International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 5146-5158

https://doi.org/10.70517/ijhsa463442

The design of innovative path of inheritance of non-heritage music and art based on big data analysis and cultural inheritance education.

Mengsa Chang[1,*]
[1] College of Humanities and Arts, Xi'an International University, Xi'an, Shaanxi, 710077, China
Corresponding authors: (e-mail: changmengsa2021@163.com).

Abstract: As an important carrier of Chinese culture, the inheritance and innovation of non-heritage music faces multiple challenges in the digital era. Using big data analysis and complex network communication dynamics as the core tools, this paper explores the digital communication paths and educational practice strategies of non-heritage music by constructing an information cascade model in social networks. The study first starts from the practical significance of integrating non-heritage music into college education, and clarifies its communication value in cultural heritage and youth groups. It also combines node cascade feature modeling, including temporal relationship and preference similarity analysis and complex network propagation dynamics model (SI, SIS, SIR and threshold model), to quantify the propagation law and diffusion threshold of non-heritage music content in social networks. Based on the empirical data of Weibo and Tik Tok platforms, Monte Carlo simulation and numerical iterative experiments are conducted to reveal the spatio-temporal evolution characteristics and propagation mechanism of non-legacy music information under different network topologies (SF and RR networks). The empirical study and Monte Carlo simulation experiments reveal the spatio-temporal evolution characteristics of non-heritage music information dissemination: structured forwarding dependence (84.46% of the initial forwarding dependence on the attention relationship), the lifecycle bimodal characteristics (two forwarding peaks within 72 hours), and the regulatory mechanisms of the dissemination parameters (α, γ, μ) on the diffusion efficiency. Experiments show that the cascade incremental prediction performance of the proposed complex network dynamics model on Aminer, SinaWeibo, and Twitter datasets significantly outperforms that of the existing methods (e.g., MSE is reduced to 2.862, and RMSPE is 0.483), which verifies its potential for application in high-precision prediction and optimization of propagation strategies.

Index Terms cultural heritage, non-heritage music, node cascade features, propagation dynamics modeling

I. Introduction

Intangible cultural heritage is the essence of national culture, the crystallization of national wisdom and the soul of national spirit, which not only has significant historical and cultural value, spiritual aesthetic value, but also has the value of scientific research and economic and social value that cannot be underestimated. And music non-heritage with unique art form and cultural connotation, become the bond of national spiritual identity [1]. As a carrier of national historical memory and cultural identity, non-heritage music has been continued through oral singing and transmission from generation to generation, with distinctive ethnicity and regionality, while carrying the expression of emotions, life wisdom and aesthetic pursuit of the ancestors [2], [3]. However, in the current wave of globalization and modernization, the intangible cultural heritage of music is facing unprecedented challenges, and its survival and development space has been squeezed.

On the one hand, the most effective mode of national culture protection and inheritance is not education [4]. The society will teach and record all kinds of traditional culture, no matter it is in the form of knowledge, art, or folklore, custom, language, etc., which exist in the people, through education, and will protect and pass this on [5], [6]. Colleges and universities, as the core position of cultural inheritance and innovation, assume the key functions of non-heritage education, research and dissemination, and contemporary college students are also the main force of non-heritage music inheritance [7]. However, as far as education is concerned, there are also a variety of modes, and what is the most appropriate and effective mode to use still needs to be recognized in practice.

On the other hand, the rapid development of digital technology provides a new path for the inheritance and innovation of intangible cultural heritage [8]. Digital means can not only break through the time and space limitations of traditional inheritance methods, but also activate the inherent vitality of ICH in an innovative form, so that it can be reintegrated into people's lives in the context of the new era [9], [10]. The application of big data technology, which can obtain the relevant data of non-heritage music in various forms such as text, pictures, audio and video, can establish a complete and clear portrait of non-heritage, which will certainly have an important and positive impact on the protection, inheritance and dissemination of non-heritage [11]-[13].

The article takes big data analysis as the core tool, combines with the theory of complex network communication dynamics, constructs the information cascade model in social networks, and aims to provide theoretical support and technical paths for the digital communication and educational practice of non-heritage music. Firstly, we start from the practical significance of integrating non-heritage music into college education, and clarify its communication value in cultural inheritance and youth groups. Subsequently, focusing on the information dissemination mechanism in social networks, we quantify the dissemination law of NRM content in digital platforms through cascade feature modeling of nodes, including temporal relationship and preference similarity analysis, and construct a dynamic model of NRM content in digital dissemination. Finally, combining the complex network propagation dynamics models (e.g., SI, SIS, SIR and threshold models), we explore the diffusion threshold and propagation efficiency of NRM information in different network topologies, and introduce the heterogeneous mean-field method to provide a more adaptive theoretical framework for the precise optimization of the NRM propagation strategy through the correlation analysis of degree distribution and node state.

II. Social network communication modeling and its application in non-heritage music education

II. A. The significance of integrating non-heritage music into music education in colleges and universities

Integrating non-heritage music into music education in colleges and universities can make the two promote and develop each other. As an important position for cultural inheritance, colleges and universities have the responsibility to incorporate these precious cultural heritages into the teaching system, so that the ancient music culture can be continued among the young generation and avoided to be forgotten in the wave of modern society.

China has rich and diverse non-heritage music resources, whether it is folk songs, traditional operas or instrumental performances, all of which can provide sufficient and unique materials for music teaching in colleges and universities. These materials can be transformed into curricular content, teaching cases and practical projects to broaden students' musical horizons and inspire their creativity. At the same time, the proposal of policies related to the protection of non-heritage culture provides strong support for the initiative of non-heritage culture into the campus. This provides a good policy environment and guarantee for the integration of Hebei's non-heritage music into music education in colleges and universities, which is conducive to integrating the resources of all parties and promoting the smooth development of the integration work.

II. B. Modeling cascading features of nodes

The value of non-heritage music in college education not only relies on the integration of teaching content, but also requires the use of modern technology to realize its vitality in mass communication. To this end, this section shifts the perspective to the information dissemination mechanism in social networks, and quantifies the dissemination law of NRM in digital platforms by constructing a cascade feature model of nodes.

The cascading features of nodes reveal the relationship of each node in an online social network with other nodes in the information cascade, which is crucial for predicting whether a node is infected in the information cascade and the order of infection. Two types of information cascade features are defined in this section: cascade context features and cascade affinity features.

II. B. 1) Cascading Context Modeling

The cascade context of a node should capture the temporal relationship of the node to other nodes in the information cascade, including the potential influence exerted by other nodes on that node and the temporal order in which they become infected in the information cascade. Modeling cascade contextual features requires consideration of three factors. First, the infection behavior of a node should intuitively be potentially influenced by all previously infected nodes, and this potential influence gradually decays with the gradual passage of time. Second, the cascading context of a node should be specific for a given information cascade, since a node may have very different cascading contexts in different information cascades. Third, whether a node is infected or not in a particular information cascade should be affected neither by nodes infected after it nor by nodes that have not participated in that information cascade at all. Based on these heuristic ideas, the cascade context is defined as follows:

Definition 1 (Cascade Context) Given an online social network $G=(V,E)$ containing N nodes, and a collection of information cascades $C=(C_1,C_2,\cdots,C_M)$ consisting of M information cascades occurring in G, the cascade context of all nodes in a given information cascade $C_m\left(1\le m\le M\right)$ can be defined as a matrix $X^{(m)}\in\mathbb{R}^{N\times N}$. The M information cascades $C_1,C_2,\cdots,C_M$ correspond to the M cascade context matrices $X^{(1)},X^{(2)},\cdots,X^{(M)}$. The elements $x_{u,v}^{(m)}$ of the u th row and v th column of the matrix $\mathbf{X}^{(m)}$ denote the possible influence of the node v on the node u, which is defined by the following equation:

$$x_{u,v}^{(m)}=\begin{cases}\exp\left(-\dfrac{t_u^{(m)}-t_v^{(m)}}{\tau}\right) & t_u^{(m)}<t_v^{(m)}\\ 0 & t_u^{(m)}\ge t_v^{(m)}\end{cases}\tag{1}$$

where the symbol $t_u^{(m)}$ denotes the time when the node u is infected in the information cascade C_m and the symbol τ denotes the time decay factor. The cascade context feature of node u in the information cascade C_m is the row vector $X^{(m)}$ corresponding to the matrix $x_u^{(m)}=X_{u,*}^{(m)}$.

The row vector $X^{(m)}$ of the cascade context matrix $x_u^{(m)}$ will be used as an input to the individual self-encoders in the DCE model, which quantitatively captures the temporal relationship of the node u with the other nodes in the information cascade C_m, and this temporal relationship encompasses two aspects: potential impact and relative order of infection.

II. B. 2) Cascade affinity feature modeling

As mentioned earlier, the cascade affinity between two nodes measures the similarity of their preferences regarding the information cascade, and thus can be modeled based on the number of times they have ever co-occurred in all the information cascades on the network, as defined below:

Definition 2 (Cascade Affinity) Given an online social network containing N nodes $G=\left(V,E\right)$, and a set of information cascades $C=\left(C_1,C_2,\cdots,C_M\right)$, the cascading contextual features of all nodes in this network can be modeled as a cascading affinity matrix $A\in\mathbb{R}^{N\times N}$, and the cascading affinity between two nodes u and v can be expressed as a cascading affinity matrix $A\in\mathbb{R}^{N\times N}$ in the element of row u and column v, specifically defined as the ratio of the number of information cascades in which node u and node v are jointly involved to the number of all the information cascades in the network, which can be expressed as the following equation:

$$a_{u,v}=\frac{\left|\left\{C_k\left|u\in C_k,v\in C_k,C_k\in C\right.\right\}\right|}{\left|C\right|}\tag{2}$$

Definition 2 denotes that for any two nodes in a network, if the number of information cascades they have participated in together in their history is higher, the higher the cascade affinity between them, and intuitively the more similar their preferences for information cascades are. In this sense, the cascade affinity of two nodes is related to how similar their embedding vectors are in the latent feature space.

II. C. Complex network propagation dynamics

After clarifying the micro-cascade characteristics of NRM communication, how to predict its macro-cascade efficiency becomes a key issue. This section further combines the complex network propagation dynamics model to reveal the diffusion path and potential influence scope of NRM information in different social environments from the perspective of network topology and propagation threshold.

II. C. 1) Propagation dynamics model

Complex network propagation dynamics has a wide range of applications and can be further categorized into information propagation, social propagation, and coupled information-social propagation according to the propagation object. In this subsection, we will introduce some classical models in the study of communication dynamics, including the SI model, the SIS model, the SIR model, which describes information propagation, and the threshold model, which describes social propagation.

(1) SI model

The SI model assumes that an individual is in an easy-to-propagate state at any moment in time, which is used to describe and easily influenced individuals, and stipulates that an individual will be in an influenced state once the information is propagated. The differential equation for the change in the number of individuals in the network at moment t can be expressed as:

$$\begin{cases} \dfrac{ds(t)}{dt} = -\beta s(t)i(t) \\ \dfrac{di(t)}{dt} = \beta s(t)i(t) \end{cases} \tag{3}$$

(2) SIS model

The SIS model introduces an individual recovery mechanism based on the SI model. An individual is able to spread the infectious disease to the surrounding susceptible state neighbors with the transmission probability β, and is able to recover with the probability μ, so that the state of the individual changes to the susceptible state again. Thus the differential equation for the change in the number of susceptible individuals and individuals in the network at moment t can be expressed as:

$$\begin{cases} \dfrac{ds(t)}{dt} = -\beta s(t)i(t) + \mu i(t) \\ \dfrac{di(t)}{dt} = \beta s(t)i(t) - \mu i(t) \end{cases} \tag{4}$$

(3) SIR model

The SIR model is similar to the SIS model, which also adds subsequent stages on top of the SI model.The SIR model divides nodes into susceptible state nodes, affected state nodes, and recovery state nodes. The infected state individuals can spread information to the susceptible state neighbors with a propagation probability β, and each individual can also change to the recovery state with a certain probability μ, at the same time, the individual has the immunity to the information propagation, and will not be affected again, and the subsequent propagation process will no longer participate. Until the node state in the model stabilizes, the propagation process ends. At the t moment, the differential equation for the proportion of individuals in the network in the susceptible state, influence state, and recovery state can be expressed as:

$$\begin{cases} \dfrac{ds(t)}{dt} = -\beta s(t)i(t) \\ \dfrac{di(t)}{dt} = \beta s(t)i(t) - \mu i(t) \\ \dfrac{dr(t)}{dt} = \mu i(t). \end{cases} \tag{5}$$

(4) Threshold model

There is a reinforcement effect in the social communication of information, i.e., the probability of an individual adopting information will be influenced by the surrounding neighbors. In real life, if a piece of information is adopted by more than one neighbor, the individual will have a higher probability to choose to adopt the information. In this scenario, the threshold model places a node in one of the susceptible, adoptive, and recovery states, and sets a threshold T, so that when the proportion of adoptive nodes in a susceptible node's neighbors exceeds the set value of T, or the number of messages received from the neighboring nodes exceeds the threshold value of T, then the individual chooses to adopt the message and enters into the adoptive state.

II. C. 2) Heterogeneous mean field methods

Heterogeneous mean-field theory is an important extension to the traditional mean-field approximation that makes improvements in capturing the heterogeneity of the degree distribution of nodes in a network. The heterogeneous mean-field approach treats all nodes in the network with the same degree as equivalent, and can be used to characterize the effects of factors such as weights and degree distributions on the propagation process. At moment t, assuming that the probability that a node of degree k is simultaneously an influence state is $\rho_k(t)$, it follows that the proportion of influence state nodes in the network is:

$$\rho(t) = \sum_k P(k)\rho_k(t) \tag{6}$$

where $P(k)$ is the degree distribution of the network. In an unaffiliated network, the probability $\Theta(t)$ that a node in the easy influence state is connected to an influence state neighbor can be expressed as:

$$\Theta(t)=\frac{1}{\langle k\rangle}\sum_{k}^{k_{\max}}p(k)k\rho_k(t) \tag{7}$$

where $k_{\max}$ is the value of the maximum degree in the network. This in turn yields the differential equation for $\rho_k(t)$:

$$\frac{d\rho_k(t)}{dt}=-\rho_k(t)+\beta k[1-p_k(t)]\Theta(t) \tag{8}$$

On the right-hand side of Eq. (8), the first term denotes the probability that an influenced state node of degree k is restored to the susceptible state, and the second term denotes the probability of infection of the susceptible state node. Let $\bar{\rho}(t)=[\rho_1(t),\cdots,\rho_{k_{\max}}(t)]^T$, which is obtained as $\rho_k(0)\to 0$ due to the small number of nodes being affected at the initial moment, and in the vicinity of $\rho_k(0)\to 0$, the Eq. (8) is linearly expanded and can be obtained:

$$\frac{d\bar{\rho}(t)}{dt}=C\bar{\rho}(t) \tag{9}$$

where the Jacobian matrix $C=\{C_{kk'}\}$, which contains the elements:

$$C_{kk'}=\beta\frac{kk'P(k')}{\langle k\rangle}-\delta_{k,k'} \tag{10}$$

At the initial moment, if $\rho(t)$ shows an exponential growth trend, there will be a global outburst, and at the same time the maximum eigenvalue of the matrix $\mathbf{C}$, $\beta\langle k\rangle^2/\langle k\rangle-1$, is greater than zero, which leads to the inference that the threshold of the outburst is:

$$\beta_c^{HMF}=\frac{\langle k\rangle}{\langle k^2\rangle} \tag{11}$$

where $\langle k\rangle$ and $\langle k^2\rangle$ denote the first-order and second-order moments of the degree distribution, respectively. In summary, the heterogeneous mean field can predict the propagation range and burst threshold from the degree distribution information, and it is also able to characterize the connection between the network topology and the burst threshold. However, the heterogeneous mean field has some limitations due to ignoring the specific details of the network topology and oversimplifying the influence process. For some specific and complex network structures, it may need to be combined with other models or methods to more accurately describe the propagation process in the real world.

III. Characteristics of the spatial and temporal evolution of information dissemination, mechanisms of social dissemination and cascade incremental prediction

In Chapter 2, by constructing a social network communication model and a complex network dynamics framework, this paper provides a theoretical foundation for the digital communication of non-heritage music. In order to further verify the effectiveness of the model and reveal the spatial and temporal patterns of the information dissemination of non-heritage music, Chapter 3 is based on real social platform data and simulation experiments, and conducts empirical research from the structured forwarding characteristics, life cycle dynamics, network topology influence and other dimensions, and finally verifies the predictive ability and optimization value of the model through cascading incremental prediction.

III. A. Mining the spatio-temporal evolutionary characteristics of non-heritage music information dissemination

III. A. 1) Structured forwarding study

In order to further explore the structured forwarding of information about NRM, this paper selects hot topics related to NRM and NRM content (e.g., traditional opera performances, folk instrumental music teaching videos, interviews with NRM inheritors, etc.) released by official accounts on a social platform (e.g., Weibo, Tik Tok) as a dataset, and the dataset covers 50 popular NRM content released between 2020 and 2023 The dataset contains complete user forwarding link data, underlying concern network relationship and user interaction behavior data. It contains nearly 10,000 retweeting participants and the cumulative number of retweets exceeds 70,000 times. Specifically, this paper counts the proportion of unstructured and structured forwarding in the forwarding process of popular messages about non-legacy music in terms of the size of forwarding, and the size of structured and unstructured forwarding in different forwarding depths is shown in Figure 1.

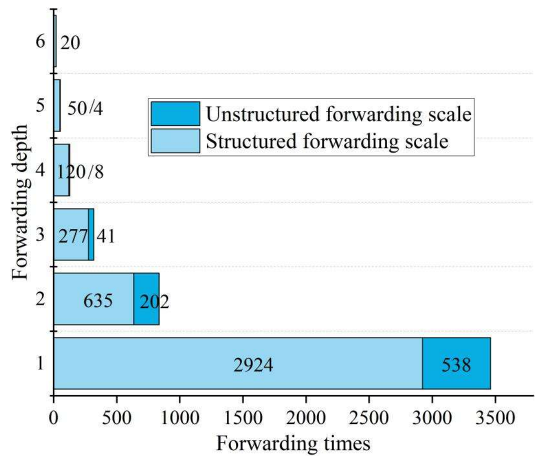

Figure: 1 The scale of structured and unstructured forwarding at different depths

Statistics found that with the increase of the number of forwarding times, the total forwarding scale decreases dramatically, for the same piece of information on nonstructured music, when the number of forwarding times is 1, the total forwarding scale is 3,462, of which 2,924 is structured forwarding scale, accounting for 84.46%, and the unstructured forwarding scale is 538, accounting for 15.54%; whereas, when the number of forwarding times is 6, the structured forwarding scale is 20, and there is no non structured forwarding size.

Statistics found that the spread of 5 messages relies on the relationship of “fans” and “followers” between users, and most of the spreading link is attached to the relationship network of users' followers; however, the spreading of the remaining 30 popular messages does not rely on the relationship of followers, and even the number of structured retweets is less than the total number of retweets, accounting for 15.54% of all retweets. The number of structured retweets is less than 5% of the total retweets, which accounts for 70% of the cases. This phenomenon questions the generalization of the study on predicting the scale of information dissemination based on the relationship between “followers” and “fans” among users, and brings new challenges to the problem of information dissemination prediction.

In order to address the above problems, this paper defines users' hidden fans as user groups that have not established direct “fans” or “followers” relationships with the users of information dissemination, but have participated in the information dissemination process of the users of information dissemination due to homogeneous interests or being in the same community. Construct user's hidden fan characteristics to quantify the user's potential spreading ability.

(1) Structured Forwarding Coefficient (HFC)

The structured forwarding coefficient in the forwarding network is used to quantify the extent to which the information relies on the “attention” and “fan” relationship in the social network, and the larger the structured forwarding coefficient, the more the information dissemination relies on the attention and fan relationship.

$$SFC(G_r) = \frac{F_S}{F} \tag{12}$$

where FS is the number of structured forwards and F is the total number of information forwards.

(2) User historical forwarding scale (HFS)

The average value of the forwarding scale generated by the user's historical information released in a certain period, which directly quantifies the scale of the user's hidden fans from the perspective of information dissemination; the larger the scale of the user's historical forwarding scale, the larger the scale of the user's hidden fan group, and the more the user's released messages have the influence of dissemination.

$$HFS\left(G_m\right)=\frac{\sum_{j=1}^{m}F_j}{n} \tag{13}$$

where Fj is the forwarding size of a particular historical posting message and n is the number of historical posting messages of the selected user.

III. A. 2) Information diffusion time feature extraction

Information disseminated in social networks undergoes a process from publication (from nothing to something), starts to be retweeted, and then stops being retweeted (from something to nothing) when the heat of the information fades. Figure 2 shows the curve of retweet size over time for popular tweets about Afrobeat on the Sina Weibo platform. The uncertainty of information life cycle increases the difficulty of quantifying the scale of information dissemination.

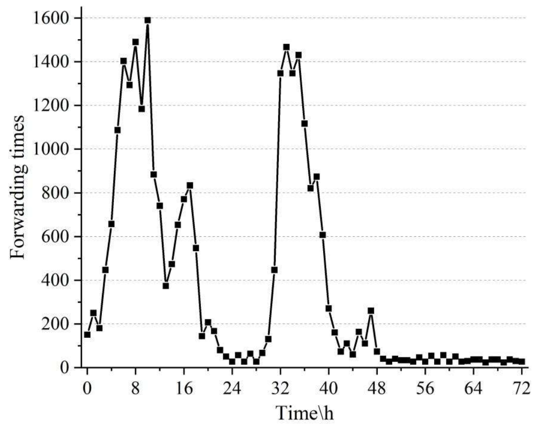

Figure 2: The life cycle length of information about intangible cultural heritage music

From Fig. 2, it can be seen that the popular microblog about this non-heritage music has two retweet peaks within 72 hours, indicating that it causes two waves of craze, which attracts the highest peak retweet peak of 1,594 retweets 10 hours after the microblog is sent, and the second time the retweet number reaches 1,478 retweets 33 hours after the microblog is sent, which reflects the persistence of the dissemination of the information about the non-heritage music.

In the process of information dissemination, this paper defines the life cycle of a piece of information as starting from Tstart when the information is released and ending with Tend when the last user retweets it. In this paper, the "life cycle time" tlife is used as a metric to measure the information dissemination time:

$$\begin{cases} T_l = T_{end} - T_{start} \\ T_{start} = \min\left\{t \middle| t \in T\right\} \\ T_{end} = \max\left\{t \middle| t \in T\right\} \end{cases} \tag{14}$$

where T denotes the sequence of time that a message is forwarded during its life cycle.Tend refers to the time of the last forwarding in the message dissemination process, the last forwarding means that it is not forwarded again within 72 hours, and Tstart is the time when the message is published.

III. B. Research on Social Communication Mechanisms on Complex Networks

Based on the quantitative analysis of the spatio-temporal evolutionary characteristics (structured forwarding dependence, life cycle duration) of NRM information dissemination in Section 3.1, Section 3.2 further combines the theory of propagation dynamics of complex networks to explore the regulatory mechanisms of network topology (SF, RR networks) and propagation parameters (α, γ, μ) on the scope and efficiency of NRM information diffusion by means of Monte Carlo simulation and numerical iterative experiments. The following is a summary of the results of the simulation and numerical iteration experiments.

III. B. 1) Experimental setup

In this part, numerical iterative experiments and Monte Carlo simulation experiments based on the theory of microscopic Markov chains will be used to investigate how the information-driven resource allocation strategy affects the dissemination process of non-heritage music information driven by cultural heritage education. Two random rule (RR) networks with different topologies were used in the experiments to model the information dissemination layer. The size of the network is set to N=1000, which means that there are 1000 nodes (individuals) in the network, and the average degree between the propagation layers satisfies <k1>=<k2>=10.

III. B. 2) Simulation analysis with different parameters

First, we investigate the time evolution of the information dissemination range for different α (probability of information dissemination), γ (probability of information recovery), and μ (dependence of information recovery on resources). In each simulation plot, the curves indicate the results of numerical iterative experiments, and the symbols indicate the results after averaging 500 independent Monte Carlo simulation experiments. At the initial moment, the seed ratio is set to $\rho AI(0) = \rho AE(0) = 0.01$. Without loss of generality, the other parameters are set to $\varepsilon = 0.2$ and $\beta = 0.4$; the time evolution of the information propagation process at different α, γ, and μ is shown in Figures 3, 4, and 5, respectively (the metrics of consideration are the proportions of the nodes that are in state X).

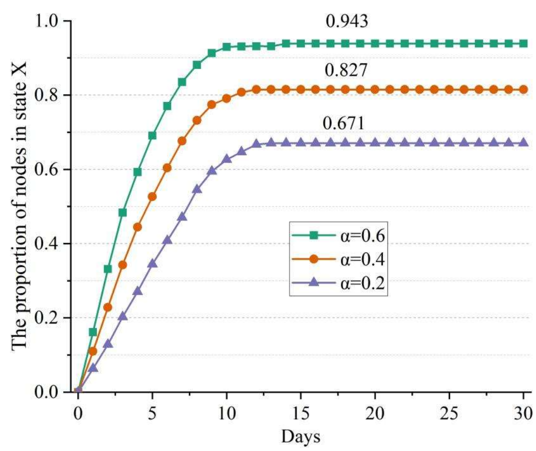

Figure 3: The time evolution of the information dissemination process at different α

The moderating effect of the information propagation probability α shows that when $\alpha = 0.6$, the proportion of nodes eventually stabilizes at 0.943. This indicates that with high propagation probability, the information gradually covers more nodes through natural diffusion and eventually reaches a higher propagation breadth. As the probability of information dissemination decreases, when $\alpha = 0.2$, the initial dissemination is slower and the final coverage is only 67.1%.

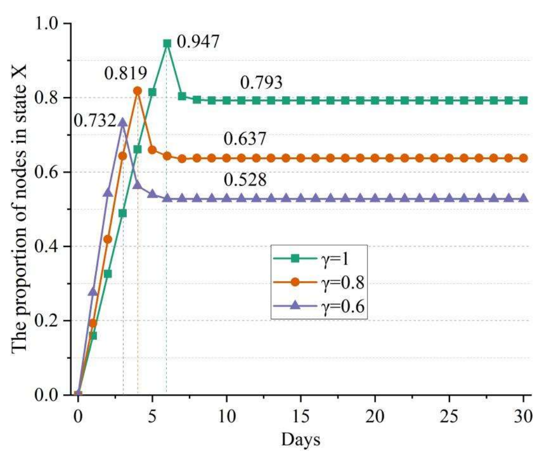

Figure 4: The time evolution of the information dissemination process at different γ

The moderating effect of information recovery probability γ on the system dynamics, with γ = 0.6, the node proportion rapidly jumps to 0.732 on the third day, and eventually stabilizes at 52.8% node proportion, indicating that a low recovery probability accelerates the exit of nodes from the propagation chain, resulting in the system rapidly reaching a steady state. When γ = 1, the node proportion grows slightly slower to 0.947 on day 6 and then stabilizes at a node proportion of 0.793, suggesting that the high recovery probability prolongs the survival time of information in the system. A low recovery probability γ accelerates system stabilization but may shorten the information life cycle; information dissemination is more persistent and efficient at high γ values.

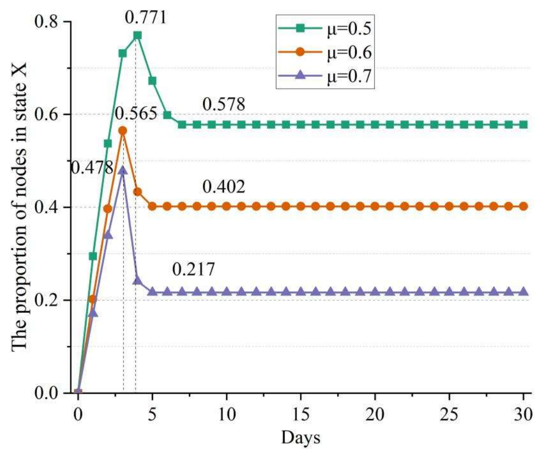

Figure 5: The time evolution of the information dissemination process at different μ

The moderating effect of the parameter μ (the degree of resource dependence on information recovery) is as follows: when μ = 0.5, the node proportion peaks at 0.771 on the third day and then drops to 0.578, indicating that low resource dependence allows nodes to recover quickly but may lead to a decrease in the efficiency of propagation at a later stage due to insufficient resources. When μ = 0.7, the proportion of nodes is always lower than 0.5, indicating that high resource dependence significantly inhibits the propagation range and needs to rely on external resource support to maintain the propagation chain. The study shows that the resource dependency μ directly affects the propagation sustainability, and at low μ values, the system has high propagation efficiency at the initial stage, but it also declines due to resource depletion.

Through simulation plots, it can be found that the Monte Carlo simulation experiments and the numerical iterative calculations of the microscopic Markov chain fit well, further verifying the accuracy of the theoretical derivation.

III. B. 3) Simulation analysis with different parameters

The article also investigates how the information-driven resource allocation model about cultural heritage education-driven information dissemination of non-heritage music evolves in different network structures. Specifically, we additionally use SF random networks with a network size of 1000 and a degree distribution index of 3 (P(k) ~ k-3) to describe the resource dissemination layer and the information dissemination layer with RR random networks and SF random networks, respectively. SF-RR in Fig. 6 indicates that the information dissemination layer is described by SF random network and the resource dissemination layer is described by RR random network. We simulated the change of information dissemination range with the growth of resource dissemination probability ε and information dissemination probability β, respectively, and the information dissemination of different network structures under different ε and β are shown in Fig. 6 and Fig. 7, respectively.

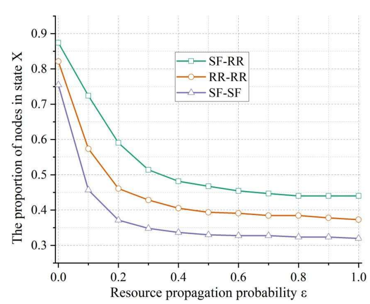

Figure 6: The dissemination situations of different network structures under different ε

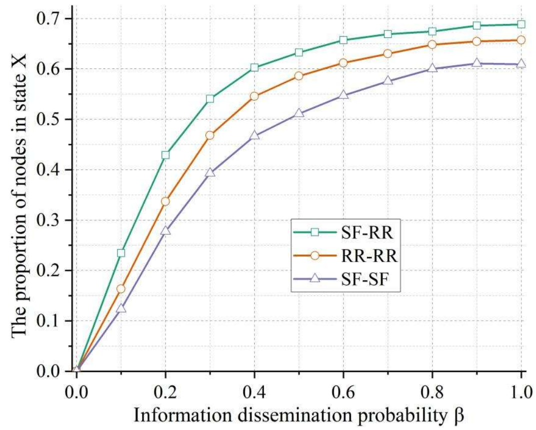

Figure 7: The dissemination situations of different network structures under different β

It is found that when the network structure of both the information dissemination layer and the resource dissemination layer are SF random networks, the spread of information is smaller, implying that the information is more easily controlled. When the resource propagation probability ε is 0, the node proportions of the three models SF-RR, RR-RR, and SF-SF are 0.875, 0.821, and 0.749, respectively, and the node proportions decrease to 0.468, 0.394, and 0.330 and regionally stabilize when ε is increased to 0.5. The node proportions increase with the increase of information propagation probability β. When the information propagation probability β increases to 1, the node proportions of the three models are 0.689, 0.657, and 0.609, respectively.A possible explanation is that there are nodes with larger degrees in the SF stochastic network, which are more likely to be given resources to restore their health in accordance with the resource allocation strategy of this study, reducing the possibility of large-degree nodes infecting more more nodes, which in turn makes the information spread less.

III. C. Cascade Incremental Forecasting Study

On the basis of clarifying the propagation mechanism and the influence of parameters, this section focuses on the practical application value of the model, and verifies the cascade incremental prediction performance of the proposed complex network dynamics model on three real datasets, namely, Aminer, SinaWeibo, and Twitter, through comparative experiments, to further confirm its superiority and generalizability compared with the existing methods from an empirical point of view.

III. C. 1) Data sets

Three real datasets in different scenarios are selected for experiments in this chapter, namely the paper citation dataset (Aminer) for predicting the citation cascade of academic papers about NRM, the Sina Weibo dataset (Sina Weibo) for predicting the re-tweeting cascade of posts about NRM, and the Twitter dataset (Twitter) for predicting the tweet dissemination cascade of tweets about NRM.).

Aminer: This dataset is an academic paper citation social network, which is widely used in cascade prediction research work, where citation cascades of papers characterize the spread of academic ideas and methods in academic social networks. The global academic paper citation social network structure G consists of citation data from 1992 to 2002; specifically, a link is established between author A and author B if author B cites a paper by author A. The citation cascade is a cascading cascade of cited papers. Thus, the users in a paper's citation cascade include the author of the paper and all authors who cite the paper. The 232 papers on NRM are selected to form the training set, and 63 and 38 papers are used as the validation and test sets. In the experiments on this dataset, the observation time length T is set to 1 year, and the prediction time interval Δt is set to 1, 2, and 3 years.

SinaWeibo: This dataset is derived from Sina Weibo, one of the most popular social networking platforms in China.

This dataset is derived from Sina Weibo, one of the most popular social networking platforms in China. 1569 posts about NRM were collected and publicly posted on the website. This dataset includes all Weibo posts with a retweet count greater than 10 and all retweets within 24 hours of publication. The post retweet cascade includes user ID, retweet time and retweet path information. This chapter uses cascades with posting times between 8:00 and 18:00 and retweets not exceeding 1000 for the experiments. Similarly, the dataset is divided into training, validation and test sets in the order of posting time, the top 80% of post forwarding cascades are taken as the training set and the remaining 20% of post forwarding cascades are equally divided into validation and test sets. In the experiments on this dataset, the observation time length T is set to 1 hour and the prediction time interval Δt is set to 1, 2 and 3 hours.

Twitter: this dataset contains information about URL links (messages) and corresponding participating users and participation times posted on the Twitter website about NRM in October 2024, and provides information about the structure of the network of inter-user attention relationships. Similarly, this chapter uses messages with more than 10 and less than 1000 retweets to conduct experiments, arranged in chronological order of posting, with the top 80% of the retweet cascades selected as the training set, and the remaining 20% of the retweet cascades equally divided into the validation set and the test set. In the experiments on this dataset, the observation time length T is set to 1 hour, and the prediction time interval Δt is set to 1, 2, and 3 hours.

III. C. 2) Contrasting models

In order to evaluate and validate the prediction performance of the proposed complex network propagation dynamics model, the following five state-of-the-art baseline models are selected for comparison in this chapter.

Features-linear: Its structural, user, temporal and content features are all commonly used and effective in cascade prediction tasks. The model mainly constructs and extracts two types of features, user features and structural features, for comparison experiments. The extracted features include: user ID and its global node degree, the mean and 90th percentile of the node degree in the cascade graph, the diameter of the cascade graph, the transmissibility, the average clustering coefficient, the number of leaf nodes, the total number of nodes, the number of triangles and the edge density. After extracting these features from the cascade information, they are fed into a linear regression model based on L2 regularization to predict cascade increments.

DeepCas: This model is the first classical model to use deep learning methods to solve the cascade prediction problem, which utilizes the structure of the cascade graph and user information to make predictions, and is an end-to-end deep learning cascade prediction method. the DeepCas model also uses recurrent neural networks to model the cascade sequences and aggregates the sequence representations through the attentional mechanism, but its attentional mechanism is limited by the mathematical aspects strong assumptions of geometric and polynomial distributions, and may not be applicable to real-world situations.

DeepHawkes: this model is based on the Hawkes process, which incorporates RNN-based user representation and path modeling as well as non-parametric time decay effects into the intensity function of the self-excited point process, combining the advantages of deep learning-like methods and generative model-like methods, and bridging the gap between predictability and interpretability of the information cascade to a certain extent.

CasCN: The idea of this model is similar to that of DeepHawkes model, which is modeled based on the self-excitation mechanism and time decay mechanism, but the main difference between the two is that CasCN model adopts Graph Convolutional Networks (GCNs) to extract the topology information of the cascade graph, and samples the cascade graph as a series of subgraphs instead of propagation paths, which can represent the structural information more accurately.

Coupled-GNN: This model represents the global propagation graph of the network based on a deep learning approach to cascade prediction. Specifically, the Coupled-GNN model models the iterative propagation process and interactions between node activation states and influence diffusion through two coupled graph neural networks, and finally pools and aggregates the activation state representations of all users to obtain the final message structure representation.

III. C. 3) Experimental setup

In this study, the predictive performance of the model is evaluated and validated using the mean square error (MSE) and root mean square percentage error (RMSPE), both of which are regression-type evaluation metrics widely used in the field of cascade prediction.The formulas for the calculation of MSE and RMSPE are given below:

$$MSE = \frac{1}{|C|}\sum_{i=1}^{|C|}\left(\hat{y}_i - y_i\right)^2 \tag{15}$$

$$RMSPE = \sqrt{\frac{1}{|C|}\sum_{i=1}^{|C|}\left(\frac{\hat{y}_i - y_i}{y_i}\right)^2} \tag{16}$$

where |C| is the total number of cascades, and y_i and $\hat{y}_i$ denote the incremental size true and predicted values of the ith cascade, respectively. It can be seen that the smaller the values of MSE and RMSPE, the better the predictive performance of the model.

Regarding the model parameters, the L2 regularization coefficient 2 is selected from the sets {1x10-8, 5x10-8, 0.01, 0.05, 0.1, 0.5, 1}, the number of hidden layer elements per GRU is selected from the sets{16,32,64}, and the learning rate of user embeddings and other variables is selected from the sets{0.0005,0.001,0.005,0.01, 0.05}. The user embedding means that the dimension is set to 50, the number of fully connected layers is set to 2, and its hidden layer dimensions are set to 32 and 16, respectively. In the process of random walk sampling, the smoothing coefficient ε is 0.01, the number of sequences K is 200, and the sequence length L is 10. In addition, the batch size of each iteration is set to 32, and when the training loss is no longer reduced after ten consecutive iterations, the training process is stopped.

III. C. 4) Analysis of results

The evaluation results of the different comparison models were tested and the overall prediction performance MSE and RMSPE values of all the models on the three different datasets are shown in Fig. 8, Fig. 9 and Fig. 10, respectively. The MSE and RMSPE metrics of the complex network dynamics model based on node cascade features proposed in this study have significantly decreased on the three datasets compared to all the other compared models, which shows that its overall prediction performance is better than the remaining five methods.

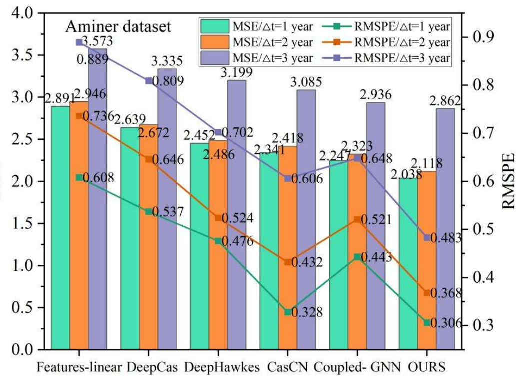

Figure 8: The predictive performance MSE and RMSPE values on Aminer

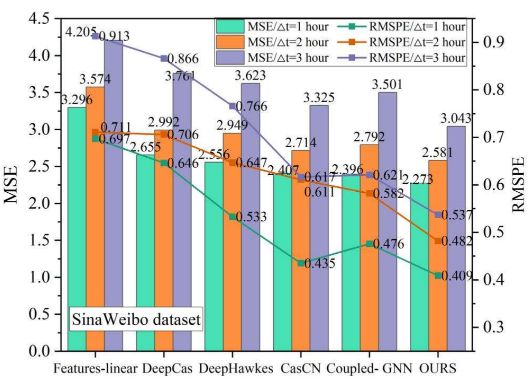

Figure 9: The predictive performance MSE and RMSPE values on SinaWeibo

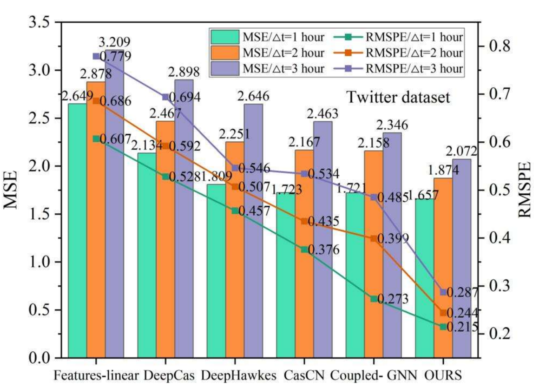

Figure 10: The predictive performance MSE and RMSPE values on Twitter

The above figures show the comparative results of the performance of different models in predicting the cascading increment of non-legacy music information on Aminer, SinaWeibo, and Twitter datasets, respectively. As a whole, the complex network propagation dynamics model proposed in this paper shows optimal performance in all datasets and time intervals (Δt = 1/2/3 years or hours), with MSE and RMSPE values significantly lower than those of the other compared models. For example, in the $\triangle t=3$ years prediction of the Aminer dataset, the MSE of this paper's model is 2.862 and the RMSPE is 0.483, while the MSE and RMSPE of the traditional linear model Features-linear are as high as 3.573 and 0.889, respectively, which is a significant difference. Similarly, the prediction error of this paper's model is always the lowest in SinaWeibo and Twitter datasets, especially in the prediction of Twitter dataset $\triangle t=3$ hours, its RMSPE is only 0.287, which is much better than other models, such as 0.485 in Coupled-GNN. Furthermore, the performance differences between the compared models reveal the limitations of the different modeling approaches: the linear model Features-linear, based on feature engineering, performs the worst by ignoring the complex propagation dynamics; Deep learning models such as DeepCas and DeepHawkes, although superior to traditional methods, are still limited by their ability to model network topology and temporal features; While the models CasCN and Coupled-GNN combined with graph neural network further improve the accuracy, the model in this paper realizes a more comprehensive capture of propagation laws by fusing the node cascade features with the theory of propagation dynamics of complex networks, thus verifying the scientific and adaptive nature of its method.

IV. Conclusion

This study systematically explores the laws and optimization strategies of digital dissemination of non-heritage music by constructing a social network dissemination model and a complex network dynamics framework. Based on the real data and simulation experiments on Sina Weibo and Tik Tok platforms, the information dissemination of non-heritage music shows significant structured forwarding dependence, with 84.46% of the initial forwarding relying on user-following relationships, and the lifecycle has a bimodal character, with the highest single-day forwarding peak reaching 1,594 times, which reveals the synergistic mechanism of the topology of social networks and the dissemination parameters.

Monte Carlo simulation further verifies that the information dissemination probability α, recovery probability γ and resource dependence μ have nonlinear regulatory effects on the dissemination scope and efficiency, and the coverage rate reaches 94.3% when $\alpha=0.6$. In addition, in the cascade incremental prediction experiments on Aminer, SinaWeibo and Twitter datasets, the proposed model outperforms the traditional linear model and deep learning methods with a significant advantage of 2.862 for MSE and 0.483 for RMSPE, proving its scientific and adaptive nature of fusing node cascade features with propagation dynamics.

References

[1] Ning, H. (2023). Analysis of the value of folk music intangible cultural heritage on the regulation of mental health. Frontiers in Psychiatry, 14, 1067753.

[2] Alivizatou-Barakou, M., Kitsikidis, A., Tsalakanidou, F., Dimitropoulos, K., Giannis, C., Nikolopoulos, S., ... & Grammalidis, N. (2017). Intangible cultural heritage and new technologies: challenges and opportunities for cultural preservation and development. Mixed reality and gamification for cultural heritage, 129-158.

[3] Howard, K. (2022). Musical instruments as tangible cultural heritage and as/for intangible cultural heritage. International Journal of Cultural Property, 29(1), 23-44.

[4] Yi, J. (2017, June). Research on inheritance of intangible cultural heritage. In 2017 2nd International Conference on Education, Sports, Arts and Management Engineering (ICESAME 2017) (pp. 1906-1910). Atlantis Press.

[5] Xu, J. (2019). Inheritance of national music culture and school music education. In Proceedings of 2019 9th International Conference on Management, Education and Information (MEICI 2019) (pp. 574-577).

[6] Yan, W. J., & Li, K. R. (2023). Sustainable cultural innovation practice: heritage education in universities and creative inheritance of intangible cultural heritage craft. Sustainability, 15(2), 1194.

[7] Chang, M. (2025). Cultural heritage education-driven utilization of NSGA algorithms to construct innovative path design for the inheritance of non-heritage music and art. J. COMBIN. MATH. COMBIN. COMPUT, 127, 5597-5615.

[8] Xu, Z., & Zou, D. (2022). Big data analysis research on the deep integration of intangible cultural heritage inheritance and art design education in colleges and universities. Mobile Information Systems, 2022(1), 1172405.

[9] Deng, J. (2023). A brief analysis of the path of intangible cultural heritage inheritance and innovative development under digital technology. Journal of Innovation and Development, 3(3), 29-32.

[10] Xu, W. (2023). Research on the communication opportunities of intangible cultural heritage under the background of big data and AI. Journal of Artificial Intelligence Practice, 6(8), 12-17.

[11] Wan, H., SIKKA, S., Wang, Z., Zhang, W., & Huang, T. (2023). Protection and Inheritance of Traditional Intangible Cultural Heritage Based on Big Data Technology: A Case Study in Guangxi. Academic Journal of Humanities & Social Sciences, 6(23), 76-85.

[12] Chang, W. (2025). The integration of artificial intelligence and ethnic music cultural inheritance under deep learning. Computer Science and Information Systems, (00), 36-36.

[13] Yang, Y., & Lei, T. (2022). The inheritance and future development direction prediction of opera culture based on cloud communication under the background of big data. Journal of Sensors, 2022(1), 1910766.