in NDCG (e.g., +0.02) may appear modest numerically, it translates to a tangible enhancement in user experience. In top- k recommendations, this improvement often indicates that highly relevant and accessible locations are successfully promoted to the top positions (e.g., moving from Rank 3 to Rank 1). For users, this minimizes the cognitive load of filtering out “popular but unreachable” destinations, thereby allowing them to discover feasible pilgrimage plans more efficiently.

4.4. Ablation Study Results

To rigorously quantify the contribution of each factor, we conducted an ablation study by systematically removing components from the full model. The quantitative results are presented in Table 4 and visualized in Figure 4.

Table 4. Ablation study results for different score combinations. The asterisk (*) indicates a statistically significant difference in ranking distributions compared to Model C (w/o Seasonality) based on the Wilcoxon Signed-Rank Test ($p < 0.05$). Statistical significance reflects differences in ranking behavior, rather than superiority in rank-based accuracy metrics.

Condition	Ranking (ID Order)	MRR	MAP	NDCG@5
Model A (Full: All Factors)	E, G, M, B, A, D, J, I, L, H, C, F, K	0.397	0.829	0.814 *
Model B (w/o Accessibility)	E, L, G, M, D, B, J, A, I, C, K, F, H	0.400	0.819	0.784
Model C (w/o Seasonality)	E, G, B, A, M, D, L, J, C, I, H, F, K	0.403	0.855	0.836
Model D (w/o Popularity)	H, G, E, M, A, B, I, J, F, C, D, K, L	0.250	0.807	0.751

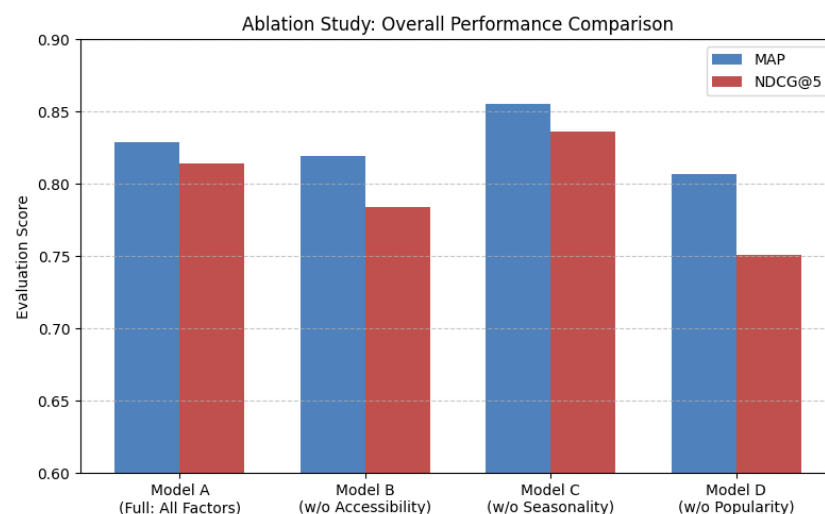


Figure 4. Ablation Study: Overall Performance Comparison. The chart displays MAP and NDCG@5 scores across four model variants. Note the drop in performance for Model D and Model B compared to Model C in terms of MAP and NDCG@5.

4.4.1. Impact of Popularity and Accessibility

As clearly illustrated in Figure 4, the performance varies significantly across different model variants.

- **Impact of Popularity (Model D):** The rightmost bars in Figure 4 show the lowest performance across all metrics (NDCG@5: 0.751). This visual drop confirms that popularity acts as the foundational filter for user interest.
- **Impact of Accessibility (Model B):** Comparing Model B with the Full Model (Model A), we observe a noticeable decline in the NDCG score (red bar), indicating that ignoring logistical constraints negatively impacts the quality of the top-5 recommendations.

4.4.2. Analysis of Seasonality Impact (Model C)

Interestingly, Model C (w/o Seasonality) achieved the highest metrics (MAP: 0.855). We attribute this phenomenon to the discrepancy between the user evaluation task and the model's design objective.

In the user study, the ranking task primarily elicited “static” content preferences—ranking locations based on their interest in the anime works and the intrinsic fame of the sites—without explicitly considering current temporal constraints. Conversely, the Full Model incorporates “dynamic” feasibility, penalizing locations with high seasonal gaps to ensure realistic travel planning. When the model downgrades a popular but off-season location, it diverges from the user's static preference ranking, resulting in a mathematical decrease in MAP.

Therefore, the lower MAP in the Full Model reflects a necessary trade-off between maximizing immediate user interest match (Static Relevance) and ensuring actionable travel timing (Dynamic Feasibility). This observation does not imply that excluding seasonality yields a superior recommendation model, but rather highlights a mismatch between static relevance-oriented evaluation metrics and planning-oriented model objectives.

It should be noted that the terms “feasibility” and “actionability” in this context refer strictly to model-imposed temporal constraints reflecting seasonal alignment, rather than empirically validated user decision outcomes or actual travel behavior.

4.4.3. Statistical Validation of Components

We further applied the Wilcoxon Signed-Rank Test to the ablation results in Table 4 to verify the contribution of individual modules. The comparison between the Full Model (Model A) and the strongest ablation variant, Model C (w/o Seasonality), reveals a statistically significant difference (Wilcoxon Signed-Rank Test, $p \approx 0.041$, $p < 0.05$). Although Model C achieved slightly higher raw metrics, the statistically significant difference indicates that the Seasonality module induces systematic changes in ranking behavior by filtering out temporally misaligned (but popular) options, rather than acting as a passive component.

4.5. Parameter Sensitivity Analysis

To justify the selection of weights ($\alpha = 0.3$, $\beta = 0.2$, $\gamma = 0.5$) used in our proposed method, we conducted a sensitivity analysis to observe how variations in the importance of “Anime Popularity” (γ) affect system performance. Since popularity is the dominant factor in user attention, we varied γ from 0.1 to 0.9 in increments of 0.2. To satisfy the constraint $\alpha + \beta + \gamma = 1.0$, the remaining weight ($1 - \gamma$) was distributed to Accessibility (α) and Seasonality (β) while maintaining the ratio of Accessibility to Seasonality constant at $\alpha : \beta = 3 : 2$. This experimental design isolates the effects of the trade-off between “content power” (Popularity) and “contextual constraints” (Accessibility and Seasonality).

Figure 5 illustrates the MAP and NDCG@5 scores across different γ values. The results reveal a clear trend: ranking performance improves significantly as γ increases from 0.1 to 0.5, reflecting the strong correlation between anime popularity and user interest. However, beyond $\gamma = 0.5$, the performance gains exhibit diminishing returns.

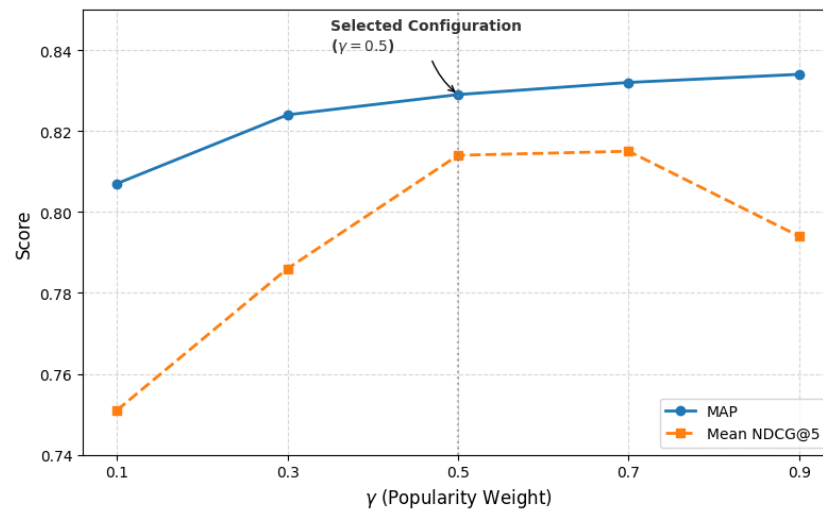


Figure 5. Sensitivity analysis of the Popularity weight (γ) against performance metrics (MAP and NDCG@5). The performance plateaus after $\gamma = 0.5$, indicating that further increasing popularity weight yields negligible gains while severely compromising accessibility constraints.

Specifically, comparing the chosen configuration ($\gamma = 0.5$) with the numerical peak ($\gamma = 0.7$):

- **Marginal Performance Gain:** Increasing γ to 0.7 yields only a negligible improvement in MAP (0.829 $\rightarrow$ 0.832) and NDCG@5 (0.814 $\rightarrow$ 0.815).
- **Significant Utility Loss:** However, this shift reduces the Accessibility weight (α) drastically from 0.30 to 0.18.

Notably, at $\gamma = 0.9$, while MAP increases slightly, NDCG@5 actually drops (0.815 $\rightarrow$ 0.794), suggesting that excessive popularity weighting harms the precision of top-ranked recommendations. This indicates that overriding spatial and temporal constraints with extreme popularity weights introduces noise into the top-tier results.

For a travel planning system, accessibility is a hard constraint—a location that is too difficult to reach offers zero utility, regardless of its popularity. Therefore, considering both the “diminishing returns” in accuracy and the degradation of top-rank precision at extremes, we identify $\gamma = 0.5$ as the optimal equilibrium point. It achieves robust ranking accuracy without sacrificing the physical feasibility of the recommended itinerary.

5. Discussion

In this section, we interpret the experimental findings from the perspective of intelligent data analysis, focusing on the hierarchical structure of user motivation and the trade-off between ranking accuracy and experience quality.

5.1. Trade-Off Analysis: Why Balance Matters

The comparison with baselines (Table 3) reveals the inherent limitations of single-metric recommendation strategies and the robustness of our integrated approach.

5.1.1. The Trap of Single Metrics

- Baseline 1 (Accessibility Only) achieved the highest NDCG@5 (0.827) but a critically low MRR (0.237). It suggests that while proximity makes for an “easy” Top-5 list, it

fails to capture the user's core motivation ("I want to see this anime"), resulting in a recommendation list that is logically convenient but emotionally irrelevant.

- Baseline 2 (Popularity Only) achieved the highest MRR (0.424) and MAP (0.834). It confirms that a user's "top choice" is almost always driven by content interest. However, its lower NDCG@5 (0.794) indicates that a purely popularity-driven list often includes distant locations that are impractical for a short itinerary, potentially forcing users to discard recommendations during actual planning.

5.1.2. The Robustness of the Integrated Model

The Proposed Method ($Score_{final}$) successfully bridges this gap. While it slightly trails Baseline 2 in MAP (0.829 vs. 0.834), it significantly improves upon the logistical feasibility (NDCG: 0.814 vs. 0.794). Unlike the baselines, which suffer from extreme weaknesses (e.g., Baseline 1's low MRR or Baseline 2's lower feasibility), the integrated model provides a robust compromise. It ensures that the recommended sites are not only popular enough to motivate a visit but also reachable enough to be included in a realistic travel plan.

5.2. The Seasonality Paradox: Planning vs. Experiencing

A counter-intuitive finding was that Model C (w/o Seasonality) achieved slightly higher ranking metrics than the Full Model (Model A: Proposed Method). As hypothesized in Section 4, this reflects the discrepancy between "Planning" and "Experiencing".

During the planning phase, as simulated in our questionnaire, users prioritize *feasibility* (Can I go?) and *motivation* (Do I want to go?) over *timing* (Is it the right season?). Consequently, seasonality adds noise to the prediction of user choices. However, for a recommendation system to be truly "context-aware," it must look beyond simple click-through rates. Recommending a winter scene in summer may satisfy a user's immediate curiosity (high ranking accuracy) but lead to a disappointing on-site experience (low satisfaction). Therefore, we argue that the Seasonal Score serves as a quality assurance mechanism. Although it does not immediately boost ranking metrics, it enhances the long-term value of the pilgrimage as a "soft constraint".

5.3. Social and Ethical Implications

Beyond technical performance metrics, the proposed recommendation framework holds significant implications for the sustainability of tourism and the understanding of user behavior.

Regarding the **Mitigation of Overtourism**, popular anime pilgrimage sites often face the challenge of "overtourism," where excessive visitor numbers disrupt local ecosystems. By integrating "Seasonality" and "Accessibility" alongside popularity, our system introduces a mechanism to potentially redistribute tourist traffic. For example, by identifying off-peak seasons (via Seasonality) or recommending accessible but less-visited locations (high Accessibility, moderate Popularity), the framework can guide tourists toward alternative options. This helps balance visitor flows, reducing the pressure on hotspots while promoting hidden gems.

From a sociological perspective, the "Popularity" module functions as more than just a metric of fame; it serves as a proxy for **Social Influence and Conformity**. In the context of technology adoption, users are more likely to accept recommendations that align with their lifestyle constraints (e.g., time availability and physical capability). Our system bridges the gap between the desire for social conformity (sharing collective memories) and logistical feasibility. However, it is essential to balance this motivation with ethical considerations. Since many anime scenes are inspired by actual residential areas, future iterations of the system should consider incorporating "residential privacy" constraints (e.g., flagging quiet zones) to ensure that tourism promotion does not infringe upon the daily lives of residents.

5.4. Limitations and Future Work

While the proposed framework demonstrates effective multi-source fusion, we acknowledge several limitations that define the scope for future research.

5.4.1. Data Reliability and Standardization Constraints

We acknowledge two primary limitations in the current data processing pipeline regarding the Cross-Platform Popularity Index (CPPI).

First, regarding **Platform Bias**, the current model employs Min-Max scaling to map diverse metric ranges into a uniform $[0, 1]$ interval. Although this method is effective for ensuring positive inputs for the weighted combination model, it is sensitive to outliers, and it does not account for varying user base distributions across different platforms (e.g., MyAnimeList vs. Bangumi). Future iterations should use Z-score standardization or robust scaling to correct for platform-specific biases and more accurately align statistical distributions.

Second, regarding **Data Reliability**, the aggregation of raw vote counts assumes that the source data is organic. Although the proposed Cross-Platform Popularity Index (CPPI) inherently mitigates single-platform anomalies (e.g., a specific site being targeted) by averaging data from four distinct user bases, it remains vulnerable to coordinated external manipulation, such as widespread “review bombing” or automated bot activity. While the current system relies on the moderation mechanisms of the source platforms, this is not a failsafe solution. Future work will necessitate the integration of an outlier detection module to identify and downweight suspicious spikes in voting patterns, ensuring that the popularity metric reflects genuine user consensus.

5.4.2. Constraints of the Mathematical Models

We acknowledge three primary mathematical limitations in the current framework components:

First, for the transportation accessibility module of the **Accessibility model**, we used a linear penalty function for both travel time (Equation (1)) and the number of transfers (Equation (2)). We explicitly acknowledge that this is a simplified first-order approximation. In real-world scenarios, travel impedance often exhibits nonlinear characteristics. For instance, the perceived physical and cognitive fatigue from making three transfers is likely much greater than the cumulative effect of three single transfers. Future iterations will use nonlinear scoring functions, such as exponential decay or logarithmic scaling, to more accurately simulate psychological travel effort.

Second, regarding the **Seasonality model**, while Section 5.2 highlights the value of seasonality as a “soft constraint,” our current implementation relies on a discrete four-season model. However, this discrete approach may not fully capture transitional periods (edge cases). For example, early March and late May are both categorized as “Spring,” yet they offer vastly different environmental experiences. To address this, future work will implement a continuous temporal model (e.g., monthly weights) and integrate real-time environmental data (e.g., temperature or vegetation indices) to handle edge cases more precisely.

Finally, concerning the **Linear Weighted Combination (LWC) model**, a significant limitation is the reliance on static weights. The current framework applies a universal weight configuration ($\alpha = 0.3, \beta = 0.2, \gamma = 0.5$) to all users, effectively treating the user base as a homogeneous group. This static approach fails to capture user heterogeneity. For example, some users prioritize convenience over popularity, while dedicated fans may endure long travel times to visit niche destinations. To address this variability, future iterations will transition to a Dynamic Weighting Framework. We plan to implement an

interactive mechanism, such as preference sliders, that allows users to clearly define their priorities. Mathematically, this transforms fixed coefficients into user-specific variables $(\alpha_u, \beta_u, \gamma_u)$, thereby enabling the system to generate personalized rankings tailored to individual constraints.

5.4.3. Limitations of the User Study

We acknowledge two primary limitations regarding the user evaluation in this study: the statistical power related to the sample size and the demographic diversity.

First, regarding the sample size ($N = 20$) and statistical power, practical constraints limited the number of participants. While prior HCI studies [34,35] suggest that samples of this size are often sufficient to reveal consistent preference trends among domain-aware users, we recognize that the ratio of participants to the number of ranked items (13 locations) constrains the degrees of freedom. Accordingly, the statistical analyses in this study, such as the Wilcoxon Signed-Rank Test results ($p < 0.05$ and specific concordance p -values detailed in Tables 3 and 4), are intended to provide supportive and exploratory evidence rather than definitive population-level inference. Although the results indicate consistent ranking behavior within this specific group, they should be interpreted with caution given the modest sample size.

Second, regarding demographic generalizability, the participants were exclusively university students from Japan and China. Specifically, the cohort consisted of 12 Chinese and 8 Japanese nationals. We conducted a qualitative comparison of ranking patterns between these two subgroups and observed broadly consistent preferences, particularly for highly popular and easily accessible pilgrimage sites. However, due to the limited sample size within each subgroup, no formal statistical tests were conducted for cross-national differences. Consequently, the findings may not fully generalize to other demographics. For instance, international tourists from Western countries may prioritize iconic or “must-visit” locations over transportation convenience due to the rarity of their visits, a preference structure that may differ from our study cohort.

To address these limitations, future studies will aim to test the system with a broader user base—increasing both the sample size to robustly assess statistical significance and the diversity of participants to validate the system’s effectiveness across broader cultural contexts.

6. Conclusions

This study established a comprehensive recommendation framework for anime pilgrimage that intelligently integrates three heterogeneous data sources: transportation accessibility, seasonal alignment, and cross-platform popularity. By synthesizing objective geospatial constraints with subjective content appreciation, the proposed system successfully addresses the complex Multi-Criteria Decision Making (MCDM) problem inherent in niche tourism planning.

This research significantly advances our preliminary work [8] by resolving its key limitations. Specifically, we refined the popularity metrics by aggregating data from four global platforms to reduce bias, and we rigorously validated the model’s effectiveness through an expanded user study and a comprehensive ablation analysis.

The experimental results demonstrate that the proposed integrated model achieves a MAP of 0.829 and an NDCG@5 of 0.814, significantly outperforming single-metric baselines. The ablation study provided critical insights into user behavior: while “Anime Popularity” and “Transportation Accessibility” serve as the primary drivers for decision-making during the planning phase, “Seasonal Alignment” functions as a qualitative enhancer that ensures

immersive experiences. This confirms that the weight configuration employed in our model effectively balances the trade-off between user interest and logistical feasibility.

Although the current framework has demonstrated strong predictive capability, avenues for future research remain to meet higher technical standards. First, regarding **Mathematical Optimization**, our goal is to transition from the current linear weighting system to nonlinear adaptive models that more effectively capture complex user preferences. Second, regarding **Data Validation**, future iterations will integrate automated outlier detection to ensure the reliability of social media metrics against manipulation. Finally, to further personalize the experience, we plan to integrate Large Language Models (LLMs) to analyze unstructured reviews. This will enable the system to differentiate subtle preferences, such as “architectural backgrounds” and “character-focused scenes”.

In summary, this research contributes a validated, scalable data analysis framework to the field of smart tourism. It bridges the gap between the virtual world of anime and the physical world of travel, providing a scientific foundation for supporting the growing global phenomenon of content tourism.

Author Contributions: Conceptualization, Y.Z. and Y.W.; methodology, Y.Z.; software, Y.Z.; validation, Y.Z. and Y.W.; formal analysis, Y.Z.; investigation, Y.Z.; resources, Y.Z.; data curation, Y.Z.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.Z. and Y.W.; visualization, Y.Z.; supervision, Y.W.; project administration, Y.W.; funding acquisition, Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by JSPS KAKENHI Grant Numbers JP21K17862 and JP25K15354.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to its nature as an anonymous user questionnaire survey, which does not involve personal data collection or invasive procedures, in accordance with the “Ethical Guidelines for Medical and Biological Research Involving Human Subjects” established by the Japanese Ministry of Health, Labour, and Welfare (MHLW) and the “Regulations for General Research Involving Human Subjects” at Yamaguchi University.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets used in this study are openly available as follows: The Bangumi Anime Database, available at <https://bangumi.tv/> (accessed on 14 February 2025); the AniList Anime Dataset, available at <https://anilist.co/> (accessed on 14 February 2025); the MyAnimeList Database, available at <https://myanimelist.net/> (accessed on 14 February 2025); and the Filmarks Anime Review Dataset, available at <https://filmarks.com/> (accessed on 14 February 2025). Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CPPI	Cross-platform popularity index
LBS	Location-Based Services
LLMs	Large language models
LWC	Linear weighted combination
MAP	Mean average precision
MCDM	Multi-criteria decision making
MRR	Mean reciprocal rank
NDCG	Normalized discounted cumulative gain
POI	Point-of-Interest

References

- Oishi, G. Introduction to the history of ANIME tourism : Newly revised edition (in Japanese). *Bull. Toyama Prefect. Univ.* **2020**, *30*, 25–35.
- Wikipedia Contributors. Seichi Junrei. 2025. Available online: https://en.wikipedia.org/wiki/Seichi_junrei (accessed on 12 November 2025).
- Squire, S.J. The cultural values of literary tourism. *Ann. Tour. Res.* **1994**, *21*, 103–120. [CrossRef]
- Jichitai Works WEB. Revitalizing the Local Area Through Anime Pilgrimages! How to Revitalize the Town Based on Case Studies (In Japanese). 2025. Available online: <https://jichitai.works/articles/2944> (accessed on 12 November 2025).
- Japan National Tourism Organization (JNTO). Inbound Visitor Statistics. 2025. Available online: <https://statistics.jnto.go.jp/en/graph/#monthly> (accessed on 26 February 2025).
- Development Bank of Japan. *Content and Regional Revitalization—100 Years of Japanese Anime, Focusing on Pilgrimages to Holy Lands*; Technical Report; Development Bank of Japan: Tokyo, Japan, 2017. (In Japanese)
- Nagata, H.; Makinojo, Y. 7 Successful Examples of “Anime x Town Revitalization”: The History of Pilgrimages and How to Revitalize Local Communities (In Japanese). 2025. Available online: https://lumiarch.ntt-east.co.jp/articles/202509_anime/ (accessed on 12 November 2025).
- Zhou, Y.; Wang, Y. Tourist Spot Recommendation for Anime Pilgrimage Considering Anime Features and Transportation Accessibility. In Proceedings of the 2025 IEEE 14th Global Conference on Consumer Electronics (GCCE), Osaka, Japan, 23–26 September 2025; pp. 557–561. [CrossRef]
- Soni, B.; Thakuria, D.; Nath, N.; Das, N.; Boro, B. RikoNet: A Novel Anime Recommendation Engine. *arXiv* **2021**. [CrossRef]
- Javaji, S.R.; Sarode, K. Hybrid Recommendation System using Graph Neural Network and BERT Embeddings. *arXiv* **2023**. [CrossRef]
- Ye, M.; Yin, P.; Lee, W.C.; Lee, D.L. Exploiting geographical influence for collaborative point-of-interest recommendation. In Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR ’11), Beijing, China, 24–28 July 2011; pp. 325–334. [CrossRef]
- Bao, J.; Zheng, Y.; Wilkie, D.; Mokbel, M. Recommendations in location-based social networks: A survey. *Geoinformatica* **2015**, *19*, 525–565. [CrossRef]
- Yuan, Q.; Cong, G.; Ma, Z.; Sun, A.; Thalmann, N.M. Time-aware point-of-interest recommendation. In Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR ’13), Dublin, Ireland, 28 July–1 August 2013; pp. 363–372. [CrossRef]
- Lian, D.; Zhao, C.; Xie, X.; Sun, G.; Chen, E.; Rui, Y. GeoMF: Joint geographical modeling and matrix factorization for point-of-interest recommendation. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’14), New York, NY, USA, 24–27 August 2014; pp. 831–840. [CrossRef]
- Alnogaithan, O.; Algazlan, S.; Aljuraiban, A.; Shargabi, A.A. Tourism Recommendation System Based on User Reviews. In Proceedings of the 2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakhr, Bahrain, 22–23 September 2019; pp. 1–5. [CrossRef]
- Ma, W.; Wang, Y. A Sentiment-Enhanced Recommendation System for Anime Pilgrimage Tours. In Proceedings of the 2024 IEEE 13th Global Conference on Consumer Electronics (GCCE), Kitakyushu, Japan, 29 October–1 November 2024; pp. 352–356. [CrossRef]
- Kuang, H.; Chang, L.; Bin, C.; Gu, T. Review of a Context-Aware Travel Recommendation System. *CAAI Trans. Intell. Syst.* **2019**, *14*, 611–618.
- Sha, Q.; Ma, L.; Chen, N. A Smart Tourism Resource Information Management Platform Based on Multi-Source Data Fusion. *Int. J. High Speed Electron. Syst.* **2025**, *34*, 2540134. [CrossRef]
- Sudarshan, K.; Bhagyalakshmi, M.V. Seasonal Tourist Recommendation Portal. *Int. J. Sci. Res. (IJSDR)* **2023**, *8*, 249–256.
- Zheng, R.; Zhen, S.; Cai, A.; Li, S.; Liu, H. Anime tourists traveling to Japan: Pilgrimage behaviors patterns and the formation of homologous emotions. *Front. Psychol.* **2025**, *16*, 1637114. [CrossRef] [PubMed]
- Putra, I.W.K.A.; Priatdana, G.Y.A. Implementation of the TOPSIS Method for a Decision Support System in Recommending Tourist Destinations in Tabanan. *J. Sist. Inf. Dan Komput. Terap. Indones. (JSIKTI)* **2025**, *6*, 460–469.
- Buasri, N.; Sangpradid, S. Enhancing Tourist Experience: A Fuzzy TOPSIS Approach to Destination Recommendation Systems. *GeoJournal Tour. Geosites* **2025**, *52*, 15–24. [CrossRef]
- NAVITIME Japan. NAVITIME Route (Totalnavi). 2025 Available online: <https://rapidapi.com/navitimejapan-navitimejapan/api/navitime-route-totalnavi> (accessed on 14 February 2025).
- Google Maps Platform. Geocoding API Overview. 2025. Available online: <https://developers.google.com/maps/documentation/geocoding/overview> (accessed on 14 February 2025).
- Wardman, M. A review of British evidence on time and service quality valuations. *Transp. Res. Part E Logist. Transp. Rev.* **2001**, *37*, 107–128. [CrossRef]

26. Guo, Z.; Wilson, N.H. Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground. *Transp. Res. Part A Policy Pract.* **2011**, *45*, 91–104. [CrossRef]
27. Schakenbos, R.; Paix, L.L.; Nijenstein, S.; Geurs, K.T. Valuation of a transfer in a multimodal public transport trip. *Transp. Policy* **2016**, *46*, 72–81. [CrossRef]
28. Wikipedia Contributors. Bangumi Fangzu Jihua (In Chinese). 2024. Available online: https://zh.wikipedia.org/zh-cn/Bangumi_ (accessed on 14 February 2025).
29. Mojimoon. Bangumi Anime Ranking Analysis. 2023. Available online: <https://github.com/mojimoon/bangumi-anime-ranking> (accessed on 14 February 2025).
30. Manami-Project. Anime-Offline-Database. 2025. Available online: <https://github.com/manami-project/anime-offline-database> (accessed on 14 February 2025).
31. Wikipedia Contributors. MyAnimeList. 2025. Available online: <https://en.wikipedia.org/wiki/MyAnimeList> (accessed on 14 February 2025).
32. Wikipedia Contributors. Filmarks (In Japanese). 2025. Available online: <https://ja.wikipedia.org/wiki/Filmarks> (accessed on 14 February 2025).
33. DOCOMO ANIME STORE Inc. Anime Viewing Numbers (Overall) Ranking (In Japanese). 2025. Available online: https://animestore.docomo.ne.jp/animestore/CR/CR00000014?ranking_type=views&period=all (accessed on 14 February 2025).
34. Nielsen, J. *Why You Only Need to Test with 5 Users*; Nielsen Norman Group: Fremont, CA, USA, 2000.
35. Faulkner, L. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behav. Res. Methods Instruments Comput.* **2003**, *35*, 379–383. [CrossRef] [PubMed]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.