Knowledge-Based Systems 301 (2024) 112246

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Unifying multimodal interactions for rumor diffusion prediction with global hypergraph modeling

Qi Zhang [a,1], Yuan Li [a,1], Jialing Zou [a], Jianming Zhu [b], Dingning Liu [c], Jianbin Jiao [b,*]

[a] School of Electronics, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing, 100049, China
[b] School of Emergency Management Science and Engineering, University of Chinese Academy of Sciences, Beijing, 100049, China
[c] School of Software, Dalian University of Technology, Dalian, 116622, China

ARTICLE INFO

Keywords:
Social network services
Rumor diffusion prediction
Deep learning
Hypergraph neural network
Multimodal interactions

ABSTRACT

A central issue in rumor surveillance and management is decoding the complex dynamics of rumor propagation, with an emphasis on predicting diffusion cascades. Recent studies focus on node embedding or the sequence of dissemination in a rumor cascade based on user social interaction, while neglecting interactions between users and rumors, as well as between various rumors. Online rumor diffusion is a complex system that encompasses two fundamental components: rumors and users. A comprehensive understanding of this system's dynamics requires a global perspective. Consequently, it is necessary to develop models that capture the inherent multimodal interactions in rumor propagation. To tackle this challenge, we propose an HG2RLink framework of rumor diffusion prediction, which unifies multimodal interactions by global encoding. Specifically, our methodology begins with the establishment of a hypergraph structure. We then refine these interactions by leveraging a hypergraph neural network that aggregates users' preferences for rumors. Two high-order graphs are generated to capture latent spatial interactions of rumors. By employing a learning approach for multimodal interactions, each rumor diffusion sequence is modeled with a long-range vision field of users in a deep neural network. Moreover, to validate the effectiveness of our method, we introduce a new dataset available for further exploration of new methods. Finally, experimental results show that HG2RLink outperforms other methods with improvements ranging from 0.8% to 7.6% for the Hits@k metric and from 0.4% to 2.6% for the MAP@k metric across four diverse datasets.

1. Introduction

Rumors often refer to information whose truth and source are unreliable [1]. With the deep integration of social network services(SNS) into people's lives, massive rumors are spreading in the SNS. Some spread deeper, faster, and more permeable than real news [2]. Given the constraints of personal expertise and knowledge, it is difficult to find the veracity of information from the plethora of online content for those not privy to the truth. In recent years, the frequent occurrences of online rumors, often cause trust damage [3], public panic, social chaos, and economic volatility, and even exacerbate the impact of disasters. Consequently, the timely and precise forecasting of rumor cascades holds profound importance for emergency management authorities. Such predictions are instrumental in enhancing their comprehension of the rumor diffusion process, enabling the implementation of strategic interventions and guidance, thereby safeguarding and sustaining societal stability.

Modeling the rumor cascade usually involves the critical step of link prediction. Depending on the formulation of the problem in different applications, link prediction involves identifying missing links or forecasting potential future links [4]. Previous works on identifying missing links have predominantly relied on heuristic approaches, where the likelihood of link existence is determined by calculating similarity scores between nodes. Lin et al. [5] defined similarity based on the basic attributes of nodes, but these attributes are often concealed. Some research used the similarity [6–9] of network structure to measure the links among users. These methods excessively rely on the topological structure of the relationship network, thus they cannot predict well for datasets with missing relationships or incomplete relationship data. In addition to identifying missing links, it is also crucial to forecast the potential link for future propagation. The work in this field can be mainly divided into two methods: feature engineering-based and representation learning-based. Several studies [10–12] have extracted

* Corresponding author.
E-mail addresses: zhangqi203@mails.ucas.ac.cn (Q. Zhang), liyuan23@ucas.ac.cn (Y. Li), zoujialing@rahvolt.com (J. Zou), jmzhu@ucas.ac.cn (J. Zhu), dutldn@dlut.edu.cn (D. Liu), jiaojb@ucas.ac.cn (J. Jiao).
[1] The first two authors contributed equally.

https://doi.org/10.1016/j.knosys.2024.112246
Received 2 November 2023; Received in revised form 19 June 2024; Accepted 10 July 2024
Available online 15 July 2024
0950-7051/

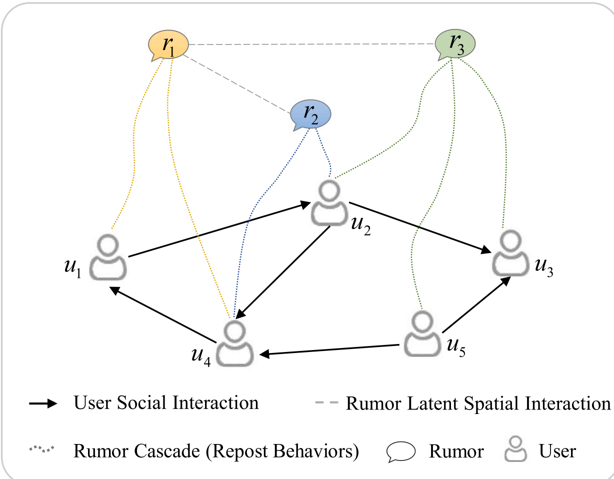

Fig. 1. A simple example of the multimodal interactions in online rumor diffusion.

cascade features from various perspectives to describe the information diffusion process. However, the cost of manual feature extraction is expensive. More importantly, these features designed in advance are not scalable, which will lead to poor model generalizability. In recent years, some techniques based on representation learning [13–15] are used for potential link prediction to obtain node embeddings for alleviating the burden of manual feature design. With the advancement of deep neural networks (DNN), several studies [16–20] commonly explored deep neural networks to study the spread of rumors cascade to predict links. Although the existing architectures of deep neural network models could well merge multiple representations of users in the network graphs, most of them only focused on users' social interaction. This limited scope overlooks the broader, multimodal interactions that occur across the network.

However, the interaction structure could be much more complicated in real-world scenarios. Numerous online rumors are spreading in the SNS simultaneously. We noticed that there are multimodal interactions between rumors and users throughout the rumor dissemination process, including various types of many-to-many relationships. For instance, a single rumor cascade can involve a multitude of users, and an individual user may contribute to the spread of different rumors. In addition, the impact of the interaction between rumors and users is of utmost importance. Given that users exhibit diverse preferences for rumors, the attractiveness of rumors to individuals naturally fluctuates [21]. When users encounter rumors that align with their interests, potential connections may be generated between individuals and rumors. Additionally, there exists a degree of similarity among rumors, suggesting that different rumors can engender potential spatial interactions. Previous studies focused primarily on the social relationships of users, neglecting the direct interactions between rumors and users, as well as the subtle interactions between rumors themselves. In this case, the traditional graph structure has trouble reflecting many relationships among users bringing about rumor diffusion. Consequently, it is crucial to adopt a multimodal modeling approach from a global perspective to effectively capture and analyze the complex dynamics of rumor diffusion. This comprehensive method will enable a more accurate understanding of how rumors spread and influence user behavior. This is also the primary motivation of our research.

Based on the above motivation, our research is driven by three key objectives to enhance the early detection and monitoring of rumor diffusion: (1) Our foremost goal is to develop a predictive framework that is not only more accurate and dependable but also robust enough to offer substantial support for the management and regulation of online rumors. (2) We are committed to enhancing the scalability and generalization of our framework by exploring and using deep learning techniques without relying on manual feature extraction. (3) Most importantly, we aim to demonstrate that the integration of comprehensive global information can significantly enhance the prediction of rumor diffusion by modeling the multimodal interactions between users and rumors.

Therefore, we introduce a hypergraph structure to construct multimodal interactions in online rumor diffusion. These complex interactions, as shown in Fig. 1, include user social interactions, rumor latent spatial interactions, and rumor cascades. User social interaction represents the follower relationship between users. Rumor latent interaction represents a similar relationship between rumors. Rumor cascades represent the users' forwarding behaviors, which is essentially the interaction between the user and the rumor. Based on the hypergraph, we propose an HG2RLink (hypergraph to rumor link) framework to address the challenge of online rumor link prediction. This innovative framework not only integrates user preferences into a hypergraph neural network (HGNN) [22] to learn user interactions but also captures low-dimensional embedding of rumor interactions by constructing two different higher-order graphs of rumors based on the similarity of users and structures of rumor cascades. To capture the comprehensive information of interactions, we merge the embeddings of users and rumors with global attention based on the hypergraph. Then we encode the diffusion sequence of each rumor cascade and develop a DNN with multi-head attention to solve the rumor link prediction. To validate the proposed framework, four real public datasets were utilized. In conclusion, the main contributions of this study are as follows:

- **Global coding of interactions.** We propose an HG2R-Link framework aggregating users' preferences for rumor diffusion prediction which unifies multimodal interactions by global encoding.
- **A new dataset.** We build a de-identified and available dataset that fills a gap in public datasets of social relationship networks for nearly a decade.
- **Better performance.** The results of the experiment demonstrate that HG2RLink has higher generalization in addition to being more efficient.

2. Related work

Our study in this paper deals with the link prediction of rumor cascades. In this section, we will review two aspects: missing link prediction and potential link prediction.

2.1. Missing link prediction

The prediction of missing links aims to estimate the likelihood of the existence of a link between two nodes based on observed links [23]. Liben-Nowell and Kleinberg [24] developed a link prediction method that relied on node proximity, which has garnered significant attention from researchers. Subsequent studies have attempted to estimate similarity between users by considering user attributes [25–27] and behaviors [28–30]. For instance, Jiang et al. [25] utilized a topic domain dictionary to construct a weakly supervised matrix for social text, and defined similarity through comparison with a reference matrix, and Leung et al. [28] utilize big data technology to analyze massive historical behavioral data and identify the most similar nodes. However, these methods often require dealing with large amounts of textual information, which may not always be available. As a result, current research has primarily focused on analyzing network structure [31–34] to address this limitation. De et al. [33] proposed a stacked two-level learning framework that integrates local link, and community-level link density. Coskun et al. [32] introduced a novel approach that leverages the global network topology structure to enhance prediction accuracy. Butun et al. [31] addressed the significance of link direction in directed networks and proposed a direction-pattern-based approach. Singh et al. [9] employed a community detection algorithm [35] in a multiple network based on diverse relationships to ascertain the relevance of nodes, thereby computing the likelihood scores of non-existing. Recently, the work on missing link prediction extended to knowledge graph domains, where strategies [36,37] primarily involve learning low-dimensional embeddings. Additionally, some scholars [38] posit that many facts in knowledge graphs may change over time and have provided detailed summaries.

Those studies primarily focus on completing relationships within graphs. However, they lack modeling for the diffusion graphs of dynamically changing information cascades, making them unsuitable for predicting rumor propagation.

2.2. Potential link prediction

The prediction of potential links aims to forecast the cascade propagation process by offering insights into the early stages of cascades [23]. Cheng et al. [11]. investigated the patterns of cascade propagation from the perspectives of temporal and structural characteristics, demonstrating their efficacy in understanding information diffusion. Weng et al. [10] defined three sets of features, namely the influence of early adopters, community concentration, and time series, and provided evidence supporting the effectiveness of community-based structural features. However, these methods require laborious manual feature engineering, which is both time-consuming and prone to errors. Recently, several research started to utilize deep learning to build up link prediction frameworks from end to end without underlying explicit diffusion models. An RNN-based model called DeepCas [39] could automatically learn the representation of cascade graphs. DeepCon+Str [40] was a recently proposed semi-supervised model for building proximity-based cascade graphs' content and structure in order to understand how they are represented. TopoLSTM [41] produced a topology-aware embedding for each node in the diffusion prediction by including topological features in the conventional LSTM model. FOREST [42] was a reinforcement learning framework that enabled a microscopic cascade model for macroscopic diffusion prediction. CasCN [43] used a self-excitation mechanism and temporal decay mechanism to extract the topological structure of the diffusion processes. To model the preferences of users at different periods, DyHGCN [19] embedded temporal information into the heterogeneous graph and combined the structural feature of social networks and information dissemination graphs. NDM [44] made relaxed assumptions and applied the convolutional network and attention technique to the cascade prediction. HyperINF [18] used the HGNN to tackle diffusion prediction with the dynamics of user interest. MS-HGAT [45] designed modules for self-attention and memory-enhanced look-up to emphasize the interactions within the cascade and improve prediction accuracy. DisenIDP [46] captures user potential intentions from different angles and uses an attention-based encoder to extract long-term and short-term cascade influence, thus better meeting real scenarios. MIDPMS [47] models the diffusion process as a substitution system, exploring the competitive and cooperative relationships between information, the attractiveness of information to users, and the potential impact of content expectations on further diffusion.

In general, most of the current works on rumor diffusion with deep learning methods focused on learning how to depict social relation graphs and dynamic diffusion graphs, ignoring the role of rumors and the global relationship between rumors and users. Therefore, we are inspired to explore the multimodal and higher-order relationships among rumors and users.

3. Problem definition

For online rumor diffusions, both users and rumors are two essential elements that should coexist in the propagation system. It is their combined effect that drives the rapid spread of rumors. Therefore, let

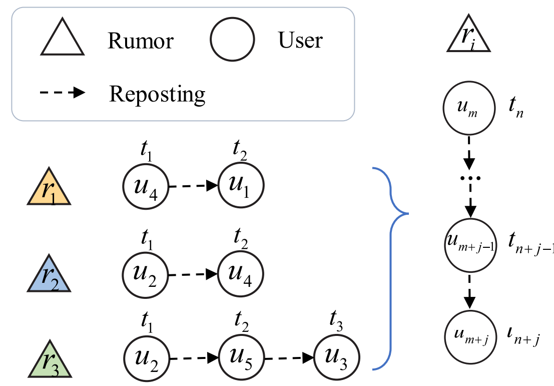

Fig. 2. Rumor cascades and link prediction task.

$H = (V, X, W)$ denote the hypergraph composed of rumors and users, where V represents the set of users and X is a set of hyperedges, $e_i \in X$ is the set of users involved in different rumors r_i. In practice, each rumor r_i corresponds to a hyperedge e_i. W denotes the weights (for a detailed construction and explanation, see Section 4.1). For user forwarding behavior, the set of rumor diffusion sequences $L = \{(r, u, t)|u \in V, r \in R\}$ is recorded, where *(r, u, t)* means that user u published or reposted rumors *r* at time *t*. Here, use Fig. 1 as a simple example of rumor diffusion. It can be seen that user u_4 and u_1 publish or repost rumor r_1, user u_2 and u_4 take action with rumor r_2, and user u_2, u_3, u_5 participate in rumor r_3. As a result, the spread of each rumor has formed a rumor cascade between several users. It can be denoted as a sub-sequence $l_r = \{(u_m, t_n), \dots, (u_{m+j-1}, t_{n+j-1}), (u_{m+j}, t_{n+j})|r \in R\}$, as shown in Fig. 2.

Thus, the issue we want to resolve is to predict the following user u_{m+j} at the time t_{n+j} in the rumor diffusion sequence l_r. Therefore, the rumor link prediction problem can be described as $u_{m+j} = argmax_{u \in V, r \in R} P(u \mid l_r, H, L)$.

4. Method

In this section, we will introduce the HG2Rlink, and its overall framework is shown in Fig. 3. Firstly, we differentiate our approach from previous studies, which focused primarily on social interaction between users, by recognizing the propagation of rumors as a system. Within this holistic system, there exist intricate interactions between users and rumors. Therefore, we addressed this by adopting hypergraph-based modeling to effectively capture these multimodal relationships within the system. Especially, we build a hypergraph of rumor interaction to reflect the higher-order relationships between rumors and users to understand the intricate interactions in the diffusion of rumors online. To reduce the dependence on empirical data, we use users' individual preferences for rumors and social relation graphs and the hypergraph as inputs of the model. Then, in this framework, there are three modules: dual-channel representation learning, global interactions encoding, and rumor diffusion prediction. Specifically, the dual channel representation learning module includes user interaction representation learning and rumor interaction representation learning. On the user side, an HGNN layer is constructed to capture the hidden features of users under three inputs. On the rumor side, the rumors are represented in vector space with few dimensions by the graph embedding layer for generated rumor high-order graphs. In the global interaction encoding module, we use the multi-head graph attention network (GAT) [48] to integrate the multimodal interactions between users and rumors through a global perspective, exploring the dependent relations among the users and the rumors. By combining hypergraph and GAT technology, we can integrate multimodal interaction and encode high-order correlations, thus better extracting representations of

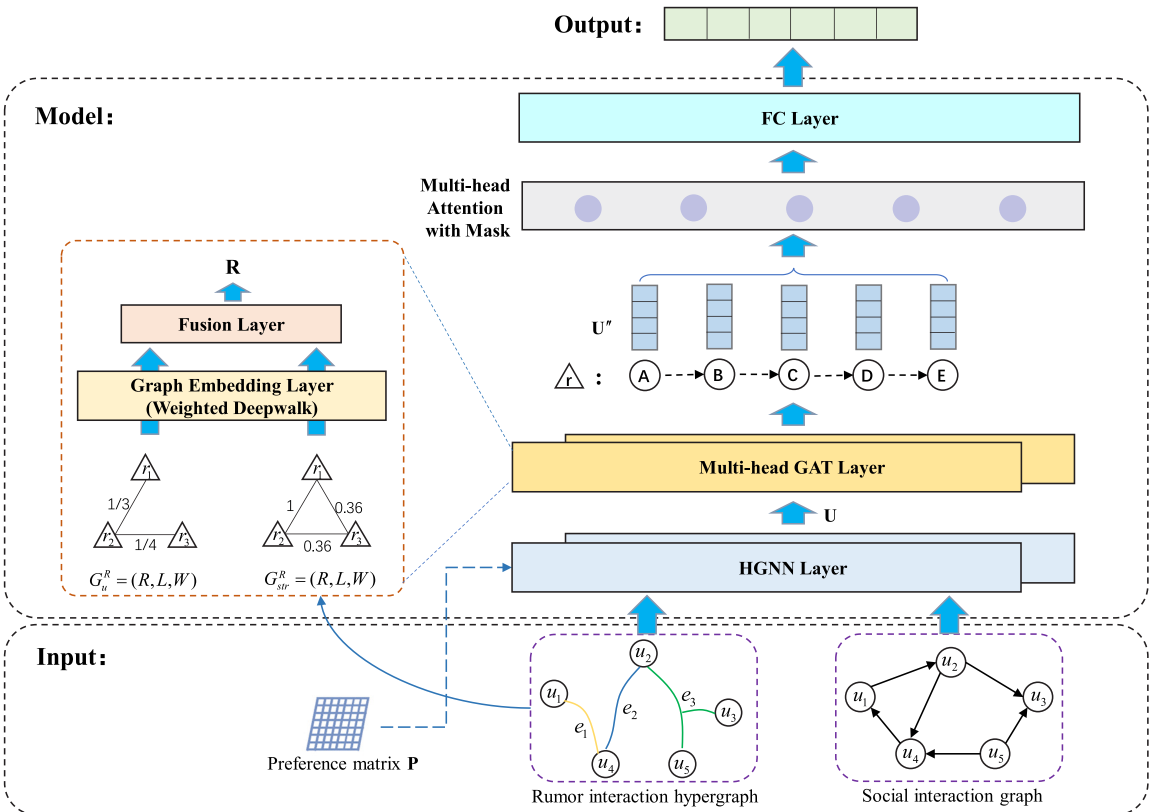

Fig. 3. The overall framework of the proposed HG2RLink. **U** and **U**″ represent user embeddings through different layers, and **R** represents rumors' embeddings. Triangles represent rumors and the circled letters represent users. The black dashed arrow indicates the order of rumor propagation.

the rumor propagation system. Finally, each rumor diffusion sequence is represented in the link prediction module. Then a classification function is learned using a multi-head attention DNN to forecast the likelihood of user participation in rumor spreading.

4.1. Modeling the rumor interaction hypergraph

Since the traditional graph structure has trouble reflecting the multimodal relationships of data, hypergraphs are introduced here. A generalized graph structure consisting of a collection of nodes and hyperedges is named a hypergraph. A hyperedge in a hypergraph can link any number of nodes, unlike an edge in a conventional graph, which only connects two nodes. Formally, a hypergraph can be represented as $H = (V, X, W)$, where V is a set of elements representing nodes, X is a set of hyperedges, $e_i \in X$ is a nonempty subset of V, and W represents a positive weight assigned to each hyperedge e_i. Obviously, hypergraphs have more advantages in modeling multimodal interactions. In our research, to retain the high-order information as much as possible, the hypergraph is used to model the complex interactions in the spread of online rumors. An online rumor is treated as a hyperedge. Each hyperedge contains several nodes representing all users who spread the rumor. Thus, the connection of rumor interaction between nodes and hyperedges is denoted by $H \in R^{|V|\times|X|}$. It should be noted that this work does not distinguish the importance and influence of different rumors, so the weight of each hyperedge is set as 1. Here, take the data in Fig. 1 as an instance to describe the procedure of constructing a rumor interaction hypergraph $H = (V, X)$. As shown in Fig. 4, we first find the user set participated in each rumor from Fig. 1, as shown in Fig. 4(a). Then use the three rumors as a set of hyperedges $X = \{e_1, e_2, e_3\} = \{(u_1, u_4), (u_2, u_4), (u_2, u_3, u_5)\}$, and take all users participated in these rumors as the node set $V = \{u_1, u_2, u_3, u_4, u_5\}$. In Fig. 4(b), curves with different colors represent the three hyperedges in the rumor interaction hypergraph.

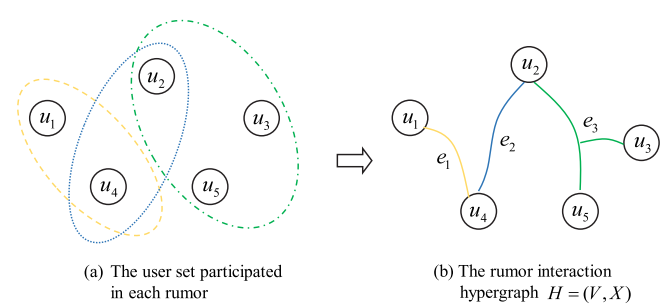

Fig. 4. Rumor interaction hypergraph construction.

4.2. Dual-channel representation learning

Based on the interactive hypergraph of rumors, we perform local interaction encoding and representation learning with the two modules of user interactions and rumors interactions.

4.2.1. Users' interactions representation learning

User behavior plays a crucial role in the spread of online rumors, and user behavior is often influenced by users themselves, their social relationships, and the rumors. According to the constructed rumor interaction hypergraph, we discover that the implicit information in the hypergraph could enhance the understanding of user behaviors. There fore, based on the rumor interaction hypergraph $H = (V, X)$ and the user social interaction graph $G = (V, E)$, we learn the embedding of nodes in these graphs through graph representation learning. For being different from the traditional graph with pairwise connections, hypergraph is unsuitable for general learning graph representations. It needs to encode the high-order relations. Inspired by Feng et al. [22], we attempt an HGNN module to update the hidden feature of the user in hypergraph with the users' embedding learned from the social relation graphs. This layer can aggregate the features of related hyperedges in the rumor interaction hypergraph to get new representing vectors of the users. In the meantime, it is crucial to simulate the aforementioned user preferences. Therefore, we extract the user preferences of different rumors to integrate into the HGNN (as shown in the input of Fig. 3).

First, we use the Deepwalk algorithm [49] to get the embeddings of nodes in graph $G = (V, E)$, denoted as u. Then, from hypergraph $H = (V, X)$, we extract the incidence matrix $\mathbf{H} \in R^{|V| \times |X|}$, where rows represent different nodes, columns represent different hyperedges, and the values are represented by Eq. (1).

$$h(v, e) = \begin{cases} 1, if & v \in e \\ 0, if & v \notin e \end{cases} \tag{1}$$

For each vertex $v \in V$, its degree is defined as $D(v) = \sum_{e \in X} h(v, e)$. For each hyperedge $e \in X$, its degree is defined as $D(e) = \sum_{v \in V} h(v, e)$. The number of nodes that make up each hyperedge determines its degree. $\mathbf{D}_v$ and $\mathbf{D}_e$ represent the diagonal matrices of edge degree and vertex degree, respectively. Meanwhile, users' rumor preference matrix $\mathbf{P}$ is constructed by extracting the frequency of users' participation from the rumor diffusion sequences L. To consider how user preferences affect rumor diffusion, we construct a weighted incidence matrix $\mathbf{H}'$ by adding the preference matrix $\mathbf{P}$ to the incidence matrix $\mathbf{H}$ of the hypergraph $H = (V, X)$. Finally, according to the general HGNN model, we build a hyperedge convolutional layer $f(U, W, \Theta)$ in the following formulation.

$$\mathbf{U}^{(l+1)} = \sigma(\mathbf{D}_v^{-\frac{1}{2}} \mathbf{H}' W D_v^{-\frac{1}{2}} \mathbf{H}'^T \mathbf{D}_v^{-\frac{1}{2}} \mathbf{U}^{(l)} \Theta^l) \tag{2}$$

In the above hyperedge convolutional layer, $\mathbf{U}$ represents the node feature, $\mathbf{U}^{(l+1)}$ is the hypergraph's signal at the $l+1$ layer, $\mathbf{U}^{(0)}$ represents the initial node feature $\mathbf{u}$, $\mathbf{W}$ represents the weight matrix of the hyperedge. In our work, the weights of hyperedges are all set to 1. σ is a nonlinear activation function. Θ is a learnable parameter matrix. With this HGNN model, based on the spectral convolution on the hypergraph, the user's representation $\mathbf{U}$ is well-learned.

4.2.2. Rumors' interactions representation learning

Rumor cascade can directly reflect the dynamics of rumor diffusion. Thus, we generate a diffusion cascade graph for each rumor based on the rumor interaction hypergraph. Given that user social relationships play an important role in rumor propagation, we also introduce social relationships from the rumor cascade graph. Through the analysis of the rumor cascade graph, we find that the users in cascades and the structure of cascade are two critical features of the rumors. In order to obtain a low-dimensional representation of the rumor, we have defined a new rumor high-order graph and constructed two forms of the high-order graph based on the node similarity and structure similarity of the rumor cascade graph.

Definition 1 (*Rumor High-Order Graph*).. Given a rumor interaction hypergraph $H = (V, X)$ and a social interaction graph $G = (V, E)$, extract a rumor cascade set $C^R (C_i^R \in C^R)$. We define the rumor high-order graph as $G^R = (R, L, W)$. It encodes the proximity between rumor cascades. R is the set of corresponding rumors in the C^R, L is the set of links extracted similarity between rumor cascades, and W is the weight determined by similarity metrics.

(1) User similarity in rumor cascade graph

The user's forwarding behavior can trigger a rumor cascade. It is evident that the participation of different users can lead to dynamic changes in the cascade, and the size of the cascade largely depends on the influence of the users. Therefore, the similarity of users in the rumor cascade is considered first to measure the implicit relationship between different rumors. We use the Jaccard coefficient [24] to measure the similarity of users between rumor cascades, as shown in Eq. (3). $Sim_u(c_1^R, c_2^R)$ is a ratio of the number of users who jointly take part in two rumor cascades to the total number of users participating in two rumor cascades.

$$Sim_u(c_1^R, c_2^R) = \frac{|u_{c1} \cap u_{c2}|}{|u_{c1} \cup u_{c2}|} \tag{3}$$

where U_{c1} and U_{c2} represent the participating users in two cascades c_1^R and c_2^R, respectively. If the similarity of users is more than 0, there is an edge between the corresponding rumor in $G_u^R = (R, L, W)$, and the weight of this link is the value of the user similarity. Fig. 5(a) shows this method with a simple example.

(2) Structural similarity of rumor cascade graph

Rumor cascade graphs among users can be abstracted into a variety of topologies, including a tree, star, and other diagrams. Different topologies could cause different cascade trends. Therefore, we construct the rumor high-order graph $G_{str}^R = (R, L, W)$ based on structural similarity between rumor cascades to automatically learn the characteristics of rumor diffusion's structure.

Similar to the method of degree distribution of nodes in [40], we apply Equation (4) to determine the structural similarity of the two rumor cascades.

$$Sim_{str}(c_1^R, c_2^R) = e^{-Dist(v_1, v_2)} \tag{4}$$

where v_1 and v_2 are the root users of the rumor cascades c_1^R and c_2^R. $Dist(v_1, v_2)$ is calculated by Eq. (5), and represents a sum of the distances between users in a degree-order at all potential k distances from root users.

$$Dist(v_1, v_2) = \sum_{k=0:K} l(s(D_k(v_1)), s(D_k(v_2))) \tag{5}$$

here, $D_k(v)$ is the set of users with the distance (hop count) $k (k > 0)$ to user v in the c_i^R. Then $D_0(v)$ represents user v itself, and $D_1(v)$ contains all the neighbors of v. $s(D_k(v))$ denotes the ordered sequence of degrees of $D_k(v)$, $D_k(v) \subset V$. Noting that the K should equal to the smaller of two distances k_1 and k_2 in Eq. (5), and k_i is the maximum distance from all users of c_i^R to the root user v_i. $l(s_1, s_2)$ measures the distance between sequences s_1 and s_2. Since s_1 and s_2 may be of different lengths, the Euclidean distance could not effectively calculate the distance between these two sequences. To solve this problem, we adopt the dynamic time warping (DTW) algorithm [50], which is used to measure the similarity of two-time series of different sizes. Therefore, after calculating the value of $l(\cdot)$ by DTW algorithm, the structural similarity $Sim_{str}(c_1^R, c_2^R)$ between cascades can be obtained. Fig. 5(b) illustrates this construction process. Specifically, we get three cascades c_1^R, c_2^R and c_3^R from Fig. 1. u_4, u_2 and u_2 are the root nodes of three cascades respectively. For c_1^R, $D_0(u_2) = (u_2)$, $D_1(u_2) = (u_4)$ and $s(D_0(u_4)) = (1)$, $s(D_1(u_4)) = (0)$. For c_2^R, $D_0(u_2) = (u_2)$, $D_1(u_2) = (u_4)$ and $s(D_0(u_2)) = (1)$, $s(D_1(u_2)) = (0)$. For c_3^R, $D_0(u_2) = (u_2)$, $D_1(u_2) = (u_5)$, $D_2(u_2) = (u_3)$ and $s(D_0(u_2)) = (1)$, $s(D_1(u_2)) = (1)$, $s(D_2(u_2)) = (0)$. Then by DTW algorithm, we get the value $Dist(\cdot)$. Finally, Eq. (4) is used to get the results. It should be pointed out that in order to ensure that $G_{str}^R = (R, L, W)$ is not a fully connected graph, a link exists in $G_{str}^R = (R, L, W)$, only if the similarity of the two cascades is not less than the average value of all structural similarities in our experiment.

Through the acquisition of the above two rumor high-order graphs, we also use the Deepwalk algorithm to mine information on graph structure. But unlike the user social relationship graph in the user channel, the rumor high-order graphs are weighted graphs. Thus, we design a weighted Deepwalk algorithm to break through the limitation of the original Deepwalk algorithm suited for unweighted graphs. We design a transition probability matrix Φ constructed with weights in $G^R = (R, L, W)$ as prior knowledge for random walks. Suppose in $G^R = (R, L, W)$ the rumor node r_0 has n neighbors, and w_{0i} is the weight of each adjacent edge. Then, each value in Φ is calculated by Eq. (6).

$$p_{0i} = \frac{w_{0i}}{\sum_{i=1}^{n} w_{0i}} \tag{6}$$

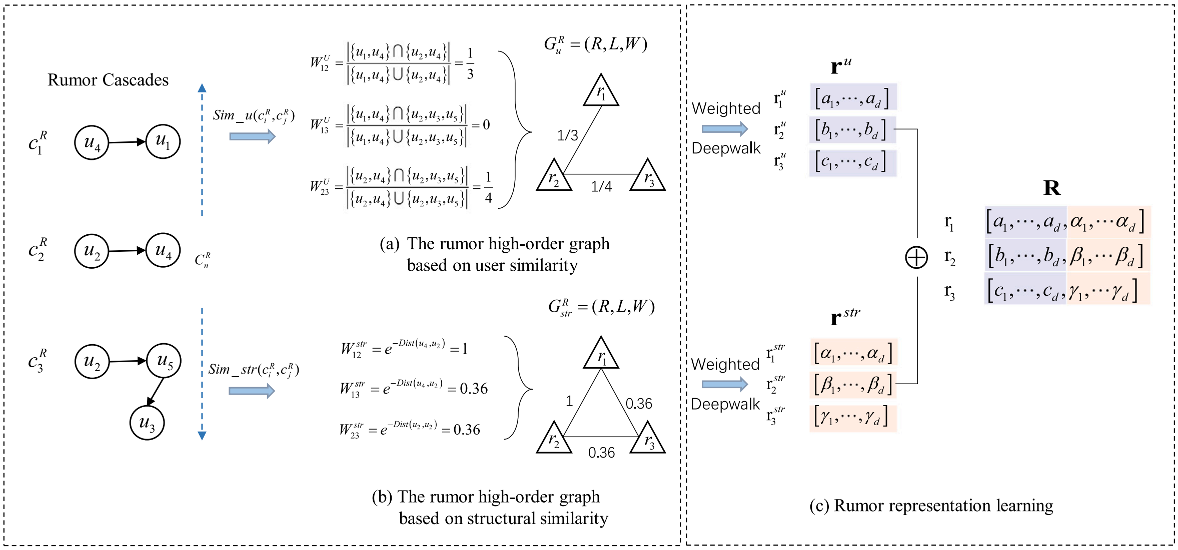

Fig. 5. The generation of high-order rumor graphs and rumor representation learning.

In the process of sampling the node sequence using random walk, neighbor nodes are selected with the transition probability of the seed node, and the process is repeated before the preset length of the sequence is reached. After sampling, the sequence is input into the word2vec [51] model for training. After learning the representations of two rumor high-order graphs respectively, the low-dimensional embeddings $\mathbf{r}^u$, $\mathbf{r}^{str}$ of the rumors are obtained. Both embeddings are important, thus they are composed into the ultimate representation of rumors $\mathbf{R} = [\mathbf{r}^u : \mathbf{r}^{str}]$, as shown in Fig. 5(c).

4.3. Global interaction encoding

We have separately learned about user representation and rumor representation, but the relationship between the representations has been ignored so far. In this section, we will study how to encode multi-modal interactions in rumor diffusion prediction to further investigate the global dependence between users and rumors.

We revisit the rumor interaction hypergraph with a global view. If rumors are regarded as heterogeneous nodes on the corresponding hyperedges, there are two types of relations in Fig. 4(b): rumor-centered relation (such as $u_3 - r_3 - u_5$) and user-centered relation (such as $r_1 - u_4 - r_2$). In order to capture the implicit relation in the rumor interaction hypergraph, we utilize the GAT model by paying attention to the rumor neighbors of users. However, two learned representations of users and rumors are in different semantic spaces. Thus, we first need to align two spaces for further processing, which can be formulated as

$$\begin{aligned} u_i' &= \mathbf{W}_u u_i, \\ r_j' &= \mathbf{W}_r r_j \end{aligned} \tag{7}$$

where $\mathbf{W}_u \in \mathbb{R}^{(d_u \times d)}$ and $\mathbf{W}_u \in \mathbb{R}^{(d_r \times d)}$ are learned parameters.

To encode two relationships into the user representation, a masked-attention mechanism is used here, which only considers the first-order information of neighbors. In other words, each user has a weight of attention regarding their rumor neighbors. The attention coefficient is defined as follows

$$a_{ij} = softmax(LeakReLU(a^T [u_i'; r_j'])) \tag{8}$$

here, $a \in R^{2d \times 1}$ is a learnable parameter, and in the LeakyReLU function [52], the slope of the negative part of the axis is set to 0.2 to ensure that all negative axis information is not lost. Then, a weighted sum of the rumors' features gives a new feature for each user in the following formulation

$$u_i'' = \sigma\Big(\sum_{j \in N(u_i)} a_{ij} \mathbf{W} r_j'\Big) \tag{9}$$

here, $j \in N(i)$ indicates the rumor neighbor of user i. The attention mechanism is enhanced to a Multi-head Attention [53] in order to obtain multiple representations from different relations and enhance the model's capacity for fitting. Specifically, K independent attention mechanisms (Eq. (9)) are executed, and their features are then combined. Finally, the representation of the output feature that follows may be produced as follows

$$u_i'' = \big\Vert_{k=1}^{K} \sigma\Big(\sum_{j \in u_i} a_{ij}^k \mathbf{W}^k r_j'\Big) \tag{10}$$

where $\Vert$ represents concatenation operation, a_{ij}^k and $\mathbf{W}^k$ are the results obtained by executing the kth attention mechanism. More details are shown in Algorithm 1. This algorithm takes a hypergraph and local representations of users and rumors as inputs to generate a global latent representation $\mathbf{U}'$ of users.

Algorithm 1 Global relational coding algorithm

Input: $H = (V, X, W)$; user id u_i; rumor id r_j; weight vector $\mathbf{W}$; Neighborhood function $N(\cdot)$; user representation $\mathbf{U}$; Rumor Representation $\mathbf{R}$
Output: New user representation $\mathbf{U}''$

```
for each i ∈ V, j ∈ X do
    for u_i ∈ N(r_j) do
        Calculate u_i' and r_j' by Eq. (7)
        Calculate a_ij by Eq. (8)
    end for
    Calculate u'' by Eq. (10)
end for
return U''
```

4.4. Rumor diffusion link prediction

For predicting the subsequent users who will generate interactions in the rumor cascades, we apply the learned final user representation $\mathbf{U}''$ to generate the diffusion sequence representation $\mathbf{U}^L = [u_a'', u_b'', u_c'', \dots]$ for each rumor.

To consider the internal dependency and correlation within the rumor diffusion sequence, we utilize a multi-head attention mechanism, which has better memory and could also acquire more information with longer distances than other recurrent neural networks (RNNs). In order to make each user in the sequence only focus on the users before them, we construct a mask matrix $\mathbf{M}$ which is a triangular matrix. The entity of mask matrix $\mathbf{M}$ is denoted as

$$m_{ij} = \begin{cases} 0 & i \le j \\ -\infty & otherwise \end{cases} \tag{11}$$

To be specific, given the same set of queries $\mathbf{Q} \in \mathbb{R}^{d_k}$, keys $\mathbf{K} \in \mathbb{R}^{d_k}$, and values $\mathbf{V} \in \mathbb{R}^{d_v}$, we first make linear projections h times to learn the transformed values of $\mathbf{Q}$, $\mathbf{K}$, $\mathbf{V}$ independently. Then, put them in the attention module in parallel. These outputs from the attention module are concatenated and transformed by another learnable linear projection. Finally, the results are put into a linear layer with residuals to produce the final output, denoted by:

$$\begin{aligned} &h_i = Attention(\mathbf{U}^L\mathbf{W}_i^Q, \mathbf{U}^L\mathbf{W}_i^K, \mathbf{U}^L\mathbf{W}_i^V), \\ &Attention(\mathbf{Q},\mathbf{K},\mathbf{V}) = softmax(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} + \mathbf{M})\mathbf{V}, \\ &MultiHead(\mathbf{Q},\mathbf{K},\mathbf{V}) = Concat(h_1, h_2, \dots, h_h)\mathbf{W}^o, \\ &Z = h + LeakyReLU(h, \mathbf{W}^Z) \end{aligned} \tag{12}$$

Here, $\mathbf{W}_i^Q, \mathbf{W}_i^K \in \mathbb{R}^{d\times d_k}$, $\mathbf{W}_i^V \in \mathbb{R}^{d\times d_v}$, $\mathbf{W}^o \in \mathbb{R}^{hd_v\times d}$, $\mathbf{W}^Z \in \mathbb{R}^{d\times d}$, $\mathbf{Z} \in \mathbb{R}^{L\times d}$, h represents how many parallel attention layers there are.

The resulting representation is projected onto the desired space of class probability using fully connected layers:

$$\hat{y} = \mathbf{W}_2 ReLU(\mathbf{W}_1\mathbf{Z}^T + b_1) + b_2 \tag{13}$$

where, $\hat{y} = \mathbb{R}^{L\times|V|}$, $\mathbf{W}_1 \in \mathbb{R}^{d\times d}$, $W_2 \in \mathbb{R}^{|V|\times d}$, b_1 and b_2 are learnable parameters.

To optimize the model, we select the cross-entropy as the loss function, which is obtained as

$$J(\theta) = -\sum_{i=2}^{L}\sum_{j=1}^{|V|} y_{ij} log(\hat{y}_{ij}) \tag{14}$$

where, $y_{ij} = 1$ means that the rumor was forwarded by the jth user at the ith moment, otherwise it has not.

5. Experiment

5.1. Experimental objectives

Based on the results of these experiments, we want to assess the following three experimental objectives in this section:

- **Objective 1: Evaluation on overall performance.** We aim to compare the effectiveness and generalization of the HG2RLink with other state-of-the-art methods.
- **Objective 2: The rationality of parameter selection.** We desire to find out whether the multi-head attention parameter selection is reasonable due to its optimal performance to the model.
- **Objective 3: Ablation study.** We use the ablation research to evaluate the worth of the modules in our system.

5.2. Datasets

To fill a gap in the public datasets of online social networks for nearly a decade and spark further interest in the exploration of new methods, we introduce a new real-world dataset for predicting rumor dissemination. We implement the evaluation experiments on our new dataset as well as three public datasets, including Twitter [54], Memetracker [55], and Douban [56].

The new dataset was collected from Weibo, which is the mainstream Chinese social media platform. The data collection started at the end of 2022 and continued until spring 2023. We randomly selected 10 seed users and obtained two layers of followees from each seed user and their followees, totaling over 10,600 users and 29 million follow relationships. Then we obtained the public attributes and posted messages of all users, as well as the diffusion sequences of messages (more than 10 reposts), totaling 130,000 messages. Finally, we removed identifiable personalized content from the data. In the experiment, we randomly extracted 1908 diffusion sequences from the Weibo dataset as rumor cascades. We further extracted corresponding users and their following relationships from these diffusion sequences.

Table 1
Descriptive statistics of data.

Dataset	#Nodes 2	#Edges	#Cascades
Weibo	8538	74,699	1908
Twitter	10,236	257,706	2768
Memetracker	3353	1,818,096	3719
Douban	15,596	506,810	5499

Twitter is a popular social media around the world. This dataset was obtained using the Twitter API and includes over 3 million tweets, social relationships, and user retweet behaviors over three weeks in the autumn of 2010. From this dataset, we extracted several user retweet behaviors as rumor cascades and social relationships as social graphs for the experiment.

Memetracker is a website that tracks trends and creates maps of the daily news cycle by examining several news articles and blog posts each day. This dataset collected blogs and news articles from millions of online news websites over a period of three months in April 2009. In our experiment, each website URL is regarded as a node, and each meme is treated as a diffusion sequence. Note that there is no social relationship between two memes. Referring to previous work (Wang et al. 2017), we adopted this method to generate edges and get more than 2.7 million edges in the dataset.

Douban is a Chinese social media platform that allows users to share their comments on books, music, and movies. This dataset contains three different types of user behaviors. These different user behaviors make up more than 750 million user social connections. We consider these behavior data as participation in information dissemination.

We randomly sampled a portion of data from four datasets as our experimental data. The precise statistics utilized in the experiment are shown in Table 1. The sign # represents the number of nodes edges and cascades. We randomly selected 80% of the cascades and corresponding nodes and edges for training and 20% for testing.

5.3. Baselines

To assess the performance of HG2RLink, we select three cutting-edge models as baselines, TopoLSTM [41], FOREST [44], DyHGCN [19], MS-HGAT [45], DisenIDP [46], MIDPMS [57]. These models are shown as follows:

TopoLSTM incorporates topological structures into the standard LSTM model, and proposes a topological recurrent neural network to extract features for diffusion prediction.

FOREST uses reinforcement learning to include the data on macroscopic diffusion in the RNN-based model of microscopic diffusion.

DyHGCN constructs a heterogeneous network with following and reposting interactions and simulates diffusion using an attention method.

MS-HGAT proposes an improved memory embedding lookup module that enables the learning of user representations that emphasize the features within the cascade.

DisenIDP introduces a self-supervised framework, which constructs intent-aware hypergraphs and performs light hypergraph convolution to adaptively activate disentangled intents, and extracts long-term and short-term cascade influences.

MIDPMS integrates macroscopic popularity and microscopic diffusion analysis, incorporating a minimal substitution neural network.

5.4. Evaluation metrics and parameter settings

Due to the large number of potential target nodes, predicting which node will be the next activated node can be viewed as a retrieval

Table 2
Experimental results on Weibo dataset (%).

Model	Hits			Maps		
	@10	@50	@100	@10	@50	@100
TopoLSTM	13.57	24.36	28.68	5.67	6.19	6.25
FOREST	15.73	22.98	26.37	9.67	10.03	10.08
DyHGCN	15.93	26.36	30.24	8.96	9.45	9.50
MS-HGAT	13.84	21.79	25.40	8.03	8.44	8.49
DisenIDP	15.27	21.71	28.15	8.31	8.64	9.19
MIDPMS	16.19	27.13	32.17	10.89	11.16	11.71
HG2RLink(**ours**)	**24.25**	**33.46**	**36.09**	**13.40**	**13.89**	**13.93**

Table 3
Experimental results on the Twitter dataset (%).

Model	Hits			Maps		
	@10	@50	@100	@10	@50	@100
TopoLSTM	16.35	32.67	34.54	13.32	13.66	13.68
FOREST	21.42	33.76	41.51	14.17	14.72	14.83
DyHGCN	23.96	37.75	45.68	15.40	16.03	16.14
MS-HGAT	31.97	48.70	58.85	20.24	21.01	21.15
DisenIDP	29.82	46.17	57.91	18.17	18.56	19.13
MIDPMS	33.48	49.37	59.40	24.26	24.99	25.11
HG2RLink(**ours**)	**36.83**	**52.58**	**60.27**	**26.83**	**27.50**	**27.61**

Table 4
Experimental results on the Memetracter dataset (%).

Model	Hits			Maps		
	@10	@50	@100	@10	@50	@100
TopoLSTM	39.24	61.83	70.41	22.17	22.93	23.04
FOREST	41.77	65.93	73.88	23.81	24.80	24.91
DyHGCN	45.97	67.10	75.37	23.56	24.59	24.71
MS-HGAT	41.12	67.61	78.16	21.72	23.02	23.15
DisenIDP	41.66	66.34	76.67	23.87	24.08	24.23
MIDPMS	46.72	69.90	79.67	24.45	24.71	25.43
HG2RLink(**ours**)	**50.48**	**74.01**	**81.23**	**24.87**	**26.03**	**26.13**

Table 5
Experimental results on the Douban dataset (%).

Model	Hits			Maps		
	@10	@50	@100	@10	@50	@100
TopoLSTM	8.13	15.18	21.16	4.81	5.03	5.32
FOREST	9.47	17.75	23.24	5.17	5.54	5.62
DyHGCN	10.88	20.54	26.74	5.60	6.04	6.12
MS-HGAT	11.93	21.76	28.67	5.99	6.43	6.53
DisenIDP	13.16	23.54	30.53	6.16	6.47	7.09
MIDPMS	14.87	26.19	35.17	7.91	8.24	8.73
HG2RLink(**ours**)	**18.34**	**33.77**	**41.91**	**9.33**	**10.04**	**10.15**

problem. The evaluation metrics we employ for the current diffusion cascades are mean average precision (MAP) and hit ratio (Hits), with Hits@k and MAP@k designating the top k predictions, respectively. Hits@k is used to evaluate the accuracy of the top-k candidates predicted by the model. MAP@k is a mean measure that takes into account both prediction accuracy and relative order. The higher value for the two measures indicates that the analyzed model performed better, on average. The k are $\{10, 50, 100\}$.

Pytorch is used to implement all of the experiment's code. Adam method [58] is used to update the model's parameters. The learning rate is set to 1e−3, the training algorithm's iterations are set to 50, and the training batch size is set to 32. The user and rumor embedding dimensions are both set to 128. Two-layer hypergraph convolution is used to learn the user embedding. The dimensions of convolutional kernel are set to $\{64, 128\}$. We use two layers of multi-head GAT to acquire the global representation of users with rumor embedding. In a multi-head GAT module, the first layer has 8 heads, while the second layer contains 1 head. The dimensions of the two layers are set to $\{8, 128\}$. In the link prediction module, we set the size to 128 and the number of heads in multi-head attention to 8. The first fully connected layer's output dimension is set to 128, and the second fully connected layer's output dimension is the number of users. The output dimension is set to 128 for the first fully connected layer and to the number of users for the second layer. In different modules, we use different dropouts. The settings of dropout are $\{0.3, 0.1, 0.5\}$ in m ulti-head GAT, multi-head attention, and fully connected layer, respectively. Every experiment is run on a computer with a Windows 64-bit system, Intel(R) Core(TM) i7-9700 Processor, CPU @3.00 GHz, and 32 GB memory size.

5.5. Comparison results and performance analysis

5.5.1. Evaluation on overall performance

Effectiveness analysis. On the four experimental data-sets, we assess how well the proposed HG2RLink and these baselines perform. Tables 2, 3, 4, and 5 provide the experimental findings. To ensure the comparability of experimental results, we run all baselines on the extracted datasets instead of directly citing the data in their original papers. It can be observed that there are discrepancies between the performance of these baselines on the extracted datasets and the results reported in their original papers. Specifically, on the Memetracker dataset, the performance of all models improved because we filtered out a large number of extremely short cascades. These short cascades do not provide much help for the model to learn complex implicit relationships between different users. On the Douban dataset, however, the performance of all models decreased. By observing the statistical information of the dataset, this phenomenon can be easily explained. In the extracted new dataset, the ratio of users to cascades increased (2.1 in the original dataset compared to 2.7 in the new dataset), which means there are more implicit relationships between users in fewer cascades. As for the Twitter dataset, this ratio did not change (remained at 3.6), and as a result, the model's performance did not show significant variations. Considering the experiment's findings, the most obvious trend is that the values of Hits@k and MAP@k increase with the number of retrieved candidate nodes. We have the following more detailed observations:

In general, the HG2RLink consistently achieves improvement on the four datasets compared with all baselines. On the Weibo dataset, our model has shown gains of over 4% on Hits@k, as well as over 2% on Map@k. On the Twitter dataset, our model performs 3% better than the baseline models on Hits@10 and Hits@50 and 2% better on MAP@k. On the Memetracker dataset, we can observe that in addition to our model, the performances of other baselines are not bad, this is because of the fact that all cascades in this dataset are short so that all models can easily capture their diffusion structure. In comparison, our model has improved by over 3.7% on Hits@10, over 4.1% on Hits50. On the Douban dataset, our model achieves more than a 3% improvement in the Hits@k metric and more than a 1.4% improvement in the MAP@k metric compared to the other model. According to the combined findings of the four datasets, our model shows an average increase in the Hits@k metric of 4.4% and the MAP@k metric of 1.8%.

Evaluation of generalization. From the model perspective, TopoLSTM only constructs the diffusion sequence into a graph structure, but ignores social network structure. Although FOREST considers social network structure, it only models the diffusion sequence as a sequential pattern, which is insufficient to model the complex diffusion behavior and users' dynamic preferences. DyHGCN takes into account both social relationships and users' dynamic preferences. It can also be seen from the experimental results that their prediction effectiveness is successively improved. Meanwhile, it reflects that network structure and users' dynamic preferences are both important for sequence diffusion prediction. MS-HGAT captures user dynamic preferences using temporal attention mechanisms and highlights the feature within the

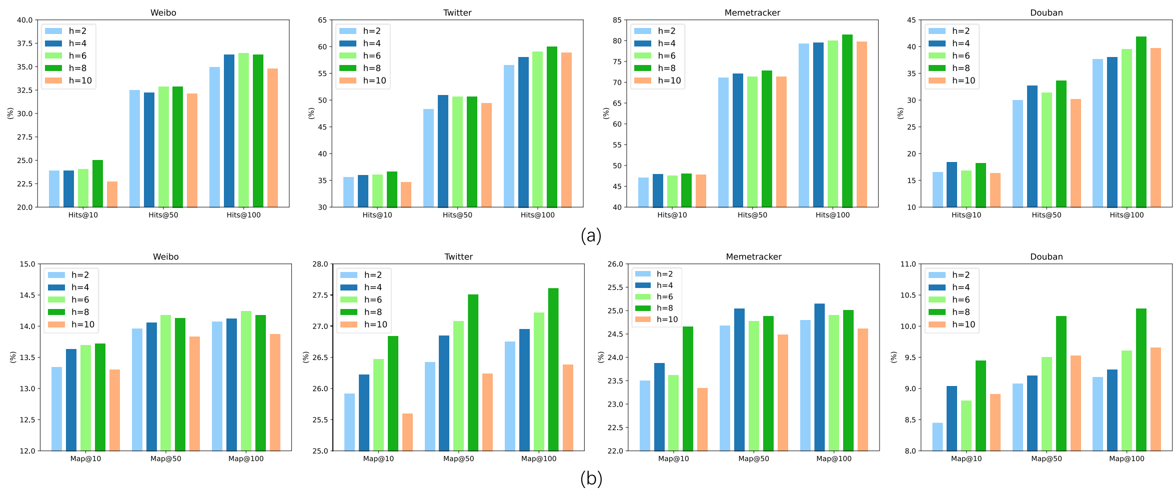

Fig. 6. Number of heads on the Weibo, Twitter, Memetracker, and Douban datasets about Hits@k(a) and MAP@k(b) (%).

cascade with a memory-enhanced embedding lookup module. DisenIDP employs an intent-aware hypergraph to identify users' latent preferences, then decouples preferences through lightweight hypergraph convolution activations. MIDPMS models information dissemination as a substitution system among different messages, taking into account the lifespan of content, user preferences, and the impact of anticipated latent content. Although these methods identify user preferences in various ways, they do not explore the more complex multimodal interactions between users and cascades. Compared with these baselines, our HG2RLink framework not only considers the graph structures of diffusion sequences and social relations, but also models the interrelationships between diffusion sequences. The HG2RLink can better reflect the global preferences of users, because it can unify multimodal interactions that not only include the interaction between users, but also innovatively construct the interaction of structure and user similarities between rumors and the interaction between rumors and users. The experimental results also verify that the global representation helps to predict rumor diffusion. It is worth noting that although Weibo and Twitter have many similarities in many aspects, there are differences in the experimental results of their datasets. We analyze that this difference is mainly due to the diversity of information dissemination modes. In today's SNS, information diffusion is not strictly based on follower relationships as it was 10 years ago, but there are many potential information pushes, which may be formed by system recommendations. That is to say, even if there is no social relationship between two users, they may be pushed to each other because of the same interest or information popularity. Therefore, there are hidden and uncapturable relationships between users in the Weibo dataset, which leads to sparser user social relationships in the dataset than in the Twitter dataset. Nevertheless, our model performs the best on Weibo data, proving that our model has strong generalization ability.

5.5.2. Parameter analysis

In this section, we study how the number of heads h in the multi-head attention mechanism can affect the performance of our framework on three datasets. Thus, we conduct a sensitivity analysis on hyper-parameters h.

Fig. 6(a) shows the different Hits@k values, and Fig. 6(b) shows the corresponding MAP@k values. The horizontal axis represents different evaluation metrics, the vertical axis represents the corresponding values. Different colors indicate different numbers of heads. From these results, we can see that starting from 2 heads, the performance almost gradually improves with increasing number of heads until at $h = 8$, and then starts to decline. This also verifies that multi-head attention allows the model to jointly attend and capture information from different subspaces, thus enhancing the effectiveness of the model. However, experience shows that more heads are not always better. When the

Table 6
Ablation study on Weibo dataset (%).

Model	Hits			Maps		
	@10	@50	@100	@10	@50	@100
HG2RLink	**24.25**	**33.46**	**36.09**	**13.40**	**13.89**	**13.93**
Scheme 1	16.70	23.38	27.90	11.15	11.48	11.54
Scheme 2	13.71	20.19	23.23	10.09	10.39	10.43
Scheme 3	24.82	30.11	33.10	12.78	13.01	13.67

Table 7
Ablation study on Twitter dataset (%).

Model	Hits			Maps		
	@10	@50	@100	@10	@50	@100
HG2RLink	**36.83**	**52.58**	**60.27**	**26.83**	**27.5**	**27.61**
Scheme 1	16.43	25.17	30.14	7.19	7.41	8.34
Scheme 2	15.06	21.86	27.97	6.13	6.91	7.13
Scheme 3	31.88	45.69	54.68	22.93	23.55	23.67

Table 8
Ablation study on Memetracker dataset (%).

Model	Hits			Maps		
	@10	@50	@100	@10	@50	@100
HG2RLink	**50.48**	**74.01**	**81.23**	**24.87**	**26.03**	**26.13**
Scheme 1	30.32	49.17	59.41	13.15	13.57	13.95
Scheme 2	27.13	46.17	55.87	12.39	12.57	13.02
Scheme 3	48.19	68.25	76.32	19.28	19.47	20.13

Table 9
Ablation study on Douban dataset (%).

Model	Hits			Maps		
	@10	@50	@100	@10	@50	@100
HG2RLink	**18.34**	**33.77**	**41.91**	**9.33**	**10.04**	**10.15**
Scheme 1	10.25	16.68	24.02	5.24	5.93	6.26
Scheme 2	9.36	15.49	21.15	5.16	5.34	5.93
Scheme 3	17.25	27.69	35.73	7.15	7.53	8.14

number of heads exceeds a certain threshold, the memory consumption and the cumulative error grow as the number of heads increases. Additionally, the outcomes of the experiment show that when $h > 8$, the performance of the model decreases. Therefore, we choose 8 as the hyper-parameter in the multi-head attention model.

5.5.3. Ablation study

In order to further explore the importance of the relevant steps in the proposed HG2RLink, three ablation studies are implemented on the four datasets over the different parts of the framework. Table 6, 7, 8 and 9 shows the overall performance of three variant schemes and the HG2RLink. The ablation studies are conducted as following orders:

-scheme1 w/o masked multi-head attention in rumor link prediction module: Replace the masked multi-head attention model with a GCN model.

-scheme2 w/o dual-channel representation learning and global relationship coding modules: Remove these two modules, and use randomly initialized user embedding as the input of the rumor link prediction module.

-scheme3 w/o user performance in HGNN model: Remove the user performance matrix $\mathbf{P}$, and use the original incidence matrix $\mathbf{H}$ as the input in the hypergraph convolution layer of HGNN.

We first verify the impact brought by the link prediction module. When replacing the multi-head attention model with a GCN module, we can see a significant effect on all evaluation metrics. The performances drop a lot compared with the HG2RLink. The possible reason is that the GCN module only considers the information fusion of adjacent nodes and ignores the sequence information of the diffusion link. However, the multi-head attention mechanism realizes the allocation of different learning weights to different neighbors, which greatly improves the ability to capture the correlation of spatial information. In summary, the capacity to predict rumor links is improved by the multi-head attention mechanism.

In addition, we evaluate the influence of the global coding representation of users. Referring to Table 6, 7, 8 and 9, the lack of dual-channel representation learning and global relationship coding modules also brings significant performance decline on all evaluation metrics. Obviously, randomly initialized user embedding generated by Deepwalk ignores a variety of interactive relations in the process of rumor diffusion, especially the influence of rumors on users. In our framework, the dual-channel representation learning and global relationship coding modules make the distance among users reposted identical rumors close to each other. The combination of global relationships provides a complementary effect from the global aspect. As a result, the findings indicate that it is crucial to clearly encode the local and global linkages for users and rumors.

Finally, we evaluate the suggested framework without taking users' performance into account. Specifically, before performing the hypergraph convolution layer of HGNN, we remove the users' performance matrix $\mathbf{P}$, and use the original incidence matrix $\mathbf{H}$ of the hypergraph $H = (V, X)$ as the input of subsequent model. Experimental results of scheme3 show that the removal of users' preferences still reduces the performance of prediction. This also reflects that the users' preferences embedded in the user representation are crucial and useful.

Through ablation studies on four diverse datasets, we observed consistent trends. This indicates that HG2RLink exhibits stable and reliable performance across various datasets. Notably, Scheme 3 has demonstrated a performance profile most closely aligned with HG2RLink, particularly excelling in the Hits@10 metric on the Weibo dataset. The challenge in this type of task lies in the precise selection of the top 10 predicted probabilities to define the hit range. Despite this, the demonstrated consistency across various scenarios significantly bolsters our confidence in HG2RLink. It highlights the model's robust generalization ability to adapt to the unique characteristics of diverse datasets.

5.6. Limitations

To investigate the model's limitations under specific case, we analyzed the node attributes of some users during the inference phase. We found that our approach still has certain limitations in the following scenarios:

Case 1: When the degree centrality or eigenvector centrality of users in social networks is very low, the model may find it difficult to accurately predict their behavior due to their low influence.

Case 2: Users who infrequently participate in reposting activities and only appear during specific information cascades present another challenge. Their sparse presence within the hypergraph impedes accurate prediction.

While our model generally shows robust performance, its predictive accuracy is diminished in marginalized scenarios. Future research should explore incorporating more behavioral data to improve the model's adaptability in these situations.

6. Conclusion and future works

With the prevalence of online social networks, once a rumor breaks out, it will quickly spread and it is necessary to timely interrupt its propagation chain. The process of rumor dissemination should be seen as a system, and considering the global information of the system is very important. Combining multimodal interactions among users and rumors can help predict rumor links that are missing or will be possibly formed in the future. Therefore, our work proposes a framework that avoids using historical data and manually designed features by exploring the mechanism of multimodal interactions among users and rumors. Specifically, we proposed a deep neural network framework with global attention based on hypergraph, namely HG2RLink. To model the interactions between users and rumors, we constructed the rumor interactive hypergraph first. Based on the hypergraph, the HG2RLink utilized a variety of machine learning techniques, such as network representation learning, HGNN, and multi-head attention mechanism, to effectively encode the interactions of users and rumors explicitly, and unify multimodal interactions with a global encoding, while considering the influence of user preferences on link prediction. For prediction, the HG2RLink applied the representation of rumor diffusion sequence with sequential characteristics generated by the users' final representation as the input to get the final prediction result. To demonstrate the superior performance of the proposed HG2RLink, we introduced a new Weibo dataset and other three public datasets compared with four baselines. The experimental results showed that the HG2RLink performed better on several different evaluation metrics on four real-world datasets. Meanwhile, the validity of the modules of the HG2RLink framework was verified through an ablation study.

This theoretical implication of our work is that we explored the possibility of multimodal interaction modeling, and unified multimodal interaction to improve the accuracy of rumor diffusion prediction. In addition, its practical implication is that our framework has good transferability and generalization ability. This means that it is a universal framework that can be applied to other types of information diffusion prediction tasks, and it has great application potential. In a word, this framework provides a new idea and method for monitoring and controlling rumors in emergency management, and this new dataset can not only be used for rumor diffusion prediction tasks, but also for other SNS analysis tasks, which is of great significance for academic research.

In the paper, we mainly focused on a variety of interactions among rumors and users. Overall, different rumors and user preferences play important roles in rumor diffusion link prediction. However, the studies on understanding the user representation from the content features and the location of rumors are not deep enough and complete. Therefore, it is potential and valuable to utilize multiple channels of data for rumor diffusion link prediction in future work.

CRediT authorship contribution statement

Qi Zhang: Writing – review & editing, Visualization, Software, Methodology, Formal analysis. **Yuan Li:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Jialing Zou:** Writing – review & editing, Formal analysis, Conceptualization. **Jianming Zhu:** Writing – review & editing. **Dingning Liu:** Investigation. **Jianbin Jiao:** Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors are very grateful for support from the Strategic Priority Research Program of the Chinese Academy of Sciences, China under Grant No. XDA27000000.

References

[1] G. Liang, W. He, C. Xu, L. Chen, J. Zeng, Rumor identification in microblogging systems based on users' behavior, IEEE Trans. Comput. Soc. Syst. 2 (3) (2015) 99–108.
[2] S. Vosoughi, D. Roy, S. Aral, The spread of true and false news online, Science 359 (6380) (2018) 1146–1151.
[3] R. Calo, C. Coward, E.S. Spiro, K. Starbird, J.D. West, How do you solve a problem like misinformation? Sci. Adv. 7 (50) (2021) eabn0481.
[4] Z. Bu, Y. Wang, H.-J. Li, J. Jiang, Z. Wu, J. Cao, Link prediction in temporal networks: Integrating survival analysis and game theory, Inform. Sci. 498 (2019) 41–61.
[5] D. Lin, et al., An information-theoretic definition of similarity, in: ICML, Vol. 98, (1998) 1998, pp. 296–304.
[6] J. Wu, G. Zhang, Y. Ren, A balanced modularity maximization link prediction model in social networks, Inf. Process. Manage. 53 (1) (2017) 295–307.
[7] F. Aghabozorgi, M.R. Khayyambashi, A new similarity measure for link prediction based on local structures in social networks, Phys. A 501 (2018) 12–23.
[8] Y. Yang, J. Zhang, X. Zhu, L. Tian, Link prediction via significant influence, Phys. A 492 (2018) 1523–1530.
[9] S.S. Singh, D. Srivastva, A. Kumar, V. Srivastava, FLP-ID: Fuzzy-based link prediction in multiplex social networks using information diffusion perspective, Knowl.-Based Syst. 248 (2022) 108821.
[10] L. Weng, F. Menczer, Y.-Y. Ahn, Predicting successful memes using network and community structure, in: Proceedings of the International AAAI Conference on Web and Social Media, Vol. 8, (1) 2014, pp. 535–544.
[11] J. Cheng, L. Adamic, P.A. Dow, J.M. Kleinberg, J. Leskovec, Can cascades be predicted? in: Proceedings of the 23rd International Conference on World Wide Web, 2014, pp. 925–936.
[12] Y. Zhang, T. Lyu, Y. Zhang, Hierarchical community-level information diffusion modeling in social networks, in: Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2017, pp. 753–762.
[13] S. Bourigault, S. Lamprier, P. Gallinari, Representation learning for information diffusion through social networks: an embedded cascade model, in: Proceedings of the Ninth ACM International Conference on Web Search and Data Mining, 2016, pp. 573–582.
[14] S. Feng, G. Cong, A. Khan, X. Li, Y. Liu, Y.M. Chee, Inf2vec: Latent representation model for social influence embedding, in: 2018 IEEE 34th International Conference on Data Engineering, ICDE, IEEE, 2018, pp. 941–952.
[15] Y. Xiao, R. Li, X. Lu, Y. Liu, Link prediction based on feature representation and fusion, Inform. Sci. 548 (2021) 1–17.
[16] Y. Xiao, Z. Huang, Q. Li, X. Lu, T. Li, Diffusion pixelation: A game diffusion model of rumor & anti-rumor inspired by image restoration, IEEE Trans. Knowl. Data Eng. 35 (5) (2022) 4682–4694.
[17] Z. Wu, M. Zhan, H. Zhang, Q. Luo, K. Tang, Mtgcn: A multi-task approach for node classification and link prediction in graph data, Inf. Process. Manage. 59 (3) (2022) 102902.
[18] H. Jin, Y. Wu, H. Huang, Y. Song, H. Wei, X. Shi, Modeling information diffusion with sequential interactive hypergraphs, IEEE Trans. Sustain. Comput. 7 (3) (2022) 644–655.
[19] C. Yuan, J. Li, W. Zhou, Y. Lu, X. Zhang, S. Hu, DyHGCN: A dynamic heterogeneous graph convolutional network to learn users' dynamic preferences for information diffusion prediction, in: Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2020, Ghent, Belgium, September 14–18, 2020, Proceedings, Part III, Springer, 2021, pp. 347–363.
[20] S. Molaei, H. Zare, H. Veisi, Deep learning approach on information diffusion in heterogeneous networks, Knowl.-Based Syst. 189 (2020) 105153.
[21] Y. Li, J. Zhu, J. Jiao, Q. Zhang, Competitive influence minimization in multi-group social networks: An opinion-based solution, IEEE Trans. Netw. Sci. Eng. 9 (4) (2022) 2617–2630.
[22] Y. Feng, H. You, Z. Zhang, R. Ji, Y. Gao, Hypergraph neural networks, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, (01) 2019, pp. 3558–3565.
[23] N.N. Daud, S.H. Ab Hamid, M. Saadoon, F. Sahran, N.B. Anuar, Applications of link prediction in social networks: A review, J. Netw. Comput. Appl. 166 (2020) 102716.
[24] D. Liben-Nowell, J. Kleinberg, The link prediction problem for social networks, in: Proceedings of the Twelfth International Conference on Information and Knowledge Management, 2003, pp. 556–559.
[25] B. Jiang, J. Liang, Y. Sha, R. Li, L. Wang, Domain dictionary-based topic modeling for social text, in: Web Information Systems Engineering–WISE 2016: 17th International Conference, Shanghai, China, November 8-10, 2016, Proceedings, Part I 17, Springer, 2016, pp. 109–123.
[26] C. Ahmed, A. ElKorany, Enhancing link prediction in Twitter using semantic user attributes, in: Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015, 2015, pp. 1155–1161.
[27] S. Kopeinik, E. Lex, P. Seitlinger, D. Albert, T. Ley, Supporting collaborative learning with tag recommendations: a real-world study in an inquiry-based classroom project, in: Proceedings of the Seventh International Learning Analytics & Knowledge Conference, 2017, pp. 409–418.
[28] C.K. Leung, F. Jiang, T.W. Poon, P.-É. Crevier, Big data analytics of social network data: who cares most about you on facebook? in: Highlighting the Importance of Big Data Management and Analysis for Various Applications, Springer, 2018, pp. 1–15.
[29] A. Shahmohammadi, E. Khadangi, A. Bagheri, Presenting new collaborative link prediction methods for activity recommendation in Facebook, Neurocomputing 210 (2016) 217–226.
[30] L. Li, J. He, M. Wang, X. Wu, Trust agent-based behavior induction in social networks, IEEE Intell. Syst. 31 (1) (2016) 24–30.
[31] E. Bütün, M. Kaya, R. Alhajj, Extension of neighbor-based link prediction methods for directed, weighted and temporal social networks, Inform. Sci. 463 (2018) 152–165.
[32] M. Coskun, M. Koyutürk, Link prediction in large networks by comparing the global view of nodes in the network, in: 2015 IEEE International Conference on Data Mining Workshop, ICDMW, IEEE, 2015, pp. 485–492.
[33] A. De, S. Bhattacharya, S. Sarkar, N. Ganguly, S. Chakrabarti, Discriminative link prediction using local, community, and global signals, IEEE Trans. Knowl. Data Eng. 28 (8) (2016) 2057–2070.
[34] P. Wang, B. Xu, Y. Wu, X. Zhou, Link prediction in social networks: the state-of-the-art, 2014, arXiv preprint arXiv:1411.5118.
[35] S.S. Singh, S. Mishra, A. Kumar, B. Biswas, CLP-ID: Community-based link prediction using information diffusion, Inform. Sci. 514 (2020) 402–433.
[36] S. Dong, X. Miao, P. Liu, X. Wang, B. Cui, J. Li, HET-KG: Communication-efficient knowledge graph embedding training via hotness-aware cache, in: 2022 IEEE 38th International Conference on Data Engineering, ICDE, IEEE, 2022, pp. 1754–1766.
[37] C. Wang, X. Wang, Z. Li, Z. Chen, J. Li, Hyconve: A novel embedding model for knowledge hypergraph link prediction with convolutional neural networks, in: Proceedings of the ACM Web Conference 2023, 2023, pp. 188–198.
[38] B. Cai, Y. Xiang, L. Gao, H. Zhang, Y. Li, J. Li, Temporal knowledge graph completion: A survey, 2022, arXiv preprint arXiv:2201.08236.
[39] C. Li, J. Ma, X. Guo, Q. Mei, Deepcas: An end-to-end predictor of information cascades, in: Proceedings of the 26th International Conference on World Wide Web, 2017, pp. 577–586.
[40] X. Feng, Q. Zhao, Z. Liu, Prediction of information cascades via content and structure proximity preserved graph level embedding, Inform. Sci. 560 (2021) 424–440.
[41] J. Wang, V.W. Zheng, Z. Liu, K.C.-C. Chang, Topological recurrent neural network for diffusion prediction, in: 2017 IEEE International Conference on Data Mining, ICDM, IEEE, 2017, pp. 475–484.
[42] C. Yang, H. Wang, J. Tang, C. Shi, M. Sun, G. Cui, Z. Liu, Full-scale information diffusion prediction with reinforced recurrent networks, IEEE Trans. Neural Netw. Learn. Syst. (2021).
[43] X. Chen, F. Zhou, K. Zhang, G. Trajcevski, T. Zhong, F. Zhang, Information diffusion prediction via recurrent cascades convolution, in: 2019 IEEE 35th International Conference on Data Engineering, ICDE, IEEE, 2019, pp. 770–781.

[44] C. Yang, M. Sun, H. Liu, S. Han, Z. Liu, H. Luan, Neural diffusion model for microscopic cascade study, IEEE Trans. Knowl. Data Eng. 33 (3) (2019) 1128–1139.
[45] L. Sun, Y. Rao, X. Zhang, Y. Lan, S. Yu, MS-HGAT: memory-enhanced sequential hypergraph attention network for information diffusion prediction, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 36, (4) 2022, pp. 4156–4164.
[46] Z. Cheng, W. Ye, L. Liu, W. Tai, F. Zhou, Enhancing information diffusion prediction with self-supervised disentangled user and cascade representations, in: Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, 2023, pp. 3808–3812.
[47] J. Li, W. Chen, Y. Liu, J. Yang, D. Zeng, Z. Zhou, Neural ordinary differential equation networks for fintech applications using internet of things, IEEE Internet Things J. (2024).
[48] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, Y. Bengio, Graph attention networks, 2017, arXiv preprint arXiv:1710.10903.
[49] B. Perozzi, R. Al-Rfou, S. Skiena, Deepwalk: Online learning of social representations, in: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014, pp. 701–710.
[50] S. Salvador, P. Chan, Toward accurate dynamic time warping in linear time and space, Intell. Data Anal. 11 (5) (2007) 561–580.
[51] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, 2013, arXiv preprint arXiv:1301.3781.
[52] A.L. Maas, A.Y. Hannun, A.Y. Ng, et al., Rectifier nonlinearities improve neural network acoustic models, in: Proc. ICML, Vol. 30, (1) Atlanta, GA, 2013, p. 3.
[53] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, in: Advances in Neural Information Processing Systems, vol. 30, 2017.
[54] N.O. Hodas, K. Lerman, The simple rules of social contagion, Sci. Rep. 4 (1) (2014) 4343.
[55] J. Leskovec, L. Backstrom, J. Kleinberg, Meme-tracking and the dynamics of the news cycle, in: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2009, pp. 497–506.
[56] E. Zhong, W. Fan, J. Wang, L. Xiao, Y. Li, Comsoc: adaptive transfer of user behaviors over composite social network, in: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2012, pp. 696–704.
[57] R. Wang, X. Xu, Y. Zhang, Multiscale information diffusion prediction with minimal substitution neural network, IEEE Trans. Neural Netw. Learn. Syst. (2023).
[58] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, 2014, arXiv preprint arXiv:1412.6980.