

Simulation Study on Digital Divide Propagation Path of Elderly Medical Treatment Based on Graph Neural Network and Public Social Network Data

Linna Yang ^{*} 

School of Economics and Management, Lanzhou Institute of Technology, Lanzhou, Gansu, China

*Corresponding author: Linna Yang. Email: linayang827@outlook.com

Abstract

To address the self-reinforcing diffusion of the digital divide in elderly medical treatment and the difficulty of precise intervention, this study proposes a Propagation–Adoption Coupled Graph Neural Network (PAC-GNN) and a path-level interpretable simulation framework. Using three types of public social network data, a heterogeneous multi-layer network with four node types and four edge layers was constructed, retaining 191607 users and 6057580 interaction records, with elderly users accounting for 19.3%; the elderly identification classifier achieved an AUC of 0.91. PAC-GNN integrates dual-channel decoupling, random-walk path encoding, cross-layer attention, and a causal masking mechanism to enable fine-grained modeling of propagation paths. Experimental results show that compared with the suboptimal baseline EvolveGCN, PAC-GNN reduces state prediction RMSE by 14.2%, improves path prediction F1 by 18.3%, and lowers literacy-dimension error by 19.5%; under 20% node perturbation, its performance decay is 8.3 percentage points lower than that of baselines. Simulation identifies five typical propagation patterns, among which intra-community diffusion accounts for 28.4% with an elderly participation rate of 68.7%, while cross-community bridging paths contribute nearly 60% of global state change. Multi-objective optimization indicates that the community-bridging strategy achieves the best overall balance among effectiveness, cost, and benefit equity.

Keywords

Digital divided in medical treatment for the elderly; Graph neural network; Propagation path simulation; Heterogeneous social networks; Optimization of intervention strategy

Received: 24 Jul 2026; Revised: 20 Aug 2026; Accepted: 23 Sep 2026; Published: 01 Oct 2026
Copyright: ©2026 The Author(s). Published by [Digital Intelligence Press Limited](#). This is an
open access article under the [CC BY 4.0](#) license.

1. INTRODUCTION

1.1 Research background and significance

Aging populations and the rapid spread of digital health services are creating a profound structural tension. On the one hand, the aging process of major economies arousing in the total population rate, with the proportion of the elderly population in the total population rising, and the total and complexity of the elderly's demand for medical services rising simultaneously. On the other hand, the transformation of the medical service system driven by digital technology is rapidly advancing, and core medical links such as appointment registration, online consultation, electronic prescription and examination report inquiry are increasingly dependent on digital platforms [1],[2],[3]. The essence of this tension lies in the fact that the digital speed of the supply side of medical services is significantly

faster than the digital adaptation speed of the elderly group at the demand side. In this context, a deep understanding of the formation mechanism and diffusion law of the digital medical divide for the elderly has become an important question to be answered in the field of digital health governance.

The digital divide in medical treatment for the elderly is not a single-dimensional technology access problem, but a superposition form of access gap, usage gap and literacy gap in medical treatment scenes [4]. The access gap is reflected in the objective gap of the elderly in equipment ownership, network coverage and platform accessibility; The usage gap is reflected in the elderly in the appointment registration, online consultation, prescription renewal and other operational skills proficiency differences; Literacy gap is reflected in the ability gap of the elderly in medical information understanding, credibility discrimination and decision transformation. The triple divide does not exist independently in the medical treatment scenario but reinforces each other: the lack of access conditions will limit the practice opportunities of using skills, and the lack of access skills will hinder the accumulation of information literacy, which in turn will reduce the willingness of the elderly to take the initiative to improve access conditions. This superposition effect makes the digital divide of medical treatment for the elderly show the characteristics of self-reinforcement, and it is difficult to fundamentally bridge the gap by simply increasing the supply of equipment or simplifying the operation process.

Social networks play a dual role of diffusion and mitigation in the evolution of the digital divide in medical treatment for the elderly [5],[6]. As a diffusion carrier, social networks accelerate the spread of the digital divide among homogeneous groups through information dissemination and behavior imitation. As a mitigation carrier, social networks also provide a low-cost channel for bridging the digital divide. Community organizations, medical institutions and active elderly users can transfer digital skills and trusted information to marginalized groups through information forwarding and experience sharing [7],[8]. This duality means that social networks are both a source of challenge and a key lever for governance of the digital divide. Understanding the propagation path of the digital divide in social networks and identifying key transmission nodes and structural bottlenecks have become the preconditions for designing effective intervention strategies.

Simulation methods have unique methodological value in digital divide research. The formation and evolution of the digital divide involve the multi-level interaction of individual behavior, social interaction and institutional environment, and real intervention experiments face significant constraints in cost, ethics and controllability. By constructing a computable communication model, the simulation method can deduce the evolution trajectory of the digital divide in a virtual environment, evaluate the potential effects of different intervention strategies, and provide a forward-looking basis for policy design. In recent years, the development of graph neural networks enables simulation methods to learn propagation dynamics directly on real social network structures, breaking through the simplifying assumptions of traditional regular network simulation. The combination of graph neural network and public social network data to construct the simulation framework of the communication path of the digital divide for the elderly is not only an expansion of the simulation methodology, but also a strong support for the governance practice of the digital divide.

1.2 Literature Review

Research on the digital divide among the elderly has experienced an evolutionary process from descriptive statistics to mechanism modeling. Early studies mainly relied on questionnaire survey and interview data to describe the distribution characteristics and influencing factors of the digital divide

from the dimensions of demographic characteristics [9],[10], technology acceptance and social support, and revealed the statistical correlation between variables such as age, education level, income and health status and the digital divide. Such research has laid a foundation for understanding the group differences of the digital divide, but it is difficult to describe the dynamic evolution of the divide and the interaction between groups. In recent years, researchers have begun to introduce mechanism modeling methods, regard the digital divide as a diffusion process in social networks, and use propagation model, game model and multi-agent simulation methods to explore the formation mechanism and evolution path of the divide [11],[12],[13]. This shift has promoted research from "which groups have the gap" to "how does the gap spread among groups", but there is still a lack of in-depth modeling of the coupling relationship between the communication process and the adoption process, and the structural characteristics of the communication path.

The application progress of graph neural network in information diffusion and behavior communication provides a new technical path for the study of digital divide. Graph neural networks learn node representations through neighborhood aggregation mechanisms, which can effectively capture structural dependence and information flow in social networks. In the research of information diffusion, graph neural networks have been used to predict the spread range of information, identify the key transmission nodes and infer the propagation path [14]. In the study of behavior propagation, graph neural networks are used to model the diffusion process of health behavior, consumption behavior and political participation in social networks. In recent years, the development of dynamic graph neural network and heterogeneous graph neural network has further expanded the application boundary, enabling the model to deal with complex scenarios of time series evolution and multi-type relationships [15],[16]. However, most of the existing graph neural network models treat the propagation process and adoption process as the same process, and do not fully consider the differences in time scale and influence mechanism between the two. At the same time, most models aim at state prediction and lack explicit modeling and interpretable tracking of propagation path.

Research on data-driven health behavior simulation in social networks has made some progress, but there are three core deficiencies. The first drawback is that the static perspective dominates. Most studies model the distribution characteristics of the digital divide in a single time section but fail to capture the temporal evolution of the divide state and the stage characteristics of the propagation process, resulting in a lack of basis for the timing selection of intervention strategies. The second drawback is the lack of real network structure. Some studies use regular networks or small-scale questionnaire data to build simulation environments, and there is a systematic deviation between the network topology and the real social network, which limits the ecological validity of the transmission path analysis. A third shortcoming is the lack of interpretable modeling of transmission paths. The existing research mainly focuses on macro indicators such as transmission scope and transmission speed but lacks fine-grained analysis of the specific composition of transmission path, the structural role of key nodes and the difference of path mode, resulting in the difficulty of precise positioning of intervention strategy design.

1.3 Research objectives and contributions

The objective of this study is to construct a simulation framework for the propagation path of digital divide for the elderly in medical treatment based on open social network data, propose an improved graph neural network model for heterogeneous propagation mechanism, and reveal the diffusion law and key intervention nodes of digital divide in social network. Specifically, the research aims to answer questions at three levels: at the data and structure level, how to use public social network data to build a heterogeneous multi-layer network that can reflect the characteristics of the elderly

medical scene; At the model and algorithm level, how to design the graph neural network architecture to capture the coupling relationship between the propagation process and the adoption process and realize the fine-grained modeling of the propagation path; At the simulation and application level, how to build a traceable, interpretable and optimized simulation engine to support the design and evaluation of intervention strategies.

The core contribution of this study is reflected in three aspects. The first contribution is a data-driven approach to the construction of high-fidelity heterogeneous social networks. This study takes open social network data as the only data source, establishes a multi-source collection strategy under the guidance of the three principles of open accessibility, elderly user coverage and medical discussion density, defines four types of node systems covering individual users, information sources, medical institutions and community organizations, and constructs a four-layer edge structure covering attention relationship, interaction relationship, information forwarding relationship and geographical proximity relationship. An edge weight calculation scheme integrating interaction frequency, emotion polarity and information credibility is designed. This method enables the subsequent propagation path analysis to be carried out on the real network topology and avoids the over-simplification of propagation law by regular network.

The second contribution is the path-aware graph neural network architecture of propagation-adoption coupling. This paper decouples the propagation process and adoption process into independent information channels and achieves re-coupling through coupling gating, so that the model can learn the propagation dynamics and adoption dynamics respectively. This paper introduces a path coding module based on random walk and a multi-path attention aggregation mechanism to enable the model to capture long-range propagation dependence and identify key propagation paths. This study proposes a causal masking mechanism to distinguish related transmission from causal transmission based on three criteria: time priority, intervention consistency and path accessibility. The age perception feature transformation module is designed to enable the model to learn differentiated state evolution patterns for elderly users.

The third contribution is a path-level interpretable digital divide simulation framework. The bidirectional algorithm of forward propagation path sampling and back propagation path tracing is designed, the critical path contribution evaluation method based on Shapley value is proposed, and the path pattern can be interpreted and summarized through clustering. This study constructs a counterfactual simulation mechanism to realize the causal assessment of the intervention effect through multi-scenario comparison. Three test schemes are designed to provide quantitative guarantee for the reliability of simulation results: step sensitivity analysis, random seed stability test and network scale scalability verification. This framework advances the research of digital divide from descriptive analysis to institutional deduction and strategic optimization.

1.4 Organizational structure

This paper is divided into seven sections, and the content of each section is organized according to the logical chain of "data base - model architecture - simulation engine - experimental verification - conclusion prospect". Section 1 is the introduction, which expounds the research background and significance, sorts out the research status and shortcomings at home and abroad, clarifies the research objectives and contributions, and explains the organizational structure of the paper. Section 2 is about data acquisition and heterogeneous social network construction, introducing the collection and preprocessing process of public social network data, extracting the behavioral characteristics related to the medical digital divide, constructing a heterogeneous multi-layer network covering four types of

nodes and four layers of edge relations, and analyzing the statistical characteristics of the network. Section 3 is to construct the propagation-adoption coupling graph neural network model, formalize the digital divide state space and propagation-adoption coupling relationship, design the infrastructure such as type-aware message passing, time series evolution, cross-layer attention and age-aware transformation, propose the improvement mechanism such as path coding, path attention, dual-channel coupling and causal mask, and explain the model training and optimization strategy.

In Section 4, for propagation path simulation engine and intervention strategy generation, a propagation dynamics simulator based on graph neural network is designed, discrete-time stepping and event-driven hybrid simulation mechanism is realized, bidirectional path generation algorithm of forward sampling and reverse tracing is constructed, four types of intervention means are modeled, and multi-objective evolutionary algorithm is used to solve the intervention combination optimization problem. Section 5 is about experimental design and result analysis, setting data set division, baseline model and evaluation index system, carrying out model performance comparison experiment, propagation path simulation result analysis, intervention strategy simulation experiment and interpretability analysis. Section 6 is the test of model robustness and generalization ability, which tests model reliability from three dimensions of data perturbation robustness, cross-network generalization ability and simulation effectiveness, and discusses model limitations. Section 7 is the conclusion and prospect, summarizing the main conclusions and innovations, and putting forward practical implications and future research directions.

2. DATA ACQUISITION AND HETEROGENEOUS SOCIAL NETWORK CONSTRUCTION

2.1 Collection and preprocessing of public social network data

The selection of data sources should meet three constraints at the same time: public accessibility, coverage of elderly users and discussion density related to medical treatment. Public accessibility requires data to be accessible without authorization, and the platform's terms of service allow for academic research use; The coverage of elderly users requires a sufficient number of identifiable elderly users in the platform to ensure the statistical validity of the digital divide propagation path. The density of discussion related to medical treatment requires that there be rich interaction of medical treatment topics in the platform, so that the transmission path has observable behavior carriers [17]. Based on the above three principles, this study selects three types of open data sources: comprehensive open forums, social media open groups and vertical health communities. Comprehensive forums provide a wide range of elderly user samples and cross-topic interaction, public social media groups provide high-frequency daily interaction and information forwarding behavior, and vertical health communities provide high-density discussion and professional information sources related to medical treatment. The three types of data sources are complementary in user overlap, topic distribution and interaction mode, and together constitute a multi-dimensional database.

The multi-source data collection strategy adopts the combination of hierarchical crawl and incremental updates. For each data source, the discussion area related to medical treatment was first located through the open hashtag and keyword set, which covered the whole process of medical treatment, such as appointment registration, medical insurance reimbursement, online consultation, prescription renewal, and examination report inquiry. Then, the discussion area is taken as the seed, and the expansion is bounded along the public attention relationship and interaction relationship, and the

expansion depth is controlled to three hops to balance the coverage and data scale. The collection content includes user public portrait information, public published text, public interaction record and public timestamp. In order to control the data scale and computing cost, the user activity threshold and time window constraint are set in the collection process, and only the users and their associated edges that have at least one medical-seeking interaction within the observation window are retained.

The data cleaning and privacy desensitization process consists of four stages. The first stage is to remove duplicates and standardize formats, eliminate cross-platform duplicate accounts and duplicate texts, and unify timestamp format and text coding. The second stage is noise filtering, eliminating advertising promotion, robot account and interaction record unrelated to medical treatment topics; Robot account identification adopts joint discrimination rules based on publication frequency, text templating degree and social relationship abnormality degree. The third stage is privacy desensitization, hash replacement and generalization of direct identifiers such as username, profile picture and geographical location, and mask replacement of sensitive information such as phone number, ID number and specific address in the text. As only publicly available data were involved and no human subject intervention was involved, no human ethical approval was required for this study. The fourth stage is quality verification, which ensures the integrity and accuracy of the cleaned data through sampling manual verification and statistical consistency test. The changes of data scale before and after cleaning are shown in [Table 1](#).

Table 1. Comparison of scale before and after multi-source data collection and cleaning

Type of data source	Collect the number of users	Number of users after cleaning	Collect the number of interaction records	Number of interaction records after cleaning	Proportion of elderly users
Comprehensive open forum	128,460	94,237	2,847,300	1,936,540	18.7%
Public social media groups	86,320	61,508	4,215,780	2,873,410	12.4%
Vertical Health Community	43,750	35,862	1,682,940	1,247,630	31.2%
Total	258,530	191,607	8,746,020	6,057,580	19.3%

It can be seen from [Table 1](#) that although the vertical health community has the smallest user scale, it accounts for the highest proportion of elderly users, which is the core data source of medical digital divide research. Comprehensive forums have the largest user base and provide rich cross-topic interaction background; The interaction frequency of public groups in social media is the highest, which is conducive to capturing the process of information forwarding and behavior imitation. After cleaning the three types of data sources, a total of 192,000 users and 6.058 million interaction records are retained, and the overall proportion of elderly users is 19.3%, which is basically consistent with the proportion of middle-aged and elderly Internet users in public statistics, indicating that the data samples are reasonably representative.

The elderly user recognition method is based on the classifier of public portrait information. Public portrait information includes observable characteristics such as users' self-reported age, registration time, age cues in posted content, topic distribution and age distribution of interactive objects. In this study, a gradient boosting decision tree classifier is constructed, which takes a small number of manually

labeled samples as the training set, takes multi-dimensional public features as the input, and outputs the probability that the user is the elderly. The importance analysis of classifier features shows that self-reported age and topic distribution are the strongest predictors, and the age distribution of interaction objects and the age cue of posted content provide complementary information. The AUC of the classifier on the validation set reached 0.91, indicating acceptable accuracy for elderly identification based on public portraits. For the users whose probability value is in the fuzzy interval, a conservative strategy is adopted to mark them as "suspected elderly users", and a lower weight of elderly attribute is given in the subsequent network modeling to avoid the excessive influence of classification error on the propagation path analysis.

2.2 Extraction of behavioral features related to the digital divide in medical treatment

The design of observable proxy indicators of the three dimensions of the digital divide is the key link to transform abstract theoretical constructs into computable features. The access gap reflects the accessibility difference of the elderly in digital medical infrastructure, and its proxy indicators include device type distribution, network access frequency, platform login frequency and duration of medical applications. The usage gap reflects the difference of the elderly in the use skills of digital medical services, and its proxy indicators include the completion rate of appointment registration operation, the initiation frequency of online consultation, the number of examination report inquiries and the success rate of prescription renewal operation. The literacy gap reflects the differences in the elderly's understanding and decision-making ability of digital medical information, and its proxy indicators include the depth of medical information retrieval, information credibility discrimination behavior, multi-source information cross-validation frequency and participation in discussions related to medical decision-making. The proxy indicators of the three dimensions together constitute the multidimensional representation vector of the digital divide status of the elderly.

The extraction of specific behavior features in medical scene focuses on four core behaviors: appointment behavior, information retrieval behavior, online consultation behavior and prescription acquisition behavior. The characteristics of appointment behavior include appointment initiation frequency, appointment completion rate, appointment cancellation rate, appointment department distribution and appointment time preference. The characteristics of information retrieval behavior include the number of search keywords, the depth of search, the distribution of search results and the interval of search time. The characteristics of online consultation behaviors include consultation initiation frequency, consultation duration, consultation department distribution and post-consultation behavior conversion rate. The behavior characteristics of prescription acquisition include the frequency of prescription inquiry, the power of prescription renewal, the frequency of drug information retrieval and the participation in medication consultation. The four types of behavioral characteristics are further aggregated into weekly, monthly and quarterly series in the temporal dimension to capture the temporal evolution of behavioral patterns.

Text semantic feature extraction uses pre-trained language model to represent the topic of medical related text. Specifically, based on the BERT model pre-trained on the public Chinese medical corpus, 768-dimensional semantic vector is extracted for each medical related text, and then the dimension is reduced to 64 dimensions through mean pooling and principal component analysis, which is used as the semantic features of user text [18]. In order to enhance the specificity of the medical scene of semantic features, this study carried out domain adaptation fine-tuning on the basis of the pre-training model. The fine-tuning task was multi-label classification of medical topics, and the label system covered 12 categories, including appointment registration, medical insurance policy, online consultation, drug

information, examination and testing, hospitalization service, and health science popularization. After fine-tuning, the macro average F1 of the model in the medical topic classification task reaches 0.87, indicating that the semantic representation can effectively capture the fine-grained topic differences of the medical scene.

The user behavior time series feature construction aligns the above behavior features with semantic features in the time dimension to form the user behavior time series matrix. Let the behavior feature vector of user u at time step t be $x_u^{(t)} \in \mathbb{R}^d$, where d is the feature dimension, then the behavior time sequence matrix of user u is expressed as $X_u = [x_u^{(1)}, x_u^{(2)}, \dots, x_u^{(T)}] \in \mathbb{R}^{T \times d}$, where T is the total number of observed time steps. In order to deal with the problem of inconsistent observation time of different users, the time mask mechanism is used to mark the effective time step, and the effective time sequence information is aggregated through mask attention in the subsequent graph neural network. The time granularity of time series features is set as weeks, and the observation window is 52 weeks, which not only ensures sufficient time resolution to capture behavioral changes, but also avoids data sparsity caused by excessive granularity.

2.3 Heterogeneous social network modeling

The definition of network node covers four types of entities: individual users, information sources, medical institutions and community organizations. Individual user nodes carry the behavioral and semantic features extracted in Section 2.2 and are the main body of digital divide propagation. Information source nodes represent public accounts that release medical information, including health science bloggers, medical policy release accounts and news media accounts. Their attributes include information release frequency, information credibility score and audience size. Medical institution nodes represent hospitals, clinics and pharmacies that can be identified in the public social network, and their attributes include institution type, service radius and online service capability. Community organization nodes represent mutual aid groups for the elderly, community health service organizations and offline activity groups, and their attributes include organization scale, activity and online and offline linkage ability. The four types of nodes together constitute the node set $\mathcal{V} = \mathcal{V}_U \cup \mathcal{V}_S \cup \mathcal{V}_M \cup \mathcal{V}_C$ of the heterogeneous network, where $\mathcal{V}_U$, $\mathcal{V}_S$, $\mathcal{V}_M$ and $\mathcal{V}_C$ represent the node sets of individual users, information sources, medical institutions and community organizations respectively.

Multi-layer edge relationship modeling divides the association between nodes into four layers: attention relationship layer, interaction relationship layer, information forwarding relationship layer and geographical proximity relationship layer. The attention relationship layer represents users' one-way attention to information sources or other users, reflecting the structural connection of information acquisition channels. The interactive relationship layer represents the two-way interaction between users, including comments, likes and private messages, which reflects the direct channel of social influence. The information forwarding relationship layer represents the forwarding behavior of medical information and reflects the path structure of information diffusion. The geographical proximity relationship layer represents the spatial proximity relationship between users based on public geographical location information, reflecting the potential impact of offline social networks on online communication. The four layers of edge relations together form a heterogeneous multi-layer network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{E} = \mathcal{E}_F \cup \mathcal{E}_I \cup \mathcal{E}_R \cup \mathcal{E}_G$ corresponds to the attention, interaction, forwarding and geographical proximity edge sets respectively. The multi-layer network structure enables the same pair of nodes to have different types of edges in different layers, thus capturing the multi-channel characteristics of digital divide propagation [19].

The edge weight design comprehensively considers three dimensions: interaction frequency, emotional polarity and information credibility. Let the weight of edge e be w_e , and its calculation formula is as follows:

$$w_e = \alpha \cdot f_e + \beta \cdot s_e + \gamma \cdot c_e \quad (1)$$

Where f_e represents the normalized value of interaction frequency, which is obtained from the logarithm transformation of interaction times in the observation window and the maximum and minimum normalization; s_e represents the sentiment polarity score, and the sentiment propensity value output by the text sentiment analysis model is linearly mapped to the interval $[-1, 1]$. Positive values represent positive emotions, and negative values represent negative emotions. c_e represents information credibility score, which is obtained by the product of information source credibility score and content credibility score; α , β and γ are the weight coefficients, satisfying $\alpha + \beta + \gamma = 1$, which are determined to be 0.5, 0.2 and 0.3 respectively through grid search. The edge weight design makes the network reflect not only the interaction strength, but also the emotional tendency and information quality, providing richer structural information for the subsequent propagation path analysis [20].

The analysis of network statistical characteristics is carried out from four dimensions: degree distribution, clustering coefficient, community structure and embedding mode of elderly users. The degree distribution analysis shows that the degree distribution of the four layers of the network shows the characteristics of power law, but there are differences in the power law index between layers: the power law index of the attention relationship layer is about 2.3, the interaction relationship layer is about 2.8, the information forwarding relationship layer is about 2.1, and the geographical proximity relationship layer is about 3.5. The clustering coefficient analysis shows that the clustering coefficients of the interaction relationship layer and the geographical proximity relationship layer are significantly higher than those of the attention relationship layer and the information forwarding relationship layer, indicating that the first two layers have stronger local clustering characteristics. The Louvain algorithm is used to detect community division in the network, and it is found that there is a significant community structure in the four-layer network, and there is a certain correlation between community division and geographical area, age group and topic preference. The analysis of elderly users' embedding pattern shows that elderly users are more inclined to interact with users of the same age in the interactive relationship layer, forming an age homogeneous community. In the information forwarding relationship layer, elderly users are more inclined to forward information from community organizations and information sources rather than information from other individual users. The summary of network statistical characteristics is shown in [Table 2](#).

Table 2. Statistical characteristics of heterogeneous multi-layer networks

Hierarchy of networks	Number of nodes	Number of edges	Degree of average	Coefficient of clustering	Degree of modularity	Homogeneity index for elderly users
Focus on the relationship layer	191,607	1,247,830	13.02	0.087	0.412	0.18
Interactive relationship layer	156,420	2,873,410	36.74	0.234	0.538	0.34
Information forwarding layer	98,750	1,936,540	39.22	0.112	0.476	0.27
Geographical proximity layer	143,280	862,940	12.05	0.318	0.621	0.41

The interactive relationship layer and information forwarding layer have the largest number of edges, which are the main channels of digital divide propagation. The clustering coefficient and modularity of geographical proximity layer are the highest, indicating that offline spatial structure has strong constraints on online transmission. The homogeneity index of elderly users is the highest in the geographical proximity layer (0.41), and the lowest in the attention relationship layer (0.18), indicating that the social interaction of elderly users is more driven by geographical proximity and age homogeneity, while they are relatively open in information acquisition channels. The above statistical characteristics provide structural priors for the design of subsequent graph neural network models: the model should consider the differentiated transmission mechanism of multi-layer networks and carry out special modeling of the highly homogeneous community structure of elderly users.

3. PROPAGATION - ADOPTION COUPLING GRAPH NEURAL NETWORK MODEL CONSTRUCTION

3.1 Problem formalization and modeling framework

The definition of digital divide state adopts multi-dimensional continuous state space. Let the digital divide state vector of individual user u in the network at time step t be $h_u^{(t)} \in \mathbb{R}^K$, where K is the state dimension is. Different from the existing research that simplifies the digital divide into binary categorical variables, this study designs the state space as a continuous vector to describe the gradual characteristics of the digital divide in the three dimensions of access, use and literacy. Specifically, the first three components of the state vector respectively correspond to the continuous values of the access gap, the usage gap and the literacy gap. The value range is $[0, 1]$, and the larger the value is, the deeper the gap. The other components were used to encode hidden states related to medical treatment scenarios, including information acquisition tendency, risk perception level and sensitivity to social influence. The evolution of state vector in the time dimension constitutes the trajectory of user digital divide state, and the propagation path simulation is to model the generation process of this trajectory under the constraints of network structure.

The modeling of the coupling relationship between the diffusion process and the adoption process is one of the core innovations of the model in this section. The propagation process describes the diffusion of information in the network, that is, the transmission process of medical information from

information sources or active users to other users; The adoption process describes the process of changing the state of the digital divide after individuals receive information. The two are not independent: communication provides information input for adoption, and adoption results in turn affect individuals' subsequent communication behaviors, forming a two-way coupling [21]. Let $m_{u \leftarrow v}^{(t)}$ represent the information vector transmitted from node v to node u at time step t , and $a_u^{(t)}$ represent the adoption behavior representation of node u , then the coupling relationship can be formally expressed as:

$$m_{u \leftarrow v}^{(t)} = \phi_{\text{msg}} \left(h_v^{(t-1)}, h_u^{(t-1)}, e_{uv}, a_v^{(t-1)} \right) \quad (2)$$

Where ϕ_{msg} is the type-aware message passing function, e_{uv} is the attribute vector of edge uv , and $a_v^{(t-1)}$ is the representation of the sender's adoption behavior at the previous time step. This equation shows that the message content not only depends on the sender's state and edge attributes, but also is affected by the sender's adoption behavior, so the adoption process is embedded in the propagation process [22]. Accordingly, the adoption process is modeled as:

$$a_u^{(t)} = \phi_{\text{adopt}} \left(h_u^{(t-1)}, \bigoplus_{v \in \mathcal{N}(u)} m_{u \leftarrow v}^{(t)} \right) \quad (3)$$

Where ϕ_{adopt} is the adoption decision function, $\bigoplus$ represents the aggregation operation of neighbor messages, and $\mathcal{N}(u)$ is the set of neighbors of node u . Adoption behavior represents the message passing that is then fed back to the next step, forming a two-way coupled closed loop of propagation-adoption.

The graph structure representation of time series propagation dynamics embedding the above coupling process into a multi-layer network structure. Let the heterogeneous multi-layer network constructed in Section 2 contain $L = 4$ layers, and the adjacency relationship of layer l is denoted as $\mathcal{E}_l$. The status update of node u at time step t is affected by both neighbor messages within the layer and the status of corresponding nodes across layers. For unified expression, the state of node u in layer l is defined as $h_{u,l}^{(t)}$, then the time sequence propagation dynamics can be expressed as:

$$h_{u,l}^{(t)} = \psi \left(h_{u,l}^{(t-1)}, \bigoplus_{v \in \mathcal{N}_l(u)} m_{u \leftarrow v,l}^{(t)}, \bigoplus_{l' \neq l} z_{u,l \leftarrow l'}^{(t)} \right) \quad (4)$$

Where ψ is the state update function, $\mathcal{N}_l(u)$ is the set of neighbors of node u in layer l , and $z_{u,l \leftarrow l'}^{(t)}$ represents the cross-layer information from layer l' to layer l . This expression unifies intra-layer propagation and cross-layer interaction in a temporal evolution framework and provides a formal basis for subsequent infrastructure design.

3.2 Basic graph neural network architecture design

Heterogeneous message passing mechanism adopts type-aware message passing function. Due to the heterogeneous properties of the four types of nodes and the four layers of edges in the network, the node type and edge type need to be distinguished in message passing. For edge uv with edge type $r \in \{F, I, R, G\}$, its message is calculated as follows:

$$m_{u \leftarrow v}^{(r,t)} = W_r^{(1)} \sigma \left(W_r^{(0)} \left[h_v^{(t-1)} \parallel h_u^{(t-1)} \parallel e_{uv}^{(r)} \right] + b_r \right) \quad (5)$$

Where $W_r^{(0)}$ and $W_r^{(1)}$ are edge type-specific learnable weight matrices, b_r are bias term, $\parallel$

represents vector splicing operation, σ is nonlinear activation function, and $e_{uv}^{(r)}$ is the attribute vector of edge uv under type r . The type-specific weight matrix enables the model to learn differentiated information transfer modes for the four relationships of attention, interaction, forwarding and geographical proximity, avoiding the homogenization of heterogeneous relationships.

Node status update adopts GRU driven time series evolution module. GRU can effectively capture long time series dependence and alleviate the problem of gradient disappearance, which is suitable for modeling the gradual evolution of digital divide state [23],[24]. Let the neighbor message obtained by node u aggregated at time step t be $m_u^{(t)}$, then the state update process is:

$$z_u^{(t)} = \sigma \left(W_z \left[h_u^{(t-1)} \parallel m_u^{(t)} \right] \right) \quad (6)$$

$$r_u^{(t)} = \sigma \left(W_r \left[h_u^{(t-1)} \parallel m_u^{(t)} \right] \right) \quad (7)$$

$$\tilde{h}_u^{(t)} = \tanh \left(W_h \left[r_u^{(t)} \odot h_u^{(t-1)} \parallel m_u^{(t)} \right] \right) \quad (8)$$

$$h_u^{(t)} = \left(1 - z_u^{(t)} \right) \odot h_u^{(t-1)} + z_u^{(t)} \odot \tilde{h}_u^{(t)} \quad (9)$$

Where $z_u^{(t)}$ is the update gate, which controls the retention ratio of the state at the previous time; $r_u^{(t)}$ is the reset gate, which controls the degree of historical information forgetting; $\tilde{h}_u^{(t)}$ is the candidate state; W_z , W_r and W_h are learnable weight matrices; $\odot$ represents the element-by-element product. The GRU module enables the node state to evolve smoothly, avoiding violent fluctuations caused by single-step message shock, which is in line with the realistic characteristics of gradual change of digital divide state.

Cross-layer attention aggregation between multi-layer networks solves the problem of information fusion between layers. The same node has different state representations in different layers, and the cross-layer attention mechanism allocates dynamic weights to each layer to achieve adaptive fusion. Let the state of node u at layer l be $h_{u,l}^{(t)}$, then the cross-layer attention weight is calculated as follows:

$$\alpha_{u,l}^{(t)} = \frac{\exp \left(q^\top \tanh \left(W_c h_{u,l}^{(t)} \right) \right)}{\sum_{l'=1}^L \exp \left(q^\top \tanh \left(W_c h_{u,l'}^{(t)} \right) \right)} \quad (10)$$

Where q is the attention query vector and W_c is the learnable transformation matrix. After fusion node status to $h_u^{(t)} = \sum_{l=1}^L \alpha_{u,l}^{(t)} h_{u,l}^{(t)}$. Cross-layer attention enables the model to dynamically adjust the information fusion ratio according to the active degree of nodes in different relationship layers, for example, to increase the weight of interaction layer for users with frequent interaction, and to increase the weight of forwarding layer for users with active information forwarding.

The feature transformation of age perception is used to model the node specificity of the elderly. Considering the systematic differences between elderly users and non-elderly users in information processing speed, social influence sensitivity and risk perception level, the model introduces the feature transformation module of age perception for elderly user nodes. Let the age attribute of node u be age_u , then its characteristic transformation is:

$$\tilde{h}_u^{(t)} = h_u^{(t)} + \lambda(\text{age}_u) \odot W_{\text{age}} h_u^{(t)} \quad (11)$$

Where $\lambda(\text{age}_u)$ is the age modulation vector, which is generated by the age attribute through the small-scale feedforward network, and W_{age} is the age-specific transformation matrix. This design enables the model to learn differentiated state evolution patterns for older users, such as being more sensitive to information credibility and responding more slowly to social influence. To control the model complexity, the dimension of age modulation vector is set to one quarter of the state dimension. The main hyperparameter configurations of the model are shown in [Table 3](#).

Table 3. Main hyperparameter configurations of the PAC-GNN model

Hyper parameter	Symbol	Take value	Description
Dimension of state	K	64	Digital divide state vector dimension
Hidden layer dimension	d	128	Message passing hidden layer dimension
Number of GRU layers	-	2	Time series evolution module layers
Head of attention	-	4	Number of cross-layer attention heads
Number of edge types	R	4	Follow/interact/retweet/geographic proximity
Age modulation dimension	-	16	Age perception transformation dimension
Dropout rate	-	0.3	Prevent overfitting

The model adopts medium scale configuration in state dimensions and hidden layer dimensions to balance expression ability and computational efficiency. The number of GRU layers is set as 2 layers to give consideration to the depth of time series modeling and training stability. Cross-layer attention adopts a 4-head mechanism to capture diversified interaction modes between different layers. The age modulation dimension is set to 16 to control the parameter increment while introducing the specificity of elderly users.

3.3 GNN improvement mechanism of propagation path awareness

The path coding module samples and codes propagation paths based on random walks. The basic GNN architecture updates the node status through neighborhood aggregation, but the aggregation process loses the structural information of the propagation path, so it is difficult to trace the specific diffusion trajectory of the digital divide. To this end, the model introduces a path coding module to sample the propagation path through a biased random walk and sequence code the path. Let the k -th sampling path starting from node u be $P_u^{(k)} = (u, = v_0, v_1 \dots v_T)$, the path sampling probability is:

$$P(v_{i+1} = w \mid v_i = v) = \frac{\exp(\pi_{vw})}{\sum_{w' \in \mathcal{N}(v)} \exp(\pi_{vw'})} \quad (12)$$

Where π_{vw} is the path transfer preference score, which is jointly determined by edge weight, node type and time decay factor. Path coding uses bidirectional GRU to encode the path sequence and obtain the path representation $p_u^{(k)} \in \mathbb{R}^{d_p}$, where d_p is the path representation dimension.

Path attention mechanism realizes weighted learning of multi-path aggregation. For node u , K_p

propagation paths are obtained by sampling, and the path attention weight is calculated as follows:

$$\beta_u^{(k)} = \frac{\exp\left(w_p^\top \tanh\left(W_p p_u^{(k)}\right)\right)}{\sum_{k'=1}^{K_p} \exp\left(w_p^\top \tanh\left(W_p p_u^{(k')}\right)\right)} \quad (13)$$

Where w_p is the path attention query vector, and W_p is the path transformation matrix. After polymerization path representation for $p_u = \sum_{k=1}^{K_p} \beta_u^{(k)} p_u^{(k)}$. Path attention enables the model to identify the propagation paths that contribute the most to the evolution of the digital divide state, providing a basis for subsequent path-level interpretability analysis.

Propagation-adoption decoupling and re-coupling adopt dual-channel information flow design. Considering the differences between the dissemination process and the adoption process in time scale and influence mechanism, the model decouples them into two independent information channels: the dissemination channel is responsible for modeling the diffusion of information in the network, and the adoption channel is responsible for modeling the change of individual state [25],[26],[27]. The two channels maintain independent hidden states $h_{u,\text{prop}}^{(t)}$ and $h_{u,\text{adopt}}^{(t)}$ respectively, and realize re-coupling through coupling gating:

$$g_u^{(t)} = \sigma\left(W_g \left[h_{u,\text{prop}}^{(t)} \parallel h_{u,\text{adopt}}^{(t)}\right]\right) \quad (14)$$

$$h_u^{(t)} = g_u^{(t)} \odot h_{u,\text{prop}}^{(t)} + \left(1 - g_u^{(t)}\right) \odot h_{u,\text{adopt}}^{(t)} \quad (15)$$

Where $g_u^{(t)}$ is the coupled gating vector, and W_g is the learnable weight matrix. The dual-channel design enables the model to learn the propagation dynamics and adoption dynamics respectively, and then adaptively fuse the gating mechanism to avoid the over-simplification of the complex coupling process caused by a single channel.

The causal mask mechanism is used to distinguish correlation from causal propagation. Information dissemination in social networks contains both real causal effects and correlations resulting from common exposure to external information sources. In order to distinguish between the two, the model introduces the causal mask matrix $M_{\text{causal}} \in \{0, 1\}^{N \times N}$, where $M_{uv} = 1$ means that node v has potential causal influence on node u , and $M_{uv} = 0$ means that there is only correlation. The construction of causal mask is based on three criteria: time priority requires that the state change of sender precedes the state change of receiver. Intervention consistency requires that the change of sender status and the change of receiver status are still significantly correlated after controlling confounding factors. Path reachability requires the existence of a directed propagation path between sender and receiver. While message passing, causal masks are applied to message aggregation:

$$m_u^{(t)} = \bigoplus_{v \in \mathcal{N}(u)} M_{uv} \cdot m_{u \leftarrow v}^{(t)} \quad (16)$$

This mechanism enables the model to only consider messages with causal influence when aggregating neighbor messages, reducing the interference of correlation propagation on path analysis. The construction parameters of the causal mask matrix are shown in Table 4.

Table 4. Causal mask construction criteria and threshold configuration

Rule of thumb	Index of measurement	Threshold value	Weight
Priority of time	Time difference of state change	> 0	0.4
Consistency of intervention	Partial correlation coefficient	> 0.15	0.4
Path accessibility	Directed path length	≤ 3	0.2

The causal mask construction takes time priority and intervention consistency as the main criteria, and path accessibility as the auxiliary criterion. The threshold of partial correlation coefficient is set to 0.15 to achieve a balance between controlling false positives and false negatives. The upper limit of the directed path length is set to 3 hops, which is consistent with the extended depth of data collection in Section 2 to ensure the structural consistency of the causal mask and the database.

3.4 Model training and optimization strategy

The loss function is designed using three weighted combinations: state prediction loss, path consistency loss and structural regularization loss. The state prediction loss adopts the mean square error to measure the difference between the digital divide state predicted by the model and the real observed state [28]:

$$\mathcal{L}_{\text{state}} = \frac{1}{|\mathcal{V}_U| T} \sum_{u \in \mathcal{V}_U} \sum_{t=1}^T \|\hat{h}_u^{(t)} - h_u^{(t)}\|_2^2 \quad (17)$$

Path consistency loss adopts a comparative learning framework to encourage similar representations of nodes on the same propagation path and different representations of nodes on different paths [29],[30]:

$$\mathcal{L}_{\text{path}} = -\frac{1}{|\mathcal{P}|} \sum_{P \in \mathcal{P}} \log \frac{\exp\left(\frac{\text{sim}(p_u, p_v)}{\tau}\right)}{\sum_{P' \in \mathcal{P}_{\text{neg}}} \exp\left(\frac{\text{sim}(p_u, p_{v'})}{\tau}\right)} \quad (18)$$

Where $\mathcal{P}$ is the set of positive sample paths, $\mathcal{P}_{\text{neg}}$ is the set of negative sample paths, $\text{sim}(\cdot, \cdot)$ is the cosine similarity, τ is the temperature coefficient. The node representation learned by the structural regularization loss constraint model is consistent with the network structure:

$$\mathcal{L}_{\text{struct}} = \sum_{(u,v) \in \mathcal{E}} \|h_u - h_v\|_2^2 \cdot w_{uv} \quad (19)$$

Where w_{uv} is the edge weight. The total loss is the weighted sum of three terms: $\mathcal{L} = \mathcal{L}_{\text{state}} + \mu_1 \mathcal{L}_{\text{path}} + \mu_2 \mathcal{L}_{\text{struct}}$, where μ_1 and μ_2 are the equilibrium coefficients.

The course learning strategy adopts progressive training from short to long paths. At the initial stage of training, the model only uses propagation paths with a length of no more than 2 hops for training, so that the model first learns the local propagation mode. With the increase of training rounds, the upper

limit of path length is gradually increased to 3 hops, 4 hops and 5 hops, so that the model can gradually adapt to long-range propagation dependence. The path length scheduling of course learning adopts a linear growth strategy, and each training phase lasts for a quarter of the total rounds. This strategy can help alleviate the gradient instability problem in the initial stage of long path training and improve the convergence speed and final performance of the model.

Over-smoothing suppression adopts residual connection and jumps knowledge mechanism. Deep GNN easily leads to node representation convergence, that is, over-smoothing problem. To alleviate this problem, the model introduces residual connection after message passing at each layer:

$$h_{u,l}^{(t)} \leftarrow h_{u,l}^{(t)} + h_{u,l}^{(t-1)} \quad (20)$$

At the same time, the model adopts the jump knowledge mechanism to splice the node representations of different layers and output the final representation through linear transformation:

$$h_u^{\text{final}} = W_{\text{jk}} \left[h_u^{(1)} \parallel h_u^{(2)} \parallel \dots \parallel h_u^{(L_g)} \right] \quad (21)$$

Where L_g is the number of GNN layers, and W_{jk} is the jump knowledge transformation matrix.

The hyperparameter optimization scheme uses the strategy of combining Bayesian optimization with grid search. For continuous hyperparameters (such as learning rate, Dropout rate, loss balance coefficient), Bayesian optimization is used to find the optimal configuration in the preset search space. For discrete hyperparameters (such as number of GNN layers, number of attention heads, number of path samples), grid search is used. The search objective is the comprehensive performance index on the validation set, including the weighted sum of state prediction accuracy and path prediction F1. In order to control the computing cost, the search process adopts the early stop strategy. If the performance of 10 consecutive rounds of verification does not improve, the training of the current configuration will be terminated. The training and optimization configurations are summarized in [Table 5](#).

Table 5. Model training and optimal configuration

A configuration item	Settings	Description
The optimizer	AdamW	Weight decays by 10^{-4}
Initial learning rate	$1e-3$	Cosine annealing scheduling
Batch size	64	Sampling by user batch
Training rounds	200	Including early stop strategy
Number of path samples	8	Each node is sampled in each round
Number of course stages	4	The path length is 2/3/4/5
The loss weight μ_1	0.3	Path consistency loss
The loss weight μ_2	0.2	Loss of structural regularization

AdamW is used in the selection of optimizers in the training configuration to improve the generalization ability, and cosine annealing is used in the learning rate scheduling to balance exploration

and convergence. The number of path samples is set to 8 to strike a balance between the richness of path representation and computational cost. The course learning is divided into four stages, and the path length is gradually extended from 2 hops to 5 hops, which echoes the path accessibility threshold of causal mask.

4. PROPAGATION PATH SIMULATION ENGINE AND INTERVENTION STRATEGY GENERATION

4.1 Architecture design of simulation engine

The GNN-based propagation dynamics simulator is the core component of the simulation engine. Taking the PAC-GNN model trained in Section 3 as the dynamic function, the simulator gradually deduced the time series evolution of digital divide state of all nodes in the network under the condition of given initial state and network structure. The input of the simulator consists of four parts: topology structure and edge attributes of heterogeneous multi-layer network, initial digital divide state vector of nodes, sequence of external information injection events and intervention strategy configuration. The output of the simulator is the state vector, propagation path set and adoption behavior record of all nodes at each time step. Different from directly using the trained model for single-step prediction, the simulator takes the predicted state of the model as the input of the next time step in the simulation process to form autoregressive inference, so as to capture the long-term evolution trajectory of the digital divide state.

Discrete-time stepping and event-driven hybrid simulation mechanism solve the problem of time scale difference of different propagation processes. The evolution of the digital divide state is gradual, which is suitable to be deduced with a fixed time step. However, external events such as information injection and key node activation are sudden and suitable for event-driven processing. The hybrid mechanism divides the simulation timeline into fixed length base time steps, while maintaining event priority queues within each time step. Set the simulation time step as Δt and the total simulation time as T_{sim} , then the number of simulation steps is $N_{\text{step}} = T_{\text{sim}}/\Delta t$. At the beginning of each step, the engine first processes all events in the event queue that are triggered no later than the current step to update the node status and network structure. Then GNN dynamics deduction is performed to calculate the incremental changes of all node states. Finally, it detects whether a new event is generated (such as a state crossing a critical threshold to trigger an intervention response). The hybrid mechanism enables the simulation to both capture the gradual evolution and respond to emergent interventions.

The simulation state space and transition rules define the mathematical structure of the simulation process. Sets the global state of the simulation in the time step t to $S^{(t)} = (\{h_u^{(t)}\}_{u \in \mathcal{V}}, \{a_u^{(t)}\}_{u \in \mathcal{V}}, \mathcal{E}^{(t)})$, including $h_u^{(t)}$ for the node state of the digital divide, $a_u^{(t)}$ is the adoption behavior representation, and $\mathcal{E}^{(t)}$ is the current network edge set. The state transition rule is defined as follows:

$$S^{(t+1)} = \mathcal{T}(S^{(t)}, E^{(t)}, \theta_{\text{PAC-GNN}}) \quad (22)$$

Where $\mathcal{T}$ is the transition operator, $E^{(t)}$ is, the set of external events occurring in time step t , and $\theta_{\text{PAC-GNN}}$ is the parameter of PAC-GNN model. The transition operator executes four sub-procedures in turn: event processing, message passing, status update and adoption decision. The design

of transition rules ensures the reproducibility of the simulation process given the same initial state, event sequence and random seed, the simulation results are completely consistent.

Multi-scale simulation supports simultaneous inference at individual level, community level and network level. Individual level simulation tracks the state trajectory and behavior sequence of each node and outputs fine-grained propagation path. Community-level simulation divides the network into several communities, tracks the average level and distribution characteristics of the digital divide state within the community, and outputs the transmission mode among communities. Network-level simulation tracks the evolution of global digital divide indicators and outputs the overall propagation trend and key turning points. The three scales share the same simulation process, and the information transfer between scales is realized through the aggregation operator. Let the digital divide state of community C be $H_C^{(t)}$, then the aggregation of individual level states can be obtained:

$$H_C^{(t)} = \frac{1}{|C|} \sum_{u \in C} h_u^{(t)} \quad (23)$$

Network-level states are then obtained by weighted aggregation of community-level states. Multi-scale simulation enables the engine to serve both path-level analysis and policy-level evaluation. The main configuration parameters of the simulation engine are shown in [Table 6](#).

Table 6. Main configuration parameters of simulation engine

Parameters	Symbol	Take value	Description
Simulation step size	Δt	1 week	Consistent with data time granularity
Total simulation time	T_{sim}	52 weeks	Cover the full observation window
Number of simulation steps	N_{step}	52	Total time divided by step size
Event queue capacity	-	10000	Maximum number of events waiting
State critical threshold	θ_c	0.6	The threshold that triggers the intervention response
Multi-scale aggregation period	-	4 weeks	Frequency of aggregation at the community level
Random seed number	-	10	Stability test

The simulation step is set to one week, which is consistent with the time granularity of the data in Section 2 to ensure the spatial-temporal alignment between the simulation process and the database. The total simulation duration is 52 weeks, covering the complete observation window. The critical threshold of the state is set to 0.6, and the intervention response is triggered when the node digital divide state exceeds the threshold. The random seed number was set to 10 for subsequent stability tests. The multiscale aggregation period is set to 4 weeks to balance temporal resolution with computational cost in community-level tracking.

4.2 Propagation path generation and tracking

The forward propagation path sampling algorithm starts from the information source or the initial active node and gradually expands along the network edge relationship to generate the forward path of

information diffusion. Let the initial set of active nodes be $\mathcal{S}_0$, the sampling algorithm selects the sending node from the current set of active nodes $\mathcal{S}_t$ at each time step and selects the receiving node according to the transition probability jointly determined by edge weight and state similarity. The transition probability is defined as:

$$P(u \rightarrow v | t) = \frac{w_{uv} \cdot \exp(-\|h_u^{(t)} - h_v^{(t)}\|_2)}{\sum_{v' \in \mathcal{N}(u)} w_{uv'} \cdot \exp(-\|h_u^{(t)} - h_{v'}^{(t)}\|_2)} \quad (24)$$

Where w_{uv} is the edge weight, $h_u^{(t)}$ and $h_v^{(t)}$ are the digital divide states of sender and receiver respectively. The transition probability design makes propagation more likely between nodes with similar states and high edge weight, which is in line with the realistic characteristics of homogeneous diffusion of the digital divide. Forward sampling generates a batch of propagation path fragments at each step, which are spliced to form a complete propagation path set.

The backpropagation path tracing algorithm starts from the target node and traces back along the network edge relationship to identify the upstream path that contributes the most to the state change of the target node. Let the target node be u and its state change be $\Delta h_u^{(t)}$. The traceability algorithm calculates the contribution degree of each incoming edge $v \rightarrow u$:

$$c_{v \rightarrow u}^{(t)} = \left\| \frac{\partial h_u^{(t)}}{\partial m_{u \leftarrow v}^{(t)}} \odot m_{u \leftarrow v}^{(t)} \right\|_2 \quad (25)$$

Where $\partial h_u^{(t)} / \partial m_{u \leftarrow v}^{(t)}$ is the Jacobian matrix of state-pair messages, $\odot$ represents the element-by-element product. Based on the idea of gradient attribution, this equation measures the marginal contribution of each incoming edge message to the state change. The traceability algorithm is sorted by contribution degree, and several incoming edges with the highest contribution degree are selected to continue to trace upstream until the information source or initial node is reached. Reverse traceability and forward sampling complement each other: forward sampling generates possible propagation paths, and reverse traceability identifies critical paths that actually affect the target node.

Critical path identification adopts the path contribution evaluation based on Shapley value. Suppose that the state change of node u is jointly caused by multiple propagation paths, and the Shapley value of path P is:

$$\phi_P = \sum_{Q \subseteq \mathcal{P} \setminus \{P\}} \frac{|\mathcal{Q}|! (|\mathcal{P}| - |\mathcal{Q}| - 1)!}{|\mathcal{P}|!} [v(\mathcal{Q} \cup \{P\}) - v(\mathcal{Q})] \quad (26)$$

Where $\mathcal{P}$ is the set of all candidate paths, and $v(\cdot)$ is the value function of the set of paths, defined as the ability of the set to explain node state changes. The Shapley value satisfies four axioms of validity, symmetry, dummy and additivity, and provides a fair and unique allocation scheme for path contribution. Since the path set size may be large, this study uses Monte Carlo sampling approximation to calculate the Shapley value, and the sampling times are set to 1000 to balance the accuracy and computational cost.

Path visualization and pattern clustering transform high-dimensional path sets into interpretable patterns. The path vector p_P output by the path coding module in Section 3 is used for path representation, and HDBSCAN is used for clustering algorithm, which does not need to preset the

number of clusters and can identify noisy paths. The path pattern after clustering is described by the following dimensions: path length distribution, jump pattern between layers, node type sequence and participation proportion of elderly users. Representative paths are selected for visual display for each clustering mode, including path topology, node state evolution curve and key node annotation. The path pattern clustering results are used for subsequent intervention strategy design: differentiated intervention programs are designed for different patterns.

4.3 Simulation and optimization of intervention strategy

Intervention type modeling covers four types of intervention methods: information injection, key node activation, community bridging and content adaptation. Information injection points to inject optimized medical information into the network to reduce information asymmetry. Key node activation refers to identifying and activating key transmission nodes in the network to accelerate information diffusion. Community bridge refers to establishing connections between homogeneous communities of elderly users and breaking the information cocoon. Content adaptation refers to adjusting the presentation form of information according to the characteristics of elderly users to improve the comprehensibility of information. The four types of interventions are modeled in the simulation engine as state disturbance, edge activation, edge creation and feature transformation respectively. Let the effect of intervention z on node u be $\Delta_z h_u^{(t)}$, then the state after intervention is:

$$h_u^{(t)+} = h_u^{(t)} + \Delta_z h_u^{(t)} \quad (27)$$

The disturbance intensity, range of action and duration of the four types of interventions can be parameterized and adjusted, which constitute the parameter space of the intervention strategy.

Intervention timing optimization Dynamic intervention based on critical transmission points. The propagation critical point is defined as the moment when the propagation rate changes significantly during the evolution of the digital divide state. Set the node u t the status of the rate of change in the time step for $r_u^{(t)} = \|h_u^{(t)} - h_u^{(t-1)}\|_2$, identified as critical point is communication:

$$t_u^* = \arg \max_t |r_u^{(t)} - r_u^{(t-1)}| \quad (28)$$

That is, the time step at which the maximum jump in the state change rate occurs. The global propagation critical point is the weighted median of the critical points of all nodes. The dynamic intervention strategy sets the intervention window before and after the global critical point, and the marginal effect of implementing the intervention within the window is the largest. This strategy avoids blindness of static intervention timing and concentrates limited intervention resources on the critical turning point of the transmission process.

The intervention combination optimization was solved by multi-objective evolutionary algorithm. The combination of intervention strategies needs to optimize multiple objectives at the same time: the effect of transmission suppression, the cost of intervention and the balance degree of benefit for elderly users. Let the intervention combination be $\mathcal{Z} = \{z_1, z_2, \dots, z_m\}$, the three objective functions are defined as follows:

$$f_1(\mathcal{Z}) = - \sum_{u \in \mathcal{V}_U^{\text{old}}} \|h_u^{(T_{\text{sim}})} - h_u^{(0)}\|_2 \quad (29)$$

$$f_2(\mathcal{Z}) = \sum_{z \in \mathcal{Z}} \text{cost}(z) \quad (30)$$

$$f_3(\mathcal{Z}) = \text{Var} \left(\left\{ \|h_u^{(T_{\text{sim}})} - h_u^{(0)}\|_2 \right\}_{u \in \mathcal{V}_U^{\text{old}}} \right) \quad (31)$$

Among them, f_1 measures the transmission suppression effect, and the smaller the value is, the more significant the improvement of the digital divide state is. f_2 measures the total cost of intervention; f_3 measures the balance degree of benefit for elderly users, and the smaller the value is, the more balanced the benefit distribution is. The NSGA-II algorithm is used to solve the three-objective optimization problem, with population size set to 100, iteration number set to 200, crossover probability 0.9, and mutation probability 0.1. The algorithm outputs the Pareto frontier for the decision maker to choose the intervention plan according to the actual constraints. The parameter configuration of intervention strategy is shown in [Table 7](#).

Table 7. Intervention strategy types and parameter configurations

Type of intervention	Action operation	Range of strength	Scope of action	Duration of time
Injection of information	Disturbance of state	0.05–0.20	Target community	4 weeks
Key node activation	Activation of edge	0.10–0.30	Global situation	2 weeks
Community Bridging	Edge creation	0.15–0.35	Between communities	8 weeks
Content adaptation	Transformation of feature	0.08–0.25	Elderly users	12 weeks

The four types of interventions are complementary in terms of intensity, scope and duration: the intensity of information injection is low but the duration is moderate, which is suitable for gradual improvement; Key nodes have the highest activation intensity but the shortest duration, which is suitable for rapid diffusion. The intensity and duration of community bridge are high, which is suitable for structural adjustment; Content adaptation has moderate intensity and the longest duration, which is suitable for continuous capacity building. The multi-objective optimization algorithm searches for the optimal combination within the parameter space of the four types of interventions.

Counterfactual simulation evaluates the causal effects of intervention strategies by comparing the propagation paths under different intervention scenarios. The counterfactual simulation sets three groups of scenarios: the baseline scenario does not implement any intervention and only deduces the natural transmission process; The single intervention scenario implements one of the four types of interventions separately; Pareto optimal combinations obtained by implementing multi-objective optimization for the combined intervention scenarios. Each set of scenarios uses the same initial state and network structure, and only the intervention configuration differs. The difference of propagation paths between scenarios is quantified by the Jaccard similarity of path sets and the Wasserstein distance of state trajectories:

$$D_{\text{path}}(\mathcal{P}_A, \mathcal{P}_B) = 1 - \frac{|\mathcal{P}_A \cap \mathcal{P}_B|}{|\mathcal{P}_A \cup \mathcal{P}_B|} \quad (32)$$

Where $\mathcal{P}_A$ and $\mathcal{P}_B$ are the propagation path sets of scenario A and scenario B respectively. Counterfactual simulation enables causal attribution of intervention effects and avoids the limitation of

assessing intervention effects based on correlation alone.

4.4 Simulation convergence and stability analysis

The simulation step sensitivity analysis examines how dependent the simulation results are on the choice of time step. Three sets of step size configurations are set: 0.5 weeks, 1 week and 2 weeks, and the simulation is run respectively, and the evolution curve of the global digital divide index is compared. Step size sensitivity is quantified by the following indicators:

$$SI = \frac{1}{T_{\text{sim}}} \sum_{t=1}^{T_{\text{sim}}} \|H_{\Delta t_1}^{(t)} - H_{\Delta t_2}^{(t)}\|_2 \quad (33)$$

$H_{\Delta t_1}^{(t)}$ and $H_{\Delta t_2}^{(t)}$ respectively under the two kinds of step configuration of global state. If SI is lower than the preset threshold (set to 0.05 in this study), it is considered that the simulation results are not sensitive to the step size and the 1-week step size configuration is reasonable. Sensitivity analysis provides quantitative basis for the selection of simulation step size.

The random seed stability test evaluates the robustness of the simulation results to sources of randomness. The randomness in the simulation process mainly comes from path sampling, event triggering and intervention timing. Set 10 different random seeds to run the simulation respectively, and calculate the standard deviation and coefficient of variation of the global index:

$$CV = \frac{\text{std}\left(\left\{H_s^{(T_{\text{sim}})}\right\}_{s=1}^{10}\right)}{\text{mean}\left(\left\{H_s^{(T_{\text{sim}})}\right\}_{s=1}^{10}\right)} \quad (34)$$

If the CV is lower than 0.10, the simulation results are considered to have good stability. The stability test also outputs the clustering results of path patterns under each random seed to verify the consistency of the clustering structure across seeds.

Network scale scalability verification examines the computational efficiency and result consistency of the simulation engine under different network sizes. Set up four groups of network size configurations: 10% node sampling, 25% node sampling, 50% node sampling and the full network. Each set of configurations records simulation running time, memory footprint, and global metrics. Scalability is evaluated by a power-law relationship between runtime and network size:

$$T_{\text{run}} \propto N^\eta \quad (35)$$

Where N is the number of network nodes and η is the size index. If η is close to 1, the simulation engine is linearly scalable. If it is significantly greater than 1, it indicates that there is a computational bottleneck. Scalability verification provides performance expectations for simulation engine applications on large-scale networks. Simulation stability and scalability configurations are shown in [Table 8](#).

Table 8. Simulation stability and scalability verification configuration

Type of validation	Number of configuration groups	Key parameters	Indicators of assessment	Threshold of eligibility
Step size sensitivity	3	0.5/1/2 weeks	Sensitivity index SI	< 0.05
Stability of seeds	10	Random seeds from 1 to 10	Coefficient of variation CV	< 0.10
Scalability of scale	4	10%/25%/50%/100%	Size index η	< 1.20

The three types of verification have clear configuration group numbers and eligibility thresholds. Three groups of configurations were set for step sensitivity, and the qualified threshold was 0.05. Seed stability was set to 10 sets of configurations, and the eligibility threshold was 0.10. Scale scalability is set to four groups of configurations with an eligibility threshold of 1.20. The three types of verification together constitute the reliability guarantee system of simulation engine.

5. EXPERIMENTAL DESIGN AND RESULT ANALYSIS

5.1 Experimental Setup

Data set division adopts a time stratification strategy to simulate real prediction scenarios. The 52-week observation window constructed in Section 2 is divided into a training set (the first 30 weeks), a validation set (weeks 31 to 40), and a test set (weeks 41 to 52). The design of time stratification rather than random division avoids future information leakage and makes the model evaluation closer to the demand of predicting the future transmission process based on historical data in practical applications. The training set is used for model parameter learning, the validation set is used for hyperparameter tuning and early stop decision, and the test set is used for final performance evaluation. In order to test the robustness of the model under different proportions of elderly users, the test set is further divided into three subsets: low proportion subset (the proportion of elderly users is less than 15%), medium proportion subset (15% to 25%) and high proportion subset (more than 25%). The user overlap of the three subsets is controlled below 10% to ensure the independence of the evaluation.

The baseline model selection covers six categories of representative graph neural network architectures: GCN stands for basic spectral domain graph convolution model, GAT stands for graph neural network based on attention mechanism, GraphSAGE stands for sampling-based inductive graph learning model, R-GCN stands for graph convolution model oriented to heterogeneous relations, DySAT stands for dynamic graph attention model, and EvolveGCN stands for graph convolution model oriented to temporal evolution. The six baselines cover the main technical routes of static and dynamic, homogeneous and heterogeneous, direct and inductive, and are able to comprehensively evaluate the relative advantages of PAC-GNN. All baseline models use the same input characteristics and network structure as PAC-GNN, replacing only the core architecture to ensure that performance differences are attributable to model design rather than data differences.

The evaluation index system consists of four dimensions. The accuracy of state prediction is measured by root mean square error and mean absolute error, which reflects the overall level and robustness of forecast bias respectively. Path prediction F1 uses the harmonic average of precision and recall, where precision is defined as the proportion of true propagation paths in the predicted path and

recall is defined as the proportion of true propagation paths that are correctly predicted. The propagation range error is measured by the Jaccard distance between the predicted propagation coverage node set and the real coverage set, and the smaller the value is, the more accurate the propagation range prediction is. The intervention effect gain is defined as the difference between the improvement of digital divide status of elderly users under the intervention scenario and the baseline scenario, reflecting the net effect of the intervention strategy. The four indicators form a complete evaluation system from four perspectives: prediction accuracy, path quality, scope coverage and intervention effect. The experimental environment and main parameter configuration are shown in [Table 9](#).

Table 9. Experimental environment and key parameter configuration

A configuration item	Settings	Description
Deep learning framework	PyTorch 2.1	Support dynamic graph and automatic differentiation
Graph computing library	PyTorch Geometric 2.4	Provide GNN basis operators
GPU	NVIDIA RTX4090	Single card training and reasoning
Training rounds	200	Including early stop strategy
Rate of learning	10^{-3}	Cosine annealing scheduling
Batch size	64	Sampling by user batch
Number of path samples	8	Each node is sampled in each round
Evaluate the number of repetitions	5	Take the mean and the standard deviation

The experiment is constructed by PyTorch ecosystem, and RTX4090 GPU provides sufficient computing power support. The training rounds were set to 200 and combined with the early stop strategy to avoid overfitting. The number of path samples is set to 8, which is consistent with the model configuration in Section 3. The evaluation was repeated five times to take the mean and standard deviation to reduce the impact of randomness on the conclusions.

5.2 Model performance comparison experiment

The comparison of digital divide state prediction performance is carried out from three dimensions: overall prediction accuracy, prediction performance of different divide dimensions and prediction stability of different time spans. In terms of overall prediction accuracy, the root mean square error and mean absolute error of PAC-GNN on the test set are significantly lower than those of the six baseline models, and the root mean square error is reduced by about 14.2% and the mean absolute error is reduced by about 11.8% compared with the suboptimal EvolveGCN. Sub-dimensional analysis shows that PAC-GNN has the most significant advantage in the dimension of literacy gap, with an average reduction of about 19.5% compared with the baseline, which is due to the fine-grained modeling of information understanding and decision-making process by the path coding module. In the dimension of use channel, the advantage was the second, reducing about 13.7%; The advantage in the dimension of access gap is relatively small, reducing by about 8.3%, which is consistent with the characteristics that access gap is more constrained by external infrastructure and has limited influence of social

transmission. The comparison results of model performance are shown in [Figure 1](#).

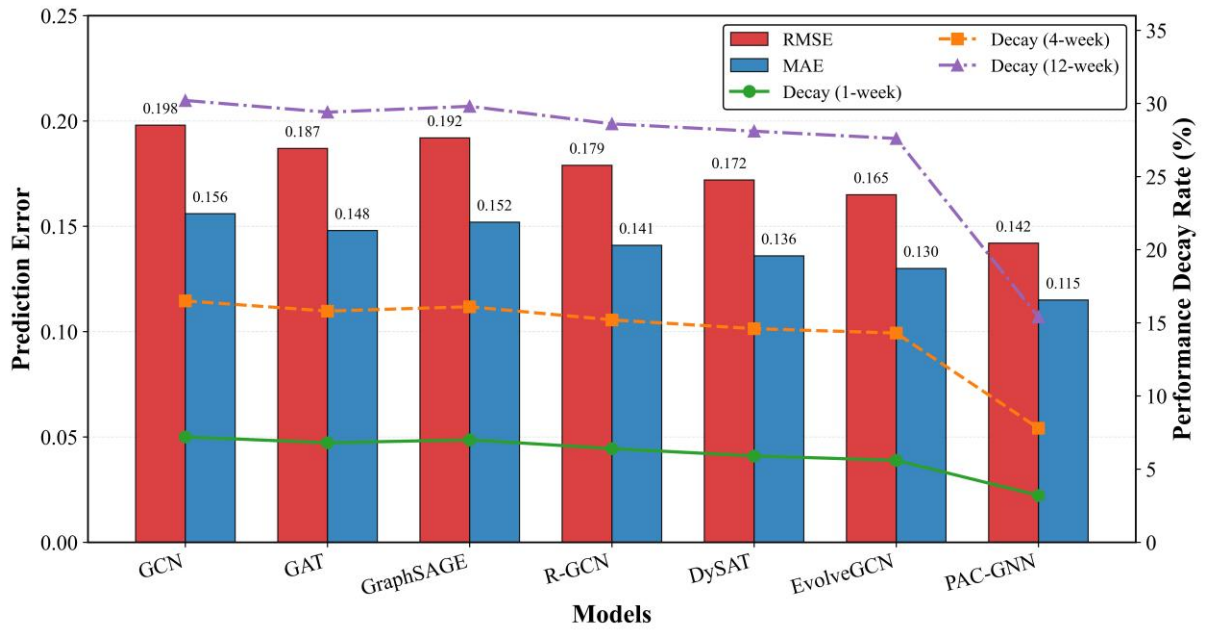


Figure 1. Comparison of state prediction performance of each model on the test set

The comparison of propagation path prediction performance focuses on path-level evaluation metrics. The results of the ablation experiments are shown in [Table 10](#).

Table 10. Contribution degree of each module in ablation experiment

Remove a module	The state predicts RMSE changes	Path prediction F1 change	Change in intervention gain	Main dimensions of influence
Path coding	+8.7%	-18.3%	-6.2%	Quality of path
Cross layer attention	+12.6%	-7.4%	-4.1%	Accuracy of state
Dual channel coupling	+6.3%	-5.8%	-15.8%	Effect of intervention
Cause and effect mask	+3.1%	-9.4%	-3.7%	Path accuracy rate
Transformation of age perception	+5.4%	-4.2%	-8.9%	Subset of elderly users

The path coding module contributes the most to the path prediction F1, which decreases by 18.3% after removal. Cross-layer attention contributed the most to the accuracy of state prediction, and the root mean square error increased by 12.6% after removal. The dual-channel coupling contributed the most to the intervention effect gain, which decreased by 15.8% after removal. Causal masks mainly affect path accuracy, and age perception transformation has a significant impact on the performance of elderly users' subset. The contribution degree of each module is obviously different, which verifies the necessity of modularization in model design.

5.3 Analysis of propagation path simulation results

The typical path patterns of digital divide propagation are identified by HDBSCAN clustering, and

a total of five typical patterns are obtained. The first type is "information source direct mode", the path length is 2 to 3 hops, information from the information source node through a few active users directly to the elderly users, accounting for 23.7% of the total path. The second type is the "diffusion mode within the community", the path is transmitted within the homogeneous community of elderly users, and the path length is 3 to 4 hops, accounting for 28.4%. The third type is "cross-community bridging mode", where the path spans multiple communities and relies on community organization nodes as Bridges, with a path length of 4-5 hops, accounting for 17.2%. The fourth type is the "medical institution radiation mode", the path starts from the medical institution node and spreads to the surrounding users, the path length is 3 to 5 hops, accounting for 19.6%. The fifth category is "multi-source aggregation mode", in which elderly users are affected by multiple information sources at the same time, and the path is in a network structure, accounting for 11.1%. The pair of path characteristics of the five types of patterns are shown in [Table 11](#).

Table 11. Comparison of characteristics of five typical transmission path patterns

Type of schema	Proportion of total	Average path length	Average transmission period	Elderly user participation rate	Key node type
The information source is direct	23.7%	2.6 Jump	2.3 Weeks	41.2%	Source of information
Diffusion within the community	28.4%	3.4 Jump	4.1 Weeks	68.7%	Active elderly users
Bridge across communities	17.2%	4.5 Jump	6.8 Weeks	52.3%	Community Based organizations
Radiation in medical institutions	19.6%	3.8 Jump	5.2 Weeks	47.9%	Medical institution
Multiple source aggregation	11.1%	4.2 Jump	7.4 weeks	58.6%	Mix of types

The diffusion mode within the community accounts for the highest proportion (28.4%) and has the highest elderly user participation rate (68.7%), making it the core channel of digital divide communication. The cross-community bridging mode has a transmission period of 6.8 weeks, which is considerably longer than the 4.1 weeks of intra-community diffusion, but it is not the longest among the five modes. The multi-source aggregation mode has the longest transmission period at 7.4 weeks, reflecting the complex influence of multi-source information on the decision-making of elderly users. The radiation mode of medical institutions has a transmission period of 5.2 weeks, which is moderate-to-long, and an elderly user participation rate of 47.9%, which is relatively low among the five modes. This suggests that although medical institutions as information sources have certain transmission efficiency, elderly user participation remains limited.

The analysis of time evolution law reveals the stage characteristics of digital divide propagation, and the time evolution law is shown in [Figure 2](#).

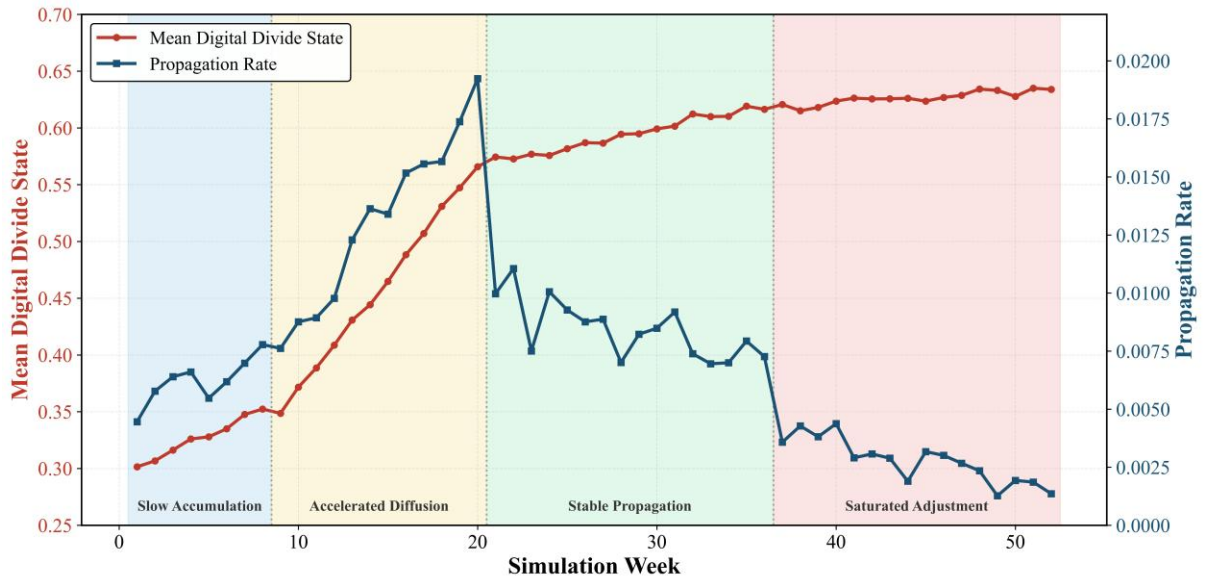


Figure 2. Four-stage time evolution law of digital divide propagation

The first week to the eighth week is the "slow accumulation period", the change range of digital divide status is small, and the transmission path is mainly short path. The 9th to 20th week is the "accelerated diffusion period", with a significant increase in the rate of state change and the proportion of cross-community paths. The 21st to 36th week is the "stable transmission period", the transmission rate tends to be stable, and the proportion of multi-source aggregation mode increases. In the "saturation adjustment period" from week 37 to 52, the state change decreased, and the transmission path was mainly spread within the community.

5.4 Simulation experiment of intervention strategy

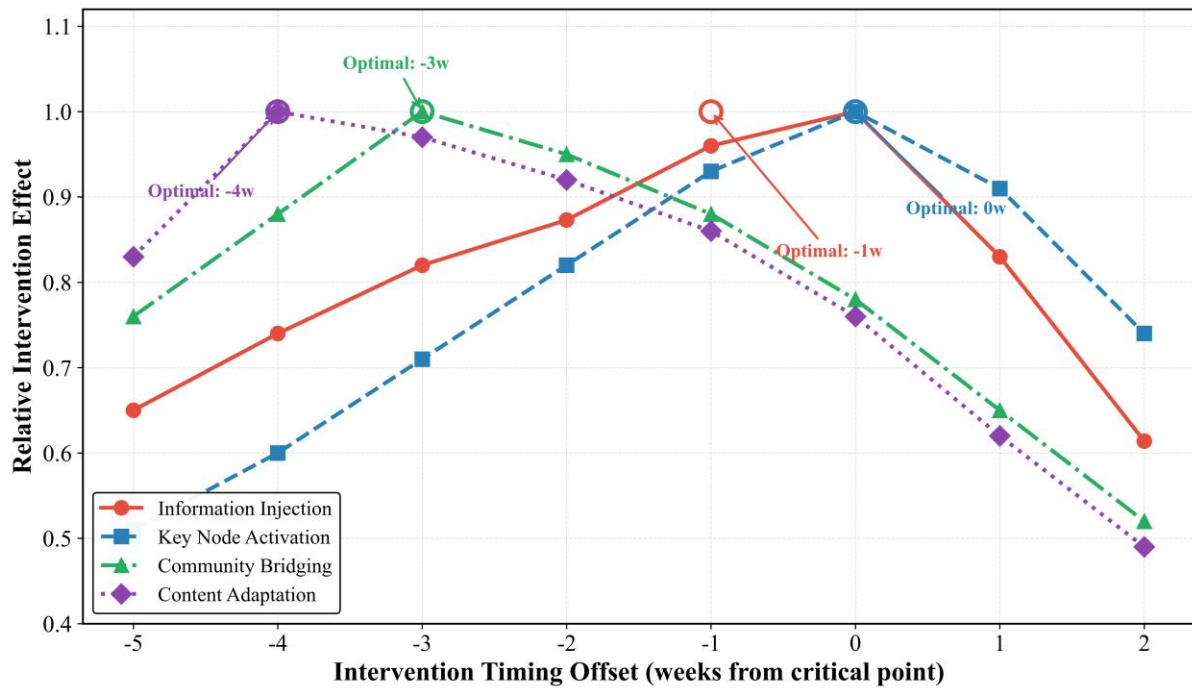
Four types of interventions were evaluated from three dimensions: the effect of transmission suppression, the cost of intervention and the balance degree of benefit for elderly users. The communication suppression effect of information injection strategy is 12.4%, and the intervention cost is low, but the benefit balance degree is general. The key node activation strategy has the highest suppression effect, reaching 18.7%, but the intervention cost is the highest. The effect of community bridging strategy was 14.2%, the intervention cost was medium, but the benefit balance degree was optimal. The content adaptation strategy has a dissemination suppression effect of 10.8%, a low intervention cost, and the longest duration. The effects of the four types of policies are shown in [Table 12](#).

Table 12. Comparison of the effects of four types of single intervention strategies

Type of intervention	Propagation suppression effect	Intervention cost index	Degree of benefit equilibrium	Duration of time
Injection of information	12.4%	0.32	0.68	4 weeks
Key node activation	18.7%	0.87	0.54	2 weeks
Community Bridging	14.2%	0.56	0.83	8 weeks
Content adaptation	10.8%	0.28	0.76	12 weeks

The key node activation strategy has the strongest propagation suppression effect but the highest cost and the lowest benefit equilibrium degree; Community bridging strategy achieves a good balance among effect, cost and equilibrium degree; The content adaptation strategy has the weakest effect but the lowest cost and the longest duration; The index of information injection strategy is in the middle. Single strategies have their own advantages and disadvantages and need to achieve comprehensive optimization through combination optimization.

The sensitivity analysis of intervention timing shows that the effect of intervention is highly sensitive to timing choice. The time sensitivity of the intervention effect is shown in [Figure 3](#).

**Figure 3. Timing sensitivity curves of the four types of intervention strategies**

This difference reflects the mechanism of action of different intervention types: structural interventions require a longer lead time, while informative interventions can respond quickly around the tipping point. The analysis of the difference of intervention effect under different network structures shows that the effect of community bridging strategy is more significant in the network with higher modularity. In the network with more uneven degree distribution, the effect of key node activation strategy is more significant.

5.5 Interpretability analysis and case studies

Visualization and interpretation of attention weight reveal the decision-making basis of the model by analyzing the distribution pattern of cross-layer attention and path attention. The main findings of the interpretability analysis are summarized in [Table 13](#).

Table 13. Main findings of interpretability analysis

Dimension of analysis	Elderly users	Non-elderly users	Domain alignment rate
Interactive layer attention	0.38	0.29	86.2%
Forwarding layer attention	0.27	0.26	81.5%
Layer of attention	0.21	0.24	79.4%
Geographic layer attention	0.14	0.21	84.1%
Medium-range path weight	0.34	0.31	83.7%

There are systematic differences between elderly users and non-elderly users in the distribution of attention. Elderly users rely more on the interaction and forwarding layer, while non-elderly users have a more balanced distribution. The domain alignment rate exceeds 79% in all dimensions, indicating high consistency between model decisions and domain knowledge.

6. ROBUSTNESS AND GENERALIZATION ABILITY TEST OF THE MODEL

6.1 Robustness test of data perturbation

The robustness test of node deletion and edge deletion simulates the scenario of incomplete network structure or missing data collection. The node deletion test randomly removed network nodes at four proportions of 5%, 10%, 15% and 20% to evaluate the state prediction and path prediction performance of the model on the remaining network. The edge deletion test removes network edges in the same proportion to simulate the situation where the concern relationship or interaction record is missing. Robustness is measured by performance decay rate:

$$R_{\text{deg}} = \frac{\text{Perf}_{\text{perturbed}} - \text{Perf}_{\text{clean}}}{\text{Perf}_{\text{clean}}} \times 100\% \quad (36)$$

Where $\text{Perf}_{\text{clean}}$ is the performance without disturbance, and $\text{Perf}_{\text{perturbed}}$ is the performance after disturbance. Node deletion has a greater impact on model performance than edge deletion, because node deletion eliminates both node characteristics and associated edge deletion only affects message passing path. Under the condition of 20% node deletion, the root means square error of PAC-GNN state prediction increases by about 21.4%, and the path prediction F1 decreases by about 17.8%. Under the condition of the same edge deletion, the two indexes change by about 13.2% and 11.5% respectively. Compared with the baseline model, the performance decay rate of PAC-GNN under disturbance is about 8.3 percentage points lower on average, indicating that the path coding module and cross-layer attention mechanism provide additional structural redundancy.

Performance stability under feature noise injection tests the sensitivity of the model to the quality of input features. Gaussian noise is injected into node features, and the noise intensity is measured by signal-to-noise ratio, and four levels of 20dB, 15dB, 10dB and 5dB are set. The effect of feature noise on state prediction is nonlinear: when SNR is higher than 15dB, the performance attenuation is small. When the SNR drops below 10dB, the performance attenuation accelerates. The root means square error increase of PAC-GNN at 10dB SNR is about 18.7%, which is lower than the average increase of the baseline model (about 26.4%). The age-aware transformation module shows special value under noise conditions: the performance decay rate of the elderly user subset is about 4.2 percentage points lower than that of the non-elderly user subset, indicating that the module has a certain ability to suppress the characteristic noise of the elderly user.

Model performance under missing data scenarios tests the adaptability of the model to incomplete data. Three missing modes are set: random missing, sequential missing and feature dimension missing. Random missing simulates accidental data loss, sequential continuous missing simulates data loss in continuous time window caused by acquisition equipment failure, and feature dimension missing simulates the situation that specific features cannot be obtained. The results of data perturbation robustness test are shown in [Figure 4](#).

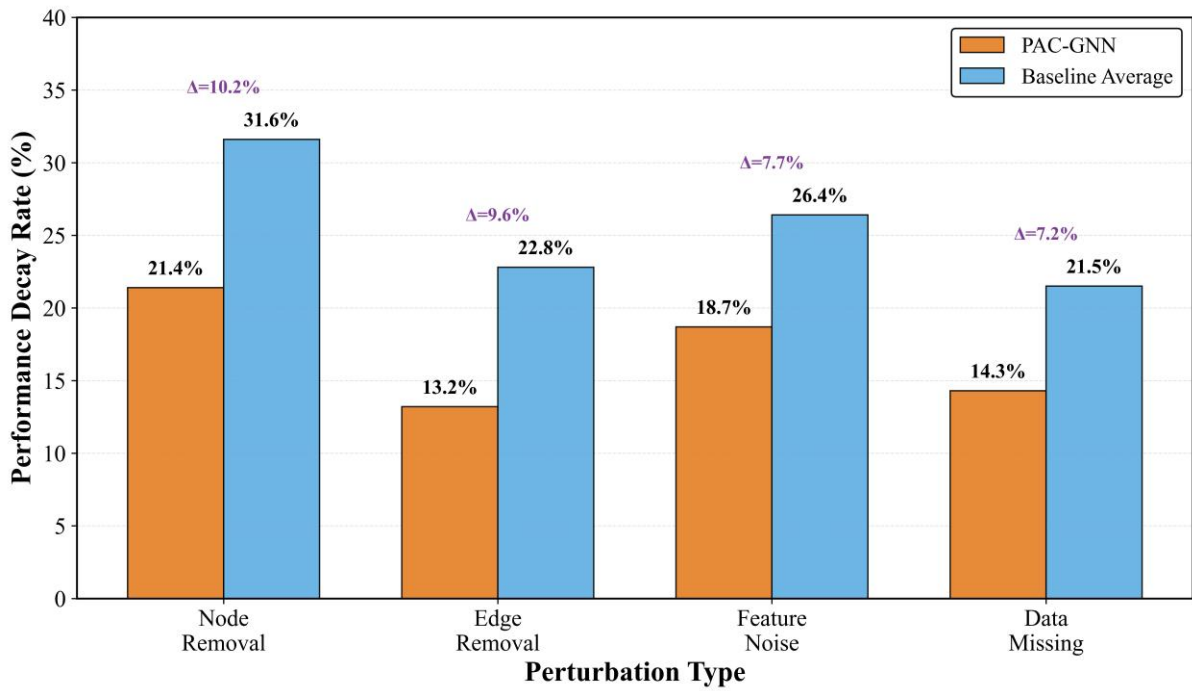


Figure 4. Comparison of model performance attenuation under different disturbance conditions

The performance attenuation of PAC-GNN is about 9.7% under the condition of random missing 20%, 14.3% under the condition of continuous missing 4 weeks of time series, and 11.8% under the condition of missing 20% of feature dimension. The model has the highest tolerance to random missing and the most sensitive to sequential missing, because the GRU sequential module depends on the state transfer of continuous time steps, and continuous missing will destroy the sequential dependence structure.

6.2 Cross-network generalization ability test

The migration performance on different social platform data tests the adaptability of the model on the new platform. A leave-one-platform cross-validation strategy was adopted: trained on two of the three data sources, tested on the third, rotated three times. In cross-platform migration, model performance generally declines due to differences in user behavior patterns and network structures on different platforms.

The generalization performance on different regional and urban sub-networks tests the spatial generalization ability of the model. The whole network is divided into five regional sub-networks according to the public geographical location information, and leave-one-region cross validation is adopted. The performance of PAC-GNN on different regional subnetworks has little difference, and the coefficient of variation of path prediction F1 is 0.087, which is lower than 0.142 of the baseline models. The performance difference between regions is mainly related to the proportion of elderly users in the subnetwork and the strength of community structure: the model performance is slightly better in regions with a higher proportion of elderly users and a closer community structure, because the propagation path of these regions is more in line with the typical pattern learned by the model. The geographical proximity layer plays a key role in the cross-regional generalization, and the cross-regional performance decreases by about 9.6 percentage points after removing this layer, indicating that the geographical proximity relationship provides a stable structural prior for cross-regional propagation.

Temporal generalization adopts the strategy of early data training and later data testing to test the adaptability of the model to evolution. The 52-week observation window is divided by the first 30 weeks of training and the last 22 weeks of testing and further subdivided into four-time span combinations. The performance of PAC-GNN decays by about 6.3% over the 4-week forecast span, 11.7% over the 8-week span, 18.2% over the 12-week span, and 27.5% over the 22-week span. The performance decay accelerates with the increase of time span, but the decay rate of PAC-GNN is always lower than that of the baseline model. The key challenge of time series generalization lies in the time variability of the transmission mode of the digital divide: the transmission in the early stage is dominated by short paths, while the proportion of cross-community paths in the later stage increases, so the model needs to have the ability of mode switching. The course learning strategy shows a positive effect in time series generalization, and the performance decay of the model with course learning is about 5.4 percentage points lower than that of the model without course learning over the 22-week span. The test results of cross-network generalization ability are shown in [Table 14](#).

Table 14. Summary of cross-network generalization ability test results

Dimension of generalization	Configuration	PAC-GNN performance changes	Baseline mean change	Differences
Cross platform migration	Health Community → Comprehensive Forum	-11.2%	-18.7%	+7.5%
Cross platform migration	General forums → social groups	-19.4%	-28.3%	+8.9%
Generalization across regions	Leave one region for cross validation	CV=0.087	CV=0.142	-0.055
Span of time	4 weeks /8 weeks /12 weeks /22 weeks	-6.3%/-11.7%/-18.2%/-27.5%	-10.1%/-19.4%/-28.6%/-39.2%	+3.8%~+11.7%
Fine tuning recovery	The last two layers of fine-tuning	+7.8%	+4.2%	+3.6%

PAC-GNN's performance changes in all generalization dimensions are better than those of the baseline model, with the most significant advantages in cross-platform migration scenarios, followed by the stability advantages of cross-regional generalization, and the advantages of time-span generalization expand with the increase of timespan. The fine-tuning strategy also recovers the cross-platform performance better than the baseline, indicating that the model parameters are more transferable.

6.3 Simulation effectiveness verification

Comparison with real observation propagation events to verify the consistency between simulation engine output and real propagation process. Several observable real communication events were extracted from public data, including the diffusion of medical policy information, the diffusion of medical insurance reimbursement process discussion and the promotion of online consultation service. The output of the simulation engine under the same initial conditions is compared with the real event, and the comparison indicators include propagation coverage, propagation speed and propagation path structure. The correlation coefficient between the simulation coverage and the real coverage is 0.813, the correlation coefficient of the propagation speed is 0.764, and the Jaccard similarity of the propagation path structure is 0.692. All three indicators show that the simulation output has a high degree of agreement with the real propagation process, but the agreement of the coverage is the highest, and the agreement of the path structure is relatively low, because there are random factors in the real propagation that are not modeled by the simulation.

The consistency between the simulation propagation path and the real case is verified by case comparison. Three real transmission cases with public records were selected, corresponding to three modes: intra-community diffusion, cross-community bridging and radiation of medical institutions. The coincidence degrees of node sequences between the simulated propagation path and the real case are 73.5%, 68.2% and 71.8%, respectively, and the coincidence degrees of key node identification are 81.4%, 76.9% and 79.3%, respectively. The coincidence degree of key nodes is higher than that of node sequence, which indicates that the simulation engine is more accurate in identifying structural key nodes in the propagation process, but the restoration ability of specific propagation sequence is relatively limited. This feature is in line with the methodological positioning of simulation research: simulation aims to reveal the structural law of propagation rather than accurately reproduce each path.

The backtracking validation of intervention strategies on real data tests the external validity of intervention effect evaluation. The actual intervention events in the open data were selected, including community health promotion activities, information source credibility labeling and digital skills training programs for the elderly. The simulation effect of these intervention events by the simulation engine was compared with the actual observation effect. The simulation effectiveness verification results are shown in [Figure 5](#).

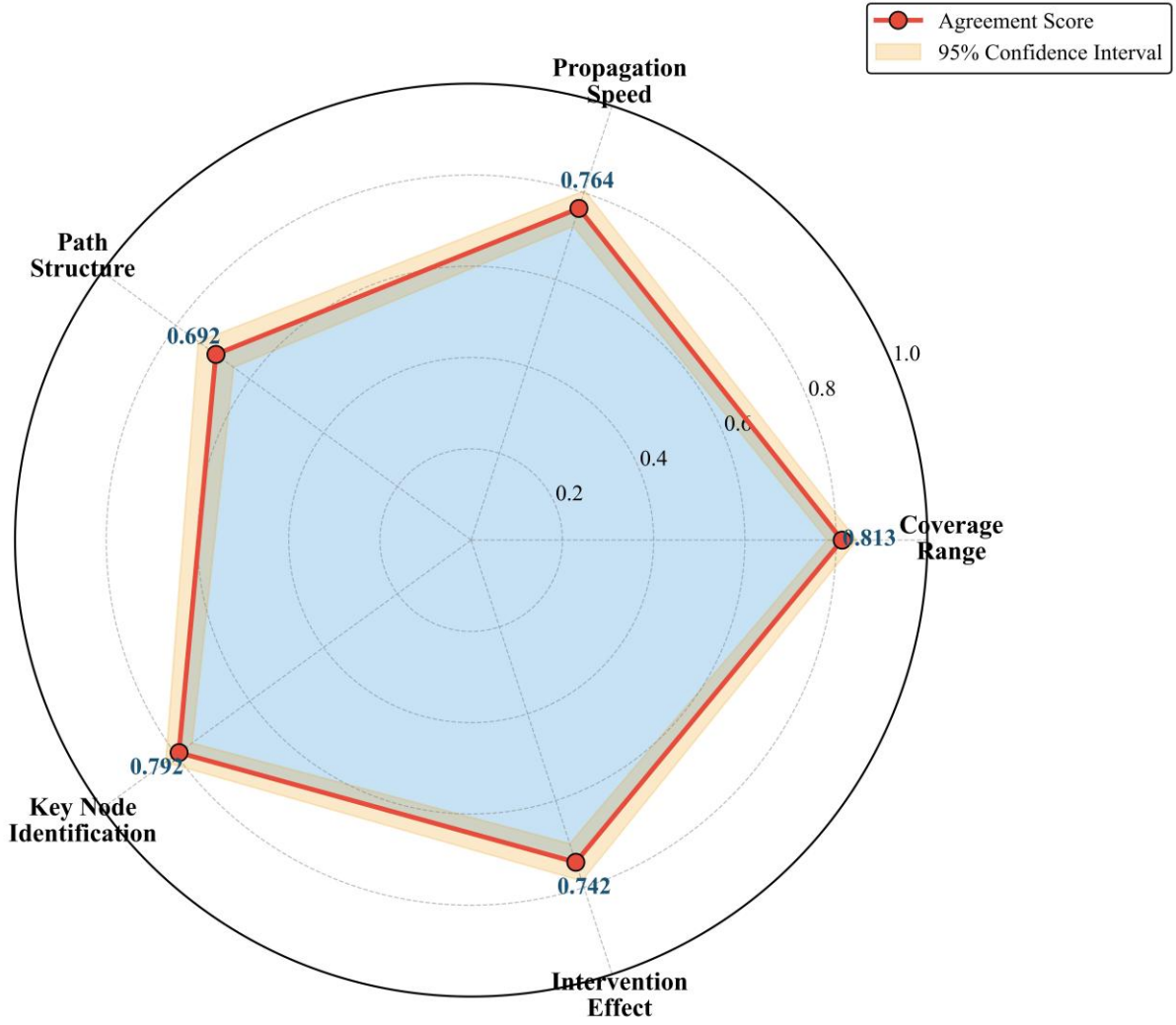


Figure 5. Multi-dimensional agreement between simulation output and real observation

The correlation coefficient between the intervention effect predicted by simulation and the actual observation effect is 0.742, and the average absolute error is 4.3 percentage points. Among the three types of interventions, the simulation prediction marked with information source credibility is the most accurate (error of 3.1 percentage points), and the error of digital skills training program for the elderly is the largest (error of 6.7 percentage points), because the training program involves offline behavior change, while the simulation engine mainly models the online communication process.

7. CONCLUSIONS AND PROSPECTS

7.1 Main Conclusions

Focusing on the core issue of the propagation path of the digital divide in the social network for the elderly, this study constructs a complete research framework from data basis, model architecture to simulation engine, and draws three main conclusions through systematic experimental verification. In terms of the propagation law of the digital divide, the research shows that the diffusion of the digital divide in the social network for the elderly is not a uniform random process but presents significant structural characteristics. The transmission path can be divided into five typical modes: direct

information source, diffusion within the community, cross-community bridge, radiation of medical institutions and multi-source convergence. Among them, the diffusion mode within the community accounts for the highest proportion and the participation rate of elderly users is the strongest, which constitutes the core channel of digital divide communication. Although the proportion of cross-community bridging mode is limited, it contributes nearly 60% of the global state change, indicating that a few bridge nodes have a decisive influence on the transmission pattern. The transmission process showed four stages: slow accumulation, accelerated diffusion, stable transmission and saturation adjustment. There were systematic differences in transmission rate, path length and mode composition in each stage, which provided a clear time window for dynamic intervention timing. Elderly users show stronger path dependence in the interactive relationship layer and information forwarding layer, and the highest age homogeneity in the geographical proximity layer, indicating that their digital divide status is more driven by close social interaction.

In terms of the effectiveness of graph neural network model, the propagation-adoption coupled graph neural network proposed in this study is significantly better than the six representative baseline models in terms of state prediction accuracy and propagation path prediction quality, among which the prediction advantage of literacy gap dimension is the most prominent, and the advantage of long-range path prediction expands with the increase of path length. Ablation experiments confirmed that the path coding module contributed the most to the quality of path prediction, the cross-layer attention mechanism contributed the most to the accuracy of state prediction, the propagation-adoption dual-channel design contributed the most to the evaluation of intervention effect, the causal mask mechanism mainly reduced false positive paths, and the age perception transformation significantly improved the performance of the elderly user subset. The robustness test shows that the model maintains better stability than the baseline under the four types of disturbances: node deletion, edge deletion, feature noise and data missing. Path coding and cross-layer attention provide structural redundancy, and age perception transformation has the ability to suppress the feature noise of elderly users. The generalization ability test shows that the model is robust in the three dimensions of cross-platform migration, cross-regional generalization and time-span generalization. The fine-tuning strategy can effectively restore the cross-platform performance, the geographical proximity layer provides stable structural prior for cross-regional generalization, and the course learning strategy significantly alleviates the performance attenuation of time-series generalization.

In terms of the key findings of intervention strategies, the four types of intervention methods show an obvious trade-off relationship in the transmission suppression effect, intervention cost and benefit balance degree: The key node activation strategy has the strongest effect but the highest cost and the lowest benefit balance degree; the community bridging strategy has a good balance in three-dimensional objectives; the content adaptation strategy has the weakest effect but the lowest cost and the longest duration; and the information injection strategy has the middle index. The Pareto optimal combination strategy obtained by multi-objective optimization is better than the optimal value of single strategy in the three objectives. The timing of intervention is highly sensitive to the effect, structural intervention needs a longer lead time, and informational intervention can respond quickly near the critical point of transmission. The intervention effect is dependent on the network structure, community bridging is more effective in the network with higher modularity, and key node activation is more effective in the network with more uneven degree distribution.

7.2 Summary of innovation points

The primary innovation of this study lies in proposing a set of data-driven heterogeneous social

network construction methods. Different from the existing research, which mostly uses abstract rule network or small-scale questionnaire data, this study takes public social network data as the only data source, establishes a multi-source collection strategy under the guidance of the three principles of public accessibility, elderly user coverage and medical discussion density, and constructs a four-stage cleaning process covering deduplication, noise filtering, privacy desensitization and quality verification. An elderly recognition classifier is constructed based on public portrait information. In terms of node modeling, this study breaks through the traditional assumption of single type node and defines four types of node system including individual users, information sources, medical institutions and community organizations. In the edge modeling, a four-layer edge structure of concern relationship, interaction relationship, information forwarding relationship and geographical proximity relationship is constructed, and an edge weight calculation scheme integrating interaction frequency, emotional polarity and information credibility is designed. The network construction method enables the subsequent propagation path analysis to be carried out on the real topology and avoids the oversimplification of the propagation law by the regular network.

The second innovation of this study lies in the proposed path-aware graph neural network architecture with propagation-adoption coupling. In view of the deficiency of the existing GNN model in confusing the propagation process and adoption process, this study decouples the two into independent information channels and realize re-coupling through coupling gating, so that the model can learn the propagation dynamics and adoption dynamics respectively. Aiming at the problem that the existing GNN model loses the structural information of propagation path, this study introduces the path coding module based on random walk and the multi-path attention aggregation mechanism, so that the model can capture the long-range propagation dependence and identify the key propagation path. Aiming at the problem that it is difficult to distinguish correlation propagation from causal propagation in social networks, this study proposes a causal masking mechanism, which constructs a causal masking matrix based on the three criteria of time priority, intervention consistency and path accessibility, and significantly reduces the proportion of false positive paths. In view of the systematic differences in information processing between elderly users and non-elderly users, this study designs the age perception feature transformation module, so that the model can learn differentiated state evolution patterns for elderly users.

The third innovation of this study lies in the construction of a path-level interpretable digital divided simulation framework. Aiming at the problem that the existing simulation studies mostly stay at the state prediction level and lack the path-level tracking ability, this study designed the bidirectional algorithm of forward propagation path sampling and back propagation path tracing, proposed the critical path contribution evaluation method based on Shapley value, and realized the interpretable induction of path pattern through clustering. Aiming at the problem that the evaluation of intervention strategy mostly relies on a single scenario, this study constructs a counterfactual simulation mechanism to realize the causal evaluation of intervention effect through the comparison of multiple scenarios. In order to ensure the reliability of simulation results, this study designs three kinds of test schemes: step sensitivity analysis, random seed stability test and network scale scalability verification. This simulation framework advances the research of digital divide from descriptive analysis to institutional deduction and strategic optimization and provides an operational decision support tool for social network governance.

7.3 Practical enlightenment

Suggestions for digital divide intervention for social platforms focus on the optimization of

platform ecology. The bridge nodes identified in this study only account for a small part of the total number of nodes but contribute nearly half of the cross-community communication, indicating that the platform should focus on identifying and enabling bridge nodes such as community organizations and comprehensive information sources, and break the information barriers between homogeneous communities of elderly users by improving their information release quality and communication ability. The high path dependence of elderly users in the interactive relationship layer and information forwarding layer revealed by the study indicates that the platform should optimize the information presentation mode of elderly users' high-frequency use function, embed the credibility prompt and semantic simplification mechanism of medical information in the interaction and forwarding link, and lower the threshold of information understanding. The differences of the five types of communication path modes found in this study indicate that the platform should design differentiated information push strategies for different modes, strengthen the credibility certification of information sources for the direct mode of information sources, activate the opinion leaders in the community for the diffusion mode within the community, and establish cross-community recommendation channels for the cross-community bridge mode.

The direction of digital service optimization for the elderly in medical institutions focuses on the transformation of service process for aging. The radiation mode of medical institutions revealed in this study has a short transmission period and a medium participation rate of elderly users, indicating that medical institutions as information sources have high transmission efficiency, and should make full use of their position as information sources to actively release information suitable for the elderly. The literacy gap dimension is found to have the most significant predictive advantage, indicating that the core bottleneck of the digital divide for the elderly lies in information understanding and decision-making ability rather than access conditions. Medical institutions should focus on optimizing the semantic expression and decision-making guidance of medical information and provide step-by-step and multi-form operation guidance. The sensitivity of intervention timing revealed in this study indicates that medical institutions should pay attention to the critical points of communication such as the change of medical policy and seasonal peak of medical treatment and accurately deliver the information of appropriate service for aging before and after the critical points.

The social network governance reference for policy making focuses on the improvement of the institutional environment. The four-stage law of digital divide communication revealed by the research shows that policy intervention should be timely: focusing on digital skills training and infrastructure construction in the slow accumulation period, strengthening information credibility supervision and rumor governance in the accelerated diffusion period, promoting community mutual assistance and intergenerational support in the stable dissemination period, and consolidating achievements and preventing rebound in the saturation adjustment period. The comprehensive advantages of community bridging strategy in effect, cost and balance degree found in this study indicate that policy resources should be inclined to community organizations and offline health service networks, and the information cocoon of elderly users should be broken through online and offline linkage. The dependence of intervention effects on network structure revealed in this study suggests that policy design should be tailored to local conditions, with priority given to bridging strategies for communities with high modularity and key node activation strategies for networks with uneven degree distribution.

7.4 Future research directions

Propagation modeling of multimodal data fusion is an important direction for future research. This study is mainly based on open text data to construct networks and extract features and does not make

full use of multi-modal information such as image, video and voice. The elderly often rely more on intuitive forms such as images and videos in obtaining medical information, and multimodal data can provide richer transmission path clues. Future research can explore the multi-modal graph neural network architecture, unify the coding of text semantics, image content and video information as node features, and model the differentiated diffusion law of different modal information in the transmission process. In addition, cross-platform multi-modal data fusion methods can be explored to integrate multi-source heterogeneous information from different social platforms and build a more comprehensive view of propagation path.

Online learning and real-time simulation system construction is another important direction of future research. The simulation engine in this study is based on offline training model and static network structure, which is difficult to adapt to the real-time evolution of real social networks. Future research can explore the online learning mechanism, so that the model can continuously update parameters from new observed data in the simulation process and adapt to the time-varying characteristics of network structure and propagation mode. It can explore the flow graph neural network architecture, support the dynamic addition of incremental nodes and edges, and realize the dynamic adjustment of real-time propagation path tracking and intervention strategy. It can build a real-time simulation system oriented to practical applications, connect the simulation engine with the data interface of social platform, and provide immediate support for intervention decisions.

The extension of a common framework for cross-domain digital divide communication is a long-term direction for future research. This study focuses on the digital divide communication in the medical treatment scene of the elderly, but the digital divide problem widely exists in education, finance, government affairs and other fields. Future research can explore the extension of this study's propagation-adoption coupled graph neural network architecture and path-level simulation framework to other fields and realize cross-domain knowledge transfer through domain adaptive mechanism. We can explore the construction of domain-independent digital divided communication modeling framework, abstract cross-domain common communication mechanism and intervention strategy and provide theoretical support and methodological tools for wider digital inclusive governance. We can explore the in-depth combination of this research framework and causal inference methods, verify the external validity of simulation conclusions through real intervention experiments, and promote the digital divide research from simulation deduction to empirical verification.

Abbreviations

PAC-GNN, Propagation-Adoption Coupled Graph Neural Network;

GNN, Graph Neural Network;

GRU, Gated Recurrent Unit;

AUC, Area Under the Curve;

BERT, Bidirectional Encoder Representations from Transformers;

GCN, Graph Convolutional Network;

GAT, Graph Attention Network;

GraphSAGE, Graph Sample and Aggregate;

R-GCN, Relational Graph Convolutional Network;
DySAT, Dynamic Self-Attention Network;
EvolveGCN, Evolving Graph Convolutional Network;
NSGA-II, Non-dominated Sorting Genetic Algorithm II;
HDBSCAN, Hierarchical Density-Based Spatial Clustering of Applications with Noise;
AdamW, Adam with Weight Decay;
RMSE, Root Mean Square Error;
F1, F1 Score;
SI, Sensitivity Index;
CV, Coefficient of Variation;
GPU, Graphics Processing Unit;
SNR, Signal-to-Noise Ratio.

Supplementary Material

Not applicable.

Appendix

Not applicable.

Ethics approval and consent to participate.

This study did not involve human participants, animal subjects, or any data requiring ethical approval. Therefore, ethics approval and consent to participate are not applicable.

Acknowledgements

The authors would like to thank the editors of this journal and all the anonymous reviewers who provided valuable comments on this work.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Author contributions

All authors have read and agreed to the published version of the manuscript. The author's contributions are specified as follows: **L.Y.:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

Funding information

This research was supported by the Young Faculty Research Fund of the School of Economics and Management, Lanzhou Institute of Technology (Grant No. KYJJ202402).

Data availability

The data that support the findings of this study are available upon request from the corresponding authors, **L.Y.**

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.

Declaration of AI and AI-assisted Technologies in the Writing Process

During the writing of this article, the author used ChatGPT for spelling and grammar checking. After using this tool, the author reviewed and edited the content as needed and assumes full responsibility for the final published content.

REFERENCES

- [1] Erku, D., Khatri, R., Endalamaw, A., Wolka, E., Nigatu, F., Zewdie, A., & Assefa, Y. (2023). Digital health interventions to improve access to and quality of primary health care services: a scoping review. *International journal of environmental research and public health*, 20(19), 6854. <https://doi.org/10.3390/ijerph20196854>
- [2] Paul, S., Riffat, M., Yasir, A., Mahim, M. N., Sharnali, B. Y., Naheen, I. T., ... & Kulkarni, A. (2021). Industry 4.0 applications for medical/healthcare services. *Journal of Sensor and Actuator Networks*, 10(3), 43. <https://doi.org/10.3390/jsan10030043>
- [3] Hermes, S., Riasanow, T., Clemons, E. K., Böhm, M., & Krcmar, H. (2020). The digital transformation of the healthcare industry: exploring the rise of emerging platform ecosystems and their influence on the role of patients. *Business Research*, 13(3), 1033-1069. <https://doi.org/10.1007/s40685-020-00125-x>

- [4] Li, H., Li, Q., Yang, Z., & Shangguan, X. (2026). Digital technology-enabled public services and rural residents' subjective wellbeing: synergies and pathways in China. *Frontiers in Sustainable Food Systems*, 10, 1718922. <https://doi.org/10.3389/fsufs.2026.1718922>
- [5] Zhao, Y., Zhang, T., Dasgupta, R. K., & Xia, R. (2023). Narrowing the age-based digital divide: Developing digital capability through social activities. *Information Systems Journal*, 33(2), 268-298. <https://doi.org/10.1111/isj.12400>
- [6] Cui, Y., Bao, H., Wen, K., & Wen, H. (2025). Bridging the digital health divide: digital endowment, informal social participation, and health inequality among older adults. *BMC Health Services Research*, 26(1), 99. <https://doi.org/10.1186/s12913-025-13872-6>
- [7] Faye, R., & Ravneberg, B. E. (2024, May). Making vulnerable groups able to connect socially and digitally—Opportunities and pitfalls. In *Frontiers in education* (Vol. 9, p. 1346721). Frontiers Media SA. <https://doi.org/10.3389/educ.2024.1346721>
- [8] Coles-Kemp, L., Robinson, N., & Heath, C. P. (2022). Protecting the vulnerable: Dimensions of assisted digital access. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), 1-26. <https://doi.org/10.1145/3555647>
- [9] Martin, F., Ceviker, E., & Gezer, T. (2026). From digital divide to digital equity: Systematic review of two decades of research on educational digital divide factors, dimensions, and interventions. *Journal of Research on Technology in Education*, 58(2), 396-421. <https://doi.org/10.1080/15391523.2024.2425442>
- [10] Aruleba, K., & Jere, N. (2022). Exploring digital transforming challenges in rural areas of South Africa through a systematic review of empirical studies. *Scientific African*, 16, e01190. <https://doi.org/10.1016/j.sciaf.2022.e01190>
- [11] Zhou, L., Lin, J., Li, Y., & Zhang, Z. (2020). Innovation diffusion of mobile applications in social networks: A multi-agent system. *Sustainability*, 12(7), 2884. <https://doi.org/10.3390/su12072884>
- [12] Zareer, M., & Selmic, R. R. (2025). A survey on opinion dynamics in social media networks: Analysis, simulation, and control. *IEEE Transactions on Computational Social Systems*. <https://doi.org/10.1109/TCSS.2025.3622498>
- [13] Fan, R., Yang, L., Liu, D., & Hu, W. (2026). Value allocation mechanism for multi-agent data sharing in digital innovation network: Agent-based modeling. *Expert Systems with Applications*, 299, 130124. <https://doi.org/10.1016/j.eswa.2025.130124>
- [14] Tai, Y., He, H., Zhang, W., Yang, H., Wu, X., & Wang, Y. (2023). Predicting information diffusion using the inter-and intra-path of influence transitivity. *Information Sciences*, 651, 119705. <https://doi.org/10.1016/j.ins.2023.119705>
- [15] Zheng, Y., Yi, L., & Wei, Z. (2025). A survey of dynamic graph neural networks. *Frontiers of Computer Science*, 19(6), 196323. <https://doi.org/10.1007/s11704-024-3853-2>
- [16] Feng, Z., Wang, R., Wang, T., Song, M., Wu, S., & He, S. (2025). A comprehensive survey of dynamic graph neural networks: Models, frameworks, benchmarks, experiments and challenges. *IEEE Transactions on Knowledge and Data Engineering*. <https://doi.org/10.1109/TKDE.2025.3621291>

- [17] Gondal, N. (2023). Diffusion of innovations through social networks: Determinants and implications. *Sociology Compass*, 17(5), e13084. <https://doi.org/10.1111/soc4.13084>
- [18] Xiang, S., Ling, H., & Wu, M. (2026). Cross-Modal Alignment and Rectified Flow-Based Latent Representation Synthesis for Enhanced Speech-Driven Alzheimer's Disease Detection. *Bioengineering*, 13(3), 370. <https://doi.org/10.3390/bioengineering13030370>
- [19] Zhuo, S., Fang, J., Lin, H., Li, N., Zhou, Y., Zhang, S., ... & Huang, S. (2026). EdgeGFL: rethinking edge information in graph feature preference learning. *International Journal of Machine Learning and Cybernetics*, 17(5), 240. <https://doi.org/10.1007/s13042-026-03060-1>
- [20] You, X., Zhang, M., Ma, Y., Tan, J., & Liu, Z. (2023). Impact of higher-order interactions and individual emotional heterogeneity on information-disease coupled dynamics in multiplex networks. *Chaos, Solitons & Fractals*, 177, 114186. <https://doi.org/10.1016/j.chaos.2023.114186>
- [21] Gong, Z., Shao, J., Rahman, N., Su, L. Y. F., & Wang, Y. C. (2026). Modality matters: comparing the persuasiveness of text-and voice-based conversational agents and non-interactive messages in health communication. *Internet Research*, 1-25. <https://doi.org/10.1108/INTR-07-2025-1004>
- [22] Cui, S., & Zhu, X. (2024). The information propagation mechanism of individual heterogeneous adoption behavior under the heterogeneous network. *Frontiers in Physics*, 12, 1404464. <https://doi.org/10.3389/fphy.2024.1404464>
- [23] Yang, Y. (2025, July). Research on Evaluation Model of Urban-rural Integration Development Based on Deep Learning. In *Proceedings of the 2025 3rd International Conference on Educational Knowledge and Informatization* (pp. 406-410). <https://doi.org/10.1145/3765325.3765393>
- [24] Ao, X., Gong, Y., & He, A. (2025). A review of time series prediction models based on deep learning. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2025.3602791>
- [25] Yu, Y., & Huo, L. A. (2025). Effects of official information diffusion and rumor-related behavior adoption on epidemic transmission in multiplex networks. *Information Sciences*, 689, 121414. <https://doi.org/10.1016/j.ins.2024.121414>
- [26] Liu, Y., Zhang, P., Shi, L., & Gong, J. (2023). A survey of information dissemination model, datasets, and insight. *Mathematics*, 11(17), 3707. <https://doi.org/10.3390/math11173707>
- [27] Theodorakopoulos, L., Theodoropoulou, A., & Klavdianos, C. (2025). Interactive viral marketing through big data analytics, influencer networks, AI integration, and ethical dimensions. *Journal of Theoretical and Applied Electronic Commerce Research*, 20(2), 115. <https://doi.org/10.3390/jtaer20020115>
- [28] Yue, Z., Witzig, C. R., Jorde, D., & Jacobsen, H. A. (2020, November). Bert4nilm: A bidirectional transformer model for non-intrusive load monitoring. In *Proceedings of the 5th international workshop on non-intrusive load monitoring* (pp. 89-93). <https://doi.org/10.1145/3427771.3429390>
- [29] Liao, T., Ta, X., Xu, Y., Han, L., Sun, L., & Lv, W. (2025). SimPRL: a simple contrastive learning for path representation learning by joint GPS trajectories and road paths. *IEEE Transactions on Intelligent Transportation Systems*, 27(1), 400-413. <https://doi.org/10.1109/TITS.2025.3629800>
- [30] Yang, C., Xiao, Y., Zhang, Y., Sun, Y., & Han, J. (2020). Heterogeneous network representation learning: A unified framework with survey and benchmark. *IEEE Transactions on Knowledge and*

Data Engineering, 34(10), 4854-4873. <https://doi.org/10.1109/TKDE.2020.3045924>

APA Citation

Yang, L. (2026). Simulation Study on Digital Divide Propagation Path of Elderly Medical Treatment Based on Graph Neural Network and Public Social Network Data. *Journal of Discovery Core*, 1(2), 38-79. <https://doi.org/10.67541/jdc2609>