IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 18 April 2024, accepted 10 May 2024, date of publication 15 May 2024, date of current version 3 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3401460

APPLIED RESEARCH

Dynamic Heterogeneous Graph Learning: An Adaptive Research Academic Network

MINGHONG YANG[1] **AND ZHENHUA YANG**[2]

[1]School of Economics and Management, Hunan Applied Technology University, Changde 415000, China
[2]School of Information Engineering, Hunan Applied Technology University, Changde 415000, China

Corresponding author: Zhenhua Yang (yzh168@hatu.edu.cn)

This work was supported in part by the Social Science Committee of Changde, Hunan, China, under Grant CSP22ZZ06.

ABSTRACT Accurately assessing the impact of academic achievements holds significant importance for scholars in their literature review process and for the retrieval and recommendation of scientific research databases. Predicting this impact presents a formidable challenge. Furthermore, scientific research and academic networks exhibit dynamic evolutionary characteristics. Over time, not only does the semantic information of keywords, journals, and other nodes undergo transformation, but the strength of connections between distinct nodes also experiences fluctuations. In recent years, with the advancement of deep learning technologies, particularly the introduction of recurrent neural networks, graph neural networks, and related architectures, a powerful tool for data representation learning has emerged. This paper adopts a dynamic graph representation learning perspective, aiming to adaptively derive vectorized representations for each node through the design of a trainable neural network, ultimately predicting the influence of academic achievements. The study centers on publicly available scientific research and academic networks like APS and AMiner, employing the citation count of papers as the evaluation metric for influence. Specifically, this research will formulate a trainable neural network from a data-driven standpoint to dynamically capture semantic information and evolutionary trends within dynamic graph structures, subsequently generating corresponding vector representations for individual nodes. Upon acquiring the semantic representation of the article, the future citation count can be forecasted through the design of a straightforward mapping function, such as a multi-layer perceptron. Furthermore, through the analysis of node representations (including authors, journals, etc.), it is possible to uncover and explore the evolutionary patterns of individuals and groups within the scientific research academic network.

INDEX TERMS Dynamic heterogeneous graph, extraction, graph neural network.

I. INTRODUCTION

Predicting the impact of academic achievements is a challenging problem [33], [34], [35], [36]. The most common approach relies on domain experts' judgment, where experts assess the quality of an article by reading its title, abstract, body, and other content, drawing upon their experience. However, this method is time-consuming and labor-intensive. In reality, the academic realm of scientific research forms a network structure, where an article is typically linked to various nodes such as journals, keywords, authors, and references. The future impact of academic achievements is influenced not only by the content of the achievements themselves, but also by the overall developmental trends of the discipline and the influence of the journals in which they are submitted. Furthermore, the scientific research and academic network exhibits dynamic evolutionary characteristics. As time progresses, not only does the semantic information of nodes like keywords and journals undergo changes, but also the strength of connections between different nodes fluctuates. As a result, conventional methods such as text analysis [16], machine learning [25], deep learning [22], and statistical analysis [19] are ineffective in addressing the impact of academic achievements.

The associate editor coordinating the review of this manuscript and approving it for publication was Bing Li.

A. THE CRUX OF THE ISSUE

The key to accurately predicting the impact of academic achievements lies in simulating the mutual relationships and dynamic evolution of nodes within the network. In fact, this can be viewed as a problem of dynamic heterogeneous graph structure representation learning [37], [38]: with the publication of papers, heterogeneous nodes such as papers, authors, and journals will establish new interaction relationships, and new nodes will emerge (such as new researchers, the creation of new journals, etc.). Designing a representation learning mechanism to extract useful semantic features for each node from this network is crucial for influence prediction.For the dynamic academic research network, this project plans to devise a multi-step update neural network for learning the representation of dynamic graphs, generating a vector representation for each type of node, and employing it for predicting paper citations. The crux of this paper lies in modeling based on dynamic graph representation learning. In recent years, with the development of deep learning technology, especially the introduction of structures like recurrent neural networks [39], [40] and graph neural networks [41], [42], a potent tool has been provided for data representation learning. Approaching from the perspective of dynamic graph representation learning, we have designed a trainable neural network to adaptively obtain vectorized representations for each node and forecast the influence of academic achievements.

B. PURPOSE OF THE STUDY

In recent years, Graph Neural Networks (GNNs) have demonstrated outstanding performance in graph feature extraction and representation learning. To better address real-world scenarios, some research works have proposed Heterogeneous Graph Neural Networks and Dynamic Graph Neural Networks. For example, [12] and [26] utilize attention-based and meta-path mechanisms to encode heterogeneous graphs. Additionally, [8], [11], and [24] employ representation learning to model heterogeneous graphs. On the other hand, [20] and [23] view dynamic graphs as a series of equally spaced sampled graph snapshots, while [4], [21], and [27] represent dynamic graphs as a series of continuous event sets. References [10], [18], [20], and [32] introduce deep learning into graph learning and employ Graph Convolutional Neural Networks to tackle dynamic graph problems.Although this study investigates the representation learning problem of dynamic heterogeneous scientific research academic networks, methods based on heterogeneous graph and dynamic graph representation learning cannot be directly applied in this paper. This is because these methods do not address the heterogeneity in the prediction of new paper citations. The design of interactive relationships and dynamic evolution laws poses a challenge.

In recent years, the number of scientific pulications has been growing in a dramatic rate.For example, the numbers of submissions and accepted papers of AAAI 2023 have increased to 8,777 and 1,721 respectively. Given the huge volume of scholarly papers, a long-standing research challenge is how to effectively evaluate the impact of scientific literature [46], [47]. A typical way to measure the impact of a scholarly paper is through the number of citations received after publication [48], reflecting the influence in the research community.In addition, the h-index and 10-index indicators based on citation volume are also used to measure the academic influence of authors, research institutions, etc. Therefore, accurately predicting the citation volume of a paper can provide decision-making assistance and basis for scenarios such as researchers reading literature, research institutions hiring researchers, and foundations issuing funds. The difficulty in predicting the number of citations for newly published papers in the future is attributed to the multitude of factors influencing citation counts. Reference [49] When predicting the citation count of a new paper, considerations extend beyond the article and its keywords, encompassing the current developmental trends of the discipline, the prominence of the author, and the journal in which the article is published, among others. Additionally, the entire scientific research and academic network undergoes dynamic changes and cannot be simply modeled statically. A crucial issue we must tackle is the ''cold start'' problem of citation forecasting. This entails achieving citation forecasts for the first few years after publication in the absence of any historical citation data. Consequently, our study cannot rely on simply constructing time series forecasts based on historical citation data.

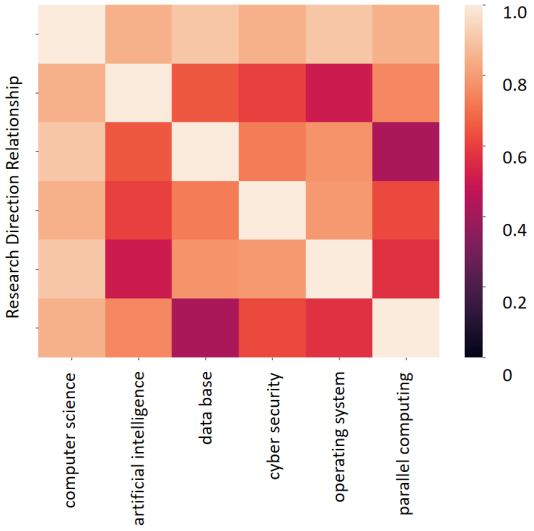


FIGURE 1. The heatmap of cosine similarity in research fields.

In this article, we utilize the SciBERT [43] model to encode and analyze the research direction.The cosine similarity of normalized vectors is employed to measure the similarity between different research fields. The resulting heatmap is depicted in Figure1. As shown in Figure1, there is a high degree of similarity between research directions, and a significant amount of interaction exists among different fields in the academic research network. Therefore, employing

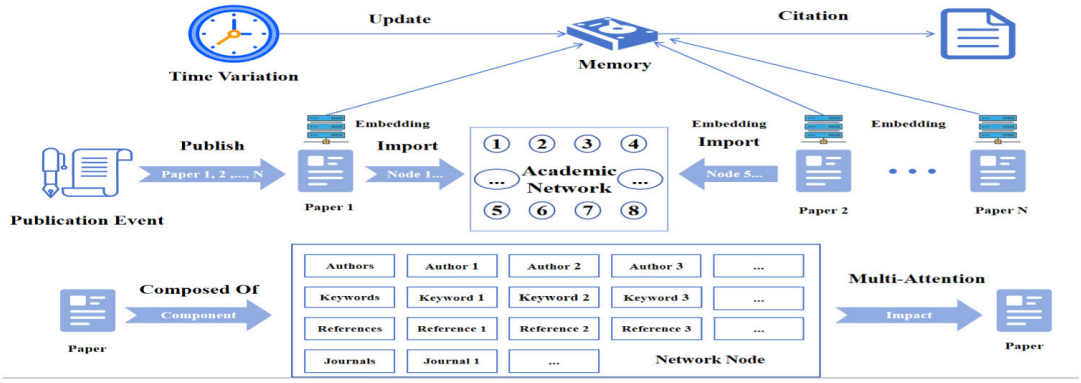


FIGURE 2. Given the publication events, the academic influence prediction model aims to predict the citation of new paper.

graph neural networks for modeling in this context is justified. As shown in Figure 2, with the change of time, publishing events continue to occur, and more and more nodes are introduced into the academic network. At the same time, through the relationships between different nodes, a multi attention mechanism is used to update the memory representation of the original nodes, and finally, the number of citations for newly published papers is obtained.

II. PROBLEM AND RELATED WORK

A. PROBLEM DEFINITION

We propose a modeling approach for the research-academic network as a dynamic heterogeneous graph. As illustrated in Figure 3, this network is centered around articles and encompasses four types of nodes and their corresponding relationships, including cited papers, authors, keywords, and journal publishers. In reality, research-academic networks undergo continual evolution. For instance, new papers are regularly published, new researchers integrate into the network, and new keywords emerge. These novel entities and their corresponding connections are integrated into the network. The dynamic evolution of the network is a sequence of article publication events. Since each article encompasses nodes such as authors, keywords, journals, references, and more, each publication event can be viewed as a subgraph centered around the recently published articles, which we refer to as the meta-information subgraph.

We formally define a dynamic heterogeneous information network as follows:

Meta Information Subgraph:

$$g_i = (p_i, A_i, P_i, V_i, K_i, t_i)$$

which indicates an event that an article was published. Among them, p_i represents the newly published article, A_i, P_i, V_i, K_i, t_i are the author collection of the new article, the reference collection, the journal conference collection (usually contains only one element), and the keyword collection. Where t_i is the publication time of the article. A scientific research and academic network that dynamically evolves to the moment of t satisfying $t_1 \leq t_2 \cdots \leq t_n \leq t$.

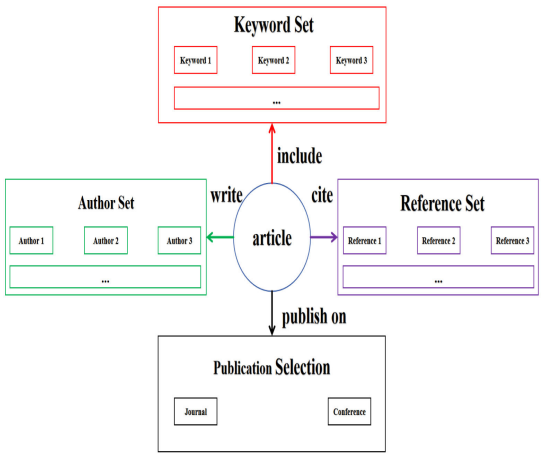


FIGURE 3. The relationship between each node of the network point.

For an article u we use to denote the cumulative citations the article has received $\hat{C}_u^l$ years after it was published.

The core idea of the solution in this paper is to maintain a dynamically changing state representation for each node u in the network to capture the dynamic change trend($S_u \in R^{d_s}$) and semantic information of nodes in the network, and to predict citations based on the state representation. The method design of this paper is intended to include three parts: message generation, node representation update, and citation generation.

B. RELATED WORK

In the current paper citation prediction problem, there exist two classical approaches. One classic method is rooted in the prediction of historical citations [1], which models the historical citations of papers as a time series. This allows for the extraction of features from citations occurring after a certain period following the publication of articles, thereby enabling the prediction of future article citations. Similar methodologies are applied in [2] and [15]. Another method pertains to the prediction of new articles. This involves using observational data from nodes associated with recently

published papers to discern the characteristics of these articles.

1) PREDICTION METHOD BASED ON HISTORICAL CITATIONS

This approach models the historical citations of papers as a time series, extracting features from citations occurring after a certain period post-publication, and predicting the citation count for future papers. Within this category, some methods parameterize the variation in article citations. Another subset of methods autonomously learns the evolution pattern from historical observational data of articles through neural networks [31]. For instance, there exist both $RNN-based$ and $LSTM-based$ machine self-learning methods.

2) FEATURE ENGINEERING

This category of methods focuses on predicting new articles without relying on historical observations. Several feature engineering-based approaches have been proposed [3], [9]. For instance, ranking-based features, including author and keyword rankings, are extracted as part of the features for new articles [29]. Following this approach, feature engineering methods extract attributes for each article from the network by manually specifying indicators such as author h-index, journal ranking, and the number of article keywords. However, manual rule design is not only time-consuming, but may also not fully account for the dynamic evolution characteristics and intricate node interactions in the network.

3) HINTS [13]

This method represents the first end-to-end prediction framework not reliant on feature engineering, as depicted in Figure 5. It views the dynamically evolving scientific research and academic network as a sequence of network snapshots sampled at uniform intervals. HINTS achieves the prediction of new papers by integrating graph neural network, recurrent neural network, and time series generation model components. The HINTS method stands as a relatively cutting-edge approach in current paper citation prediction. However, discretizing the dynamically evolving network into a series of equally spaced network snapshots leads to a loss of temporal information in the network. This is because we cannot ascertain the sequence of newly added edges in a network snapshot relative to the preceding snapshot.

In addition, this paper compares a dynamic graph framework called DyRep [44]. DyRep is a general dynamic graph network framework that models dynamic changes in graph networks. DyRep decomposes the evolution of the graph into two different dynamic processes: the associate process and the communication process. DyRep consists of three different modules: local propagation, self-propagation, and external force driving. DyRep is an inductive graph representation method that can handle the addition of new nodes and supports the addition of node and edge attributes. The model structure of DyRep is shown in Figure4.

While the aforementioned methods offer valuable insights for predicting paper citations, especially for new papers, they do not comprehensively address the modeling of intricate interaction patterns in the network or the capture of fine-grained temporal evolution laws.

III. METHODOLOGY

In this section, we commence by presenting the fundamental concept of the proposed ACPGNN (Adaptive Citation Prediction Graph Neural Network) model. Subsequently, we introduce the step-by-step framework of the model components. As illustrated in Figure 3, we conceptualize the research-academic network as a dynamic heterogeneous graph centered around articles. This network encompasses four types of nodes and their corresponding relationships. Since each article encompasses nodes such as authors, keywords, journals, and references, every article publication event can be viewed as a subgraph focused on newly published articles, which we refer to as the meta-information subgraph. Upon the occurrence of each event, the nodes and edges associated with the subgraph are integrated into the scientific research and academic network. The dynamic evolution of the network constitutes a sequence of article publication events.

We devise a multi-step update neural network to learn the representation of the dynamic scientific research academic network graph, thereby generating a vector representation for each type of node. This representation is subsequently utilized for the prediction of paper citations. Our model comprises three key components: message generation, node representation update, and citation generation. Specifically, upon the observation of a new paper being published, components are employed to encode paper-related information, capture the semantic relevance of nodes, and dynamically update their memory based on their preceding memory and the currently encoded generated messages. Simultaneously, we aggregate the representation of the paper information from static, dynamic, and hierarchical perspectives, design the sequence to generate the corresponding prediction neural network, and extract information reflecting the predicted change trend of the newly published article from the representation of all nodes in the network. This enables the analysis and forecast of the paper's citations.

A. VECTOR INITIALIZATION

For the initialization of each node in the network, we have considered factors that may affect its representation. Taking into account past perspectives, such as [33], [34], [35], and [36], it is necessary to include current research hot spots, the author's personal influence, and the quality of the article. Author nodes: Professional title, affiliation, total citation amount. Thesis nodes: date, citation count. Journal and conference nodes: impact factors, SCI level, average citation rate. Each node in the network has its own unique number and type identifier. Keyword nodes are distinguished only by unique numbers, corresponding to a specific number

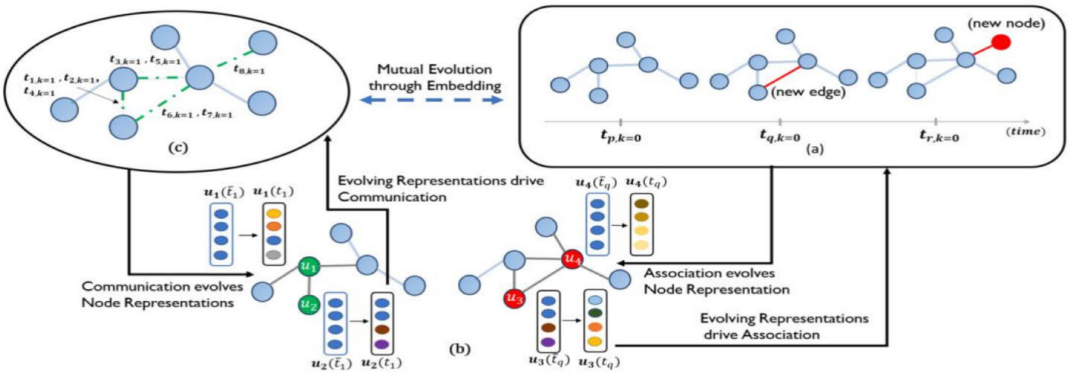


FIGURE 4. Baseline Model DyRep Structure Diagram.

table used to explain specific content. The citation count of newly published papers can be considered as 0. The citation count of newly emerging authors can be considered as 0. Neighborhood relationships in the network can be used to illustrate: the references of a paper, the articles published by the author, the keywords of the paper, the author's research direction, etc.

B. MESSAGE GENERATION

A unique perspective of our approach is to represent each new article publication as an event and learn a dynamic representation of nodes from an event-based dynamic graph, which corresponds to a chronological sequence of events. With the occurrence of meta-information subgraphs, the semantic information of corresponding nodes in the network will change. For the paper publication generation event under the timestamp, we encode the meta-information subgraph to generate update signal $m_i(k)$ for the corresponding node k involved in the event to capture this semantic information change and update the memory of the node. For time information, we design periodic encoding neural network to encode. For an edge-level interaction event between nodes i and j at time t, the message generation transfer function is as follows:

$$m_i(t) = msg_s(s_i(t^-), s_j(t^-), \Delta t, e_{ij}(t)) \quad (1)$$

$$m_j(t) = msg_d(s_j(t^-), s_i(t^-), \Delta t, e_{ij}(t)) \quad (2)$$

where$s_i(t^-)$, $s_j(t^-)$ represents the node feature representation before the t timestamp, Δt represents the time difference between the events after periodic encoding, $e_{ij}(t)$ represents the interaction event between the node i and the node j at the moment t, and msg_s, msg_d represents the learnable information transfer function. We choose to use horizontal concatenation of inputs as our function for brevity.

Batch event processing of messages may cause the same node to update nodes multiple times at the same time, so we use an aggregation method to aggregate messages. For the information $m_i(t_1), m_i(t_2), \cdots, m_i(t_b)$ of the time range $t_1, t_2, \cdots, t_b \leq t$, the message aggregation of the node at

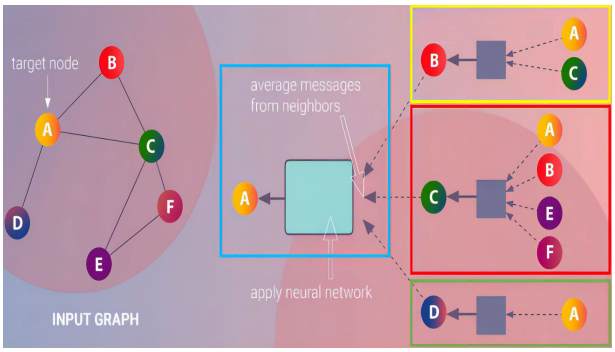


FIGURE 5. Diagram of Message Generation and Aggregation Process.

time is as follows:

$$m_i(t) = agg(m_i(t_1), m_i(t_2), \ldots, m_i(t_b)) \quad (3)$$

Among them, the aggregation message function agg must satisfy the permutation invariance to achieve that the function output is independent of the permutation of the input parameters, and needs to satisfy the differentiability to achieve the gradient transfer. We use two mechanisms of selecting the average message and the latest message as our function agg. The overall schematic is shown in the figure 5:

C. NODE REPRESENTATION UPDATE

During training, the model's memory state encompasses a vector for each node $s_i(t)$ that the model has encountered up to that point. When a node partakes in an interaction event, its memory undergoes an update. The purpose of the memory module is to encapsulate the historical information of nodes in a condensed format. This allows for the long-term retention of the topology of each node in the dynamic graph, its own previous information, and its dependencies. Simultaneously, the incorporation of global memory in the model enables the tracking of the evolution of the entire temporal network. It also facilitates predictions based on the overall graph structure level according to the global memory.

Following the dispatch of the node update message triggered by the edge interaction event, the node representation

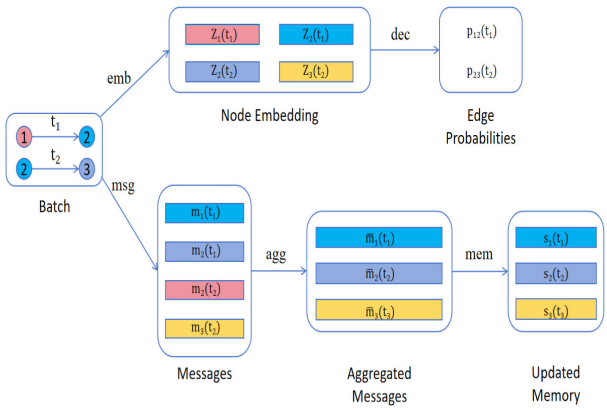


FIGURE 6. Node Representation Update Diagram.

update mechanism refines the node representation of the corresponding node pair by amalgamating the update signal. The formal representation is as follows:

$$s_i(t) = mem(\overline{m}_i(t), s_i(t^-)) \tag{4}$$

Among them $s_i(t^-)$ is the representation of the node before the message event occurs, and the update function needs to select a learnable recurrent neural network, which we select *LSTM* here as our update function.

D. ENCODE-DECODE FRAMEWORK

In this paper, the data comprises variable-length time series. Attributes of paper nodes, such as keywords and referenced papers, as well as information regarding paper authors and publication venues, are all of variable lengths. Therefore, we employ an Encoder-Decoder framework for processing. The input, represented as $X_1, X_2, \ldots, X_m$, denotes variable-length paper attributes, such as keywords. The output, represented as $Y_1, Y_2, \ldots, Y_n$, signifies the processed fixed-length vector. This allows us to map variable-length sequences into fixed-length data as input for the model. Specifically, our input sequence consists of variable-length attributes, $X_1, X_2, \ldots, X_m$, from paper attribute nodes. Each X_i represents a piece of input paper attribute data, potentially with varying lengths. Our objective is to map this variable-length paper attribute sequence into a fixed-length vector, which serves as the input for our graph neural network model in this paper. To achieve this goal, we employ an Encoder-Decoder framework, in which the encoder maps the variable-length sequence into a fixed-length vector, and the decoder decodes this vector into the output sequence $Y_1, Y_2, \ldots, Y_n$. Within this framework, we utilize recurrent neural networks to implement the functionalities of the encoder and decoder. Through this approach, we can handle paper attribute data of varying lengths and map them into fixed-length vectors, as shown in figure 7.

E. CITATION GENERATION

Based on the node state update performed by the above modules, we encode the feature of the node i at time t

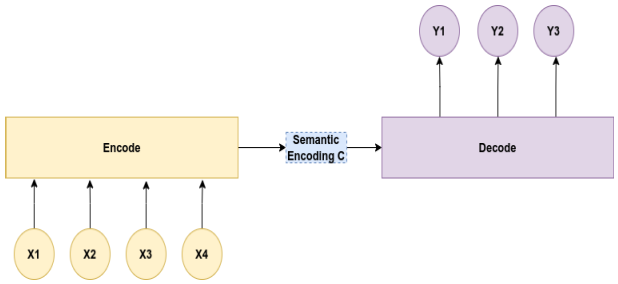


FIGURE 7. Encoder-Decoder framework.

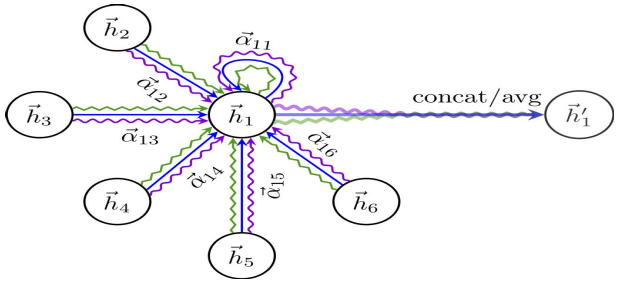


FIGURE 8. multi-head attention mechanism.

to obtain the embedded representation $z_i(t)$. The formal expression is as follows:

$$z_i(t) = emb(i,t) = \sum_{(j \in n_i^k([0,t]))} h(s_i(t), s_j(t), e_{ij}, v_i(t), v_j(t)) \tag{5}$$

Here $S_i(t), S_j(t)$ is the memory vector representation of the node i, j under time t, e_{ij} is the interaction vector representation between the representative and the node i, j, $v_i(t)$, $v_j(t)$ is the node-level event vector representation of the node i, j under time t, and h is the learnable neural network. Here we adopt a multi-layer neural network to iteratively aggregate and learn the representation of the nodes' neighbors of different orders.

$$h_i^{(l)} = MLP^l(h_i^{(l-1)}(t) \parallel \widetilde{h}_i^{(l)}(t)) \tag{6}$$

$$\widetilde{h}_i^{(l)}(t) = Attention^{(l)}(q^{(l)}(t), K^{(l)}(t)) \tag{7}$$

$$q^{(l)}(t) = h_i^{(l-1)}(t) \parallel \Theta(0) \tag{8}$$

$$K^{(l)}(t) = [h_1^{(l-1)}(t) \parallel e_{i1}(t_1) \parallel \Theta(t - t_1), \cdots, h_N^{(l-1)}(t) \parallel e_{iN}(t_N) \parallel \Theta(t - t_N)] \tag{9}$$

Here *MLP* plays the role of fusing the previous hop neighbor representation $h_i^{(l-1)}(t)$ and aggregated information $\widetilde{h}_i^{(l)}(t)$ and performing dimensional transformation, refers to the multi-head attention mechanism(As shown in the figure 8) in *TGAT*, where || represents the concatenation of vectors, $\Theta(\cdot)$ represents the periodic encoding neural network of time. For the multi-head attention mechanism, we have improved the original *TGAT*, adding the representation of node-level events and node features, so as to realize the rapid fusion of graph node structure features on the attention layer.

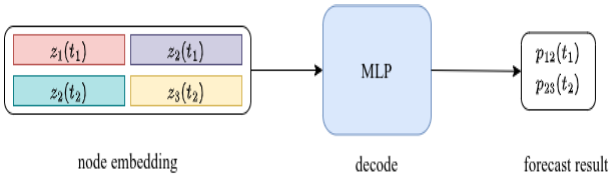


FIGURE 9. Schematic diagram for generating prediction results.

We devise a dedicated prediction neural network for sequence generation, extracting information that mirrors the anticipated change trend of newly published articles from the representations of all nodes in the network to formulate corresponding predictions, as depicted in Figure 6 below. Concerning information encompassing the global graph structure and evolution pattern, we sieve out any extraneous details utilizing a multi-layer perceptron. We then extract pertinent information regarding paper citations, which serves as our prediction outcome:

$$Prediction = MLP(z_i(t)) \tag{10}$$

F. MATHEMATICAL THEORY ANALYSIS

The dynamic graph neural network's expressive capacity is pivotal in assessing the range of problems it can effectively address. This assessment hinges on whether the network has the potential to overfit the training set, as well as the investigation into its generalization abilities, which pertain to its capacity to maintain performance levels from the training set to the test set.

A conventional approach for gauging the expressiveness of dynamic graph neural networks is through function approximation, which delves into the spectrum of functions that the neural network can accurately represent. The Universal Approximation Theorem stands as a prominent result in this context. However, the simplified feedforward structure of dynamic graph neural networks introduces specific biases related to graph characteristics, which complicates the extension of these results. Nevertheless, we can still rely on its portrayal of expressive capacity, which posits that the neural network's capabilities can be described as a set representing all the tasks it can perform. This perspective aligns with findings in computational complexity theory.

This theory categorizes problems based on their level of complexity, grouping them into sets with similar complexities. The exploration of the expressive capacity of dynamic graph neural networks is rooted in the principles of computational complexity. For instance, [28] attributed the expressive capacity of GNNs to the Graph Isomorphism Test, while [17] discussed the constraints of GNNs from the standpoint of Turing Universality. Additionally, [6] established the Subgraph Counting problem as a means to discern the capabilities of different GNNs. Furthermore, [7] revisited classic concepts, demonstrating that GNNs can approximate any function with permutation invariance on the graph.

G. COMPLEXITY ANALYSIS

Mathematically, the matrix form is represented as follows:

$$H^{(0)} = X \tag{11}$$

$$H^{(k+1)} = f(H^{(k)}, A) = \sigma(AH^{(k)}W^{(k)}) \tag{12}$$

where $A \in R^{n\times n}$ is the adjacency matrix, $H \in R^{n\times d_{in}}$ is the feature matrix, and $H \in R^{d_{in}\times d_{out}}$ is the training weight of each layer Matrix (note that the parameters are shared by all layers here). This can be divided into two steps: message aggregation and feature update. The first part of the message aggregation, that is, the left multiplication of A and H (sparse matrix multiplied by dense matrix, graph calculation), is equivalent to considering the ith row of A, that is, the ith node v_i, and the jth feature of its adjacent nodes. Aggregate to get (i, j) matrix elements.

$$(AH^{(k)})_{ij} = \sum_i a_{il}h_{lj} = \sum_{v_l\epsilon N(v_i)} h_{vlj} \Longrightarrow h_i = \sum_{v_l\epsilon N(v_i)} h_l \tag{13}$$

The second part of feature update is to multiply the aggregated features with the parameter matrix (dense matrix multiplied by dense matrix, deep learning).

$$h_i^{(k+1)} = (\sum_{v_l\epsilon N(v_i)} h_l^{(k)})W^{(k)} = \sum_{v_l\epsilon N(v_i)} h_l^{(k)}W^{(k)} \tag{14}$$

Since matrix multiplication conforms to the associative law, the order of message aggregation and feature update can be interchanged. If the aggregation is performed first and then the update is calculated, the d_i addition of d_{in}-dimension and 1 multiplication of $d_{in} \to d_{out}$ are required for node i. According to the handshake theorem, the whole graph is a total of $2|E|$ times of additions and $|V|$ times of matrices Vector multiplication (or 1 time matrix-matrix multiplication); while updating and then aggregation, the d_i addition of d_{out}-dimension and d_{in} multiplication are required for node i, and the time complexity of the whole graph is $O(|E|)$. But in fact, since the neighbor nodes will be shared by multiple nodes, the matrix-matrix multiplication is used to calculate $H(k)W(k)$ in advance, and then index aggregation is performed, and the number of multiplications is also 1.

IV. EXPERIMENTS AND ANALYSIS

In this section, we will conduct extensive experiments on real-world data to evaluate the performance of the model. We employ several statistical, classical machine learning and deep learning methods to adapt the paper impact prediction problem, and further analyze the experimental results, while we also visually demonstrate the interpretability of our method.

A. DATASETS DESCRIPTIONS

The data used in this study came from two publicly available datasets, AMiner and APS. AMiner is an open platform for disciplines established by Tsinghua University, providing a collection of publicly available papers in the field of

computer science. The APS dataset is a publicly available dataset established by the American Physical Society, which exports physics related paper data from datasets on the APS platform. These datasets are recognized as widely used open datasets. They contain a large amount of information related to the paper, including detailed information such as authors, journals, keywords, references, etc., especially in the fields of computer science and physics. In addition, the dataset also includes citation records.

We have extracted a substantial volume of paper data spanning from 2010 to 2022 within the field of computer science from the Aminer platform. This dataset encompasses a total of 1,475,248 papers, 1,123,172 distinct authors, 4,116 unique journal sources, and 40,145 distinct paper keywords. In total, these elements culminate in 14,076,743 distinct heterogeneous relationships. Leveraging this experimental data, we have constructed a corresponding academic research network to serve as the foundation for our research. This network, grounded in the academic landscape of computer science, is employed to predict the citation counts for as-yet-unpublished articles.

B. EVALUATION METRICS

We selected three evaluation indicators for the prediction effect of the previous methods and the dynamic graph representation learning method we selected to facilitate our comparative analysis of the prediction accuracy of these methods. The three metrics are Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Pearson Correlation Coefficient (PCC). The Root Mean Square Error is the square root of the squared sum of the deviations between the observed value and the true value and the ratio of the number of observations m, and is used to measure the deviation between the observed value and the true value. Mean Absolute Deviation, also known as Mean Absolute Deviation, is the average of the absolute values of the deviations of all individual observations from the arithmetic mean. The mean absolute error can avoid the problem of mutual cancellation of errors, so it can accurately reflect the size of the actual forecast error. Pearson correlation coefficient is used to measure the correlation (linear correlation) between two variables X and Y, and its value is between -1 and 1. In the field of natural sciences, this coefficient is widely used to measure the degree of correlation between two variables.

We have selected three evaluation metrics to assess the predictive performance of both previous methods and the dynamic graph representation learning method we employed. This allows for a comparative analysis of the prediction accuracy across these approaches. The three metrics are Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Pearson Correlation Coefficient (PCC).

The Root Mean Square Error is derived from the square root of the sum of squared deviations between observed and true values, normalized by the number of observations m. It quantifies the deviation between observed and true values. Mean Absolute Error, also known as Mean Absolute Deviation, represents the average of the absolute deviations of all individual observations from their arithmetic mean. This metric is advantageous in avoiding the issue of error cancellation, providing an accurate reflection of the actual forecast error magnitude. The Pearson correlation coefficient assesses the linear correlation between two variables, X and Y, yielding a value between -1 and 1. Widely applied in the natural sciences, this coefficient serves to gauge the extent of correlation between two variables.

C. COMPARED METHODS

Firstly, for the previous traditional machine learning-based approaches, we employed two sub-methods, namely LightGBM (LGB) [14] and Linear Regression (LR), to train and predict the preceding dataset. The fundamental principles underlying methods based on traditional machine learning are as follows: (1) For each recently published article, an n-dimensional feature is derived for every article through the observed edges within the preceding academic research network. (2) Subsequent to feature extraction, each article is treated as an independent vector, and conventional machine learning techniques (LR, LGB) are employed for training and citation prediction.

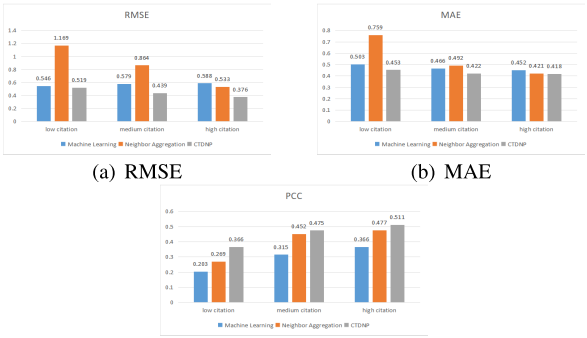
Subsequently, we carefully studied the commonly used neighbor aggregation methods in graph neural networks [45]. Capture the hidden features of the original node by learning the representation of neighbors. Due to the different characteristics of the network nodes studied in this article, we adopted neighbor aggregation through consistency calculation of weights to enhance the GNN network [30]. That is, we used the consistency score between neighbors to associate sampling, and allocated influence weights through relational attention mechanism. Here we have chosen three rules, represented as NA-P, NA-A, and NA-V, as follows: NA-P refers to treating the references of newly published papers as similar to other papers and calculating similarity using the same number of references between two papers; NA-A involves comparing newly published authors with other papers, using the same number of authors between two papers to calculate similarity; NA-V requires that articles published in the same journal as the new paper have at least one identical keyword, which is considered similar to the new paper. The similarity is calculated using the number of keywords.

Lastly, we employ DyRep for forecasting the citation count of a new paper. We sequence the papers chronologically and then utilize the paper sequence as input during the training process.

We execute the aforementioned three categories (resulting in six subdivisions) of traditional methods and the adaptive citation prediction graph neural network (ACPGNN) proposed by this project, totaling seven methods. We employ the three evaluation metrics described earlier to assess the performance of these seven methods. We set the learning rate to 0.0001. The model is trained on all datasets for 300 epochs

TABLE 1. Comparison of the results of the 7 methods.

Methods	K= 100			K = 200			K = 300		
	RMSE	MAE	PCC	RMSE	MAE	PCC	RMSE	MAE	PCC
LGB	2.1876	1.8234	0.0849	0.8939	0.7511	0.2015	0.5999	0.4663	0.3015
LR	1.6978	1.2469	0.1004	0.8369	0.7399	0.2527	0.5948	0.469	0.3204
NA-P	1.3851	1.6612	0.1212	1.1311	1.0333	0.2122	1.1442	0.878	0.2826
NA-A	2.1914	1.8767	0.0843	1.1977	1.1511	0.2204	0.5956	0.4718	0.3907
NA-V	2.2356	2.0428	0.0822	0.9534	0.8327	0.2098	0.5611	0.4284	0.4627
DyRep	2.1125	2.0128	0.0857	0.7215	0.7091	0.3074	0.4874	0.4591	0.4934
ACPGNN	2.4184	2.1233	0.0802	0.6311	0.6121	0.3558	0.4372	0.4199	0.5285



(a) RMSE (b) MAE

(c) PCC

FIGURE 10. Histogram of three methods under three citations.

and an early stopping strategy is implemented with a patience parameter set to 100. Considering the issue of overfitting, we set the dropout rate to 0.01. The model attaining the highest performance on the validation set is selected for testing. The outcomes are depicted in Table 1.

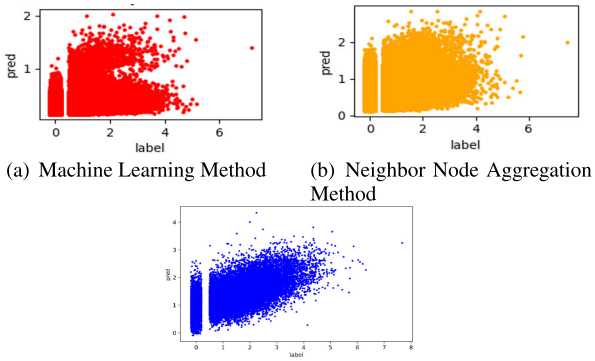
Among these, RMSE and MAE represent the root mean square error and mean absolute error, respectively. The smaller these values, the closer the predicted results are to the actual values. PCC stands for the Pearson correlation coefficient, where a value closer to 1 indicates a stronger correlation. As shown in the table above, both the RMSE and MAE values for our proposed ACPGNN method are smaller than the minimum values achieved by the five traditional methods mentioned earlier. Furthermore, the PCC coefficient is higher than the maximum values attained by the same traditional methods. Additionally, compared to DyRep, our method demonstrates faster convergence and improved outcomes. These results highlight the superior predictive performance of our proposed approach.

D. IN-DEPTH ANALYSIS OF EXPERIMENTS

The method results comparison experiment in above Section lists the numerical results of the traditional method and our ACPGNN method, but it is not intuitive enough. In order to analyze the results more carefully and in-depth, we did the following experiments to analyze the results.

1) FIRST ANALYSIS

First, we analyze the differences in the prediction effects of these three categories of methods under different citation counts. We selected the bottom 10%, the middle 10% and the top 10% of the articles ranked by citations five years after



(a) Machine Learning Method (b) Neighbor Node Aggregation Method

(c) Dynamic Graph Representation Learning Method

FIGURE 11. Actual Scatter Plot.

publication as the low-cited, medium-cited and high-cited articles collections. Calculate the average of the predicted citations of these three types of paper collections by the above three different models, and present them through a histogram, as shown in Figures 10. Among them, the first method in the figure refers to the method based on traditional machine learning (Machine Learning), the second method refers to the method based on Neighbor Aggregation, and the third method refers to the ACPGNN proposed in this paper. From the three figures listed above, we can intuitively see the accuracy of the predictions of these three types of models under different reference conditions. The errors between the predicted value and the true value of the three models when predicting the future citations of low-cited papers are larger than those of medium- and high-citations. This means that the prediction effects of the three models on papers with low citations are lower than those of medium citations and high citations. As the number of citations of a paper goes up, so does the predictive accuracy of the method. It is worth noting that the prediction effect of the proposed method based on dynamic graph representation learning is due to the previous two traditional methods in the case of low citation, medium citation and high citation. Once again, the superiority of this method in different situations is proved.

2) SECOND ANALYSIS

In terms of the correlation between the predicted value and the actual value, we use the Python plot drawing function to draw the results of different models in the form of scatter plots. We have plotted prediction-actual scatter plots for each of the three models. As shown in Figures 11 below, each point on the graph represents a prediction result, the horizontal axis is the actual value, and the vertical axis is the predicted value. Through these three scatter plots, we can intuitively see whether the correlation between the predicted value and the actual value is strong, so as to analyze the quality of the predicted results.

As can be seen from the above three figures, the dynamic graph representation learning method in this paper has a more obvious correlation between the predicted value and

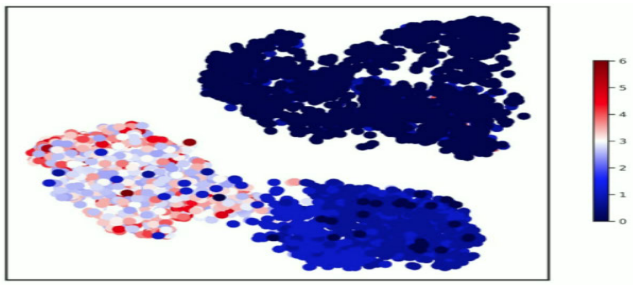


FIGURE 12. Feature distribution of paper nodes of three types of citations.

the actual citation volume compared with the two traditional methods. This also indirectly shows that the method of dynamic graph representation learning has better prediction effect than traditional methods.

3) VISUAL EXPERIMENT

To validate that the outcomes obtained by our model effectively encapsulate features pertinent to article citations, a visual analysis of the nodes post-training can be conducted. Following the completion of training, each point will possess a ''memory''. The memories from all models are saved upon training completion. Subsequently, a portion of the paper memories is selected for visualization on a two-dimensional plane. We categorize articles into three groups based on their citation rates: those in the bottom 10%, middle 10%, and top 10% by citations five years after publication. This categorization aids in making the memories of the three types of papers more distinct and discernible post-visualization. Subsequently, 1500 articles are sampled from each of these three categories, and their trained memories are visualized on a two-dimensional plane using t-SNE. The results are presented in Figure 12 below. In the figure above, nodes are color-coded based on their respective reference counts. Ranging from dark blue to dark red, this represents a gradient from low to high citations. Intuitively, the figure illustrates that paper nodes with differing citation counts (manifested in distinct colors) exhibit notably distinct feature distributions within the graph. Conversely, nodes with akin citation counts (sharing similar colors) tend to cluster together, indicating similarity in feature distributions. This confirms that the acquired memory through our training process effectively encapsulates features associated with the citation count of articles.

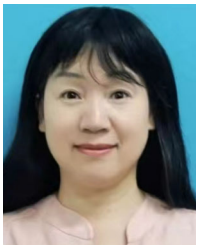
V. CONCLUSION

In this study, we formally introduce a novel task focusing on the influence of academic papers within the research academic network: the prediction of the impact a paper will have after its publication. To address this, we propose an adaptive graph neural network founded on representation learning, incorporating a multi-head attention mechanism. Building upon this, we present an improved solution for a challenging ''cold start'' time series problem, specifically, forecasting the future citation time series of a newly published paper when historical citation values are unavailable. This solution not only considers the intricate and heterogeneous interactions among articles and various nodes but also effectively retains the time series information within the network. Through experiments conducted on the Aminer and APS datasets, we demonstrate that the model surpasses all baseline methods, and we conduct a thorough analysis of the model's reliability and interpretability. Looking ahead, we aim to enhance the model's complexity to investigate more intricate interaction network models and explore further possibilities.

REFERENCES

[1] A. Abrishami and S. Aliakbary, ''Predicting citation counts based on deep neural network learning techniques,'' *J. Informetrics*, vol. 13, no. 2, pp. 485–499, May 2019.
[2] T. Amjad, A. Daud, S. Khan, R. A. Abbasi, and F. Imran, ''Prediction of rising stars from Pakistani research communities,'' in *Proc. 14th Int. Conf. Emerg. Technol. (ICET)*, Nov. 2018, pp. 1–6.
[3] C. Castillo, D. Donato, and A. Gionis, ''Estimating number of citations using author reputation,'' in *Proc. Conf. SPIRE*, Santiago, Chile, 2007, pp. 107–117.
[4] X. Chang, X. Liu, J. Wen, S. Li, Y. Fang, L. Song, and Y. Qi, ''Continuous time dynamic graph learning via neural interaction processes,'' presented at the 29th ACMCIKM, Oct. 2020. [Online]. Available: https://www.cikm2020.org/accepted-papers/accepted-research-papers/
[5] J. Chen, A. Pareja, G. Domeniconi, T. Ma, T. Suzumura, H. Kanezashi, and T. Kaler. (Feb. 13, 2020). *Patent US11537852B2*. [Online]. Available: https://patents.google.com/patent/US11537852/en
[6] Z. Chen, L. Chen, S. Villar, and J. Bruna, ''Can graph neural networks count substructures,'' in *Proc. Conf. NIPS*, Vancouver, BC, Canada, 2020, pp. 10383–10395.
[7] Z. Chen, S. Villar, L. Chen, and J. Bruna, ''On the equivalence between graph isomorphism testing and function approximation with GNNs,'' in *Proc. Conf. NIPS*, Vancouver, BC, Canada, 2019, pp. 15894–15902.
[8] Y. Dong, N. V. Chawla, and A. Swami, ''metapath2vec: Scalable representation learning for heterogeneous networks,'' in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2017, pp. 135–144.
[9] Y. Dong, R. A. Johnson, and N. V. Chawla, ''Can scientific impact be predicted?'' *IEEE Trans. Big Data*, vol. 2, no. 1, pp. 18–30, Mar. 2016.
[10] J. Gan, R. Hu, Y. Mo, Z. Kang, L. Peng, Y. Zhu, and X. Zhu, ''Multigraph fusion for dynamic graph convolutional network,'' *IEEE Trans. Neural Netw. Learn. Syst.*, early access, May 16, 2022, doi: 10.1109/TNNLS.2022.3172588.
[11] W. L. Hamilton, Z. Ying, and J. Leskovec, ''Inductive representation learning on large graphs,'' in *Proc. NIPS*, Long Beach, CA, USA, 2017, pp. 1024–1034.
[12] Z. Hu, Y. Dong, K. Wang, and Y. Sun, ''Heterogeneous graph transformer,'' in *Proc. Web Conf.*, Apr. 2020, pp. 2704–2710.
[13] S. Jiang, B. Koch, and Y. Sun, ''HINTS: Citation time series prediction for new publications via dynamic heterogeneous information network embedding,'' in *Proc. Web Conf.*, New York, NY, USA, Apr. 2021, pp. 3158–3167.
[14] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, ''LightGBM: A highly efficient gradient boosting decision tree,'' in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, vol. 30, Long Beach, CA, USA, 2017, pp. 3146–3154.
[15] K. C. Lai, S. H. Lau, J. Wong, and J. Prakash, ''Assessment on author contribution in publication: A tool for decision-makers,'' *Proc. Comput. Sci.*, vol. 91, pp. 402–411, Jan. 2016.
[16] T. Loughran and B. Mcdonald, ''Textual analysis in accounting and finance: A survey,'' *J. Accounting Res.*, vol. 54, no. 4, pp. 1187–1230, Sep. 2016.
[17] A. Loukas, ''What graph neural networks cannot learn: Depth vs width,'' presented at the ICLR, Jun. 2019.
[18] F. Manessi, A. Rozza, and M. Manzo, ''Dynamic graph convolutional networks,'' *Pattern Recognit.*, vol. 97, Jan. 2020, Art. no. 107000.

[19] A. Mazloumian, "Predicting scholars scientific impact," *PLoS ONE*, vol. 7, no. 11, pp. 149–246, Nov. 2012.
[20] A. Pareja, G. Domeniconi, J. Chen, T. Ma, T. Suzumura, H. Kanezashi, T. Kaler, T. Schardl, and C. Leiserson, "EvolveGCN: Evolving graph convolutional networks for dynamic graphs," in *Proc. AAAI Conf. Artif. Intell.*, Apr. 2020, vol. 34, no. 4, pp. 5363–5370.
[21] E. Rossi, B. Chamberlain, F. Frasca, D. Eynard, F. Monti, and M. Bronstein, "Temporal graph networks for deep learning on dynamic graphs," prensented at ICML, Jun. 2020.
[22] X. Ruan, Y. Zhu, J. Li, and Y. Cheng, "Predicting the citation counts of individual papers via a BP neural network," *J. Informetrics*, vol. 14, no. 3, Aug. 2020, Art. no. 101039.
[23] A. Sankar, Y. Wu, L. Gou, W. Zhang, and H. Yang, "DySAT: Deep neural representation learning on dynamic graphs via self-attention networks," in *Proc. 13th Int. Conf. Web Search Data Mining*, Houston, TX, USA, 2020, Jan. 2020, pp. 519–527.
[24] R. Van Belle, C. Van Damme, H. Tytgat, and J. De Weerdt, "Inductive graph representation learning for fraud detection," *Expert Syst. Appl.*, vol. 193, May 2022, Art. no. 116463.
[25] D. Wang, C. Song, and A.-L. Barabási, "Quantifying long-term scientific impact," *Science*, vol. 342, no. 6154, pp. 127–132, Oct. 2013.
[26] S. Ma, J.-W. Liu, X. Zuo, and W.-M. Li, "Heterogeneous graph gated attention network," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2021, pp. 1–6.
[27] D. Xu, C. Ruan, E. Korpeoglu, S. Kumar, and K. Achan, "Inductive representation learning on temporal graphs," in *Proc. Int. Conf. Learn. Represent.*, 2020, pp. 1–19.
[28] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?" *Presented at arXiv*. [Online]. Avaliable: https://arxiv.org/abs/1810.00826.
[29] R. Yan, J. Tang, X. Liu, D. Shan, and X. Li, "Citation count prediction: Learning to estimate future citations for literature," in *Proc. 20th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2011, pp. 1247–1252.
[30] L. Yang, Z. Liu, Y. Dou, J. Ma, and P. S. Yu, "ConsisRec: Enhancing GNN for social recommendation via consistent neighbor aggregation," in *Proc. 44th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jul. 2021, pp. 2141–2145.
[31] S. Yuan, J. Tang, Y. Hang, Y. Wang, and T. Xiao, "Modeling and predicting citation count via recurrent neural network with long short-term memory," in *Proc. 28th Int. Conf. AAAI. Artifi. Intell.*, Jul. 2014, pp. 2141–2145.
[32] S. Zhang, H. Tong, J. Xu, and R. Maciejewski, "Graph convolutional networks: A comprehensive review," *Comput. Social Netw.*, vol. 6, no. 1, p. 12, Dec. 2019.
[33] K. McKeown, H. Daume III, and S. Chaturvedi, "Predicting the impact of scientific concepts using fulltext features," *J. Assoc. Inf. Sci. Technol.*, vol. 67, no. 11, pp. 2684–2696, Jan. 2016.
[34] X. Bai, H. Liu, F. Zhang, Z. Ning, X. Kong, I. Lee, and F. Xia, "An overview on evaluating and predicting scholarly article impact," *Information*, vol. 8, no. 3, p. 73, Jun. 2017.
[35] A. P. Akella, H. Alhoori, P. R. Kondamudi, C. Freeman, and H. Zhou, "Early indicators of scientific impact: Predicting citations with altmetrics," *J. Informetrics*, vol. 15, no. 2, May 2021, Art. no. 101128.
[36] A. Clauset, D. B. Larremore, and R. Sinatra, "Data-driven predictions in the science of science," *Science*, vol. 355, no. 6324, pp. 477–480, Feb. 2017.
[37] Y. Yin, L.-X. Ji, J.-P. Zhang, and Y.-L. Pei, "DHNE: Network representation learning method for dynamic heterogeneous networks," *IEEE Access*, vol. 7, pp. 134782–134792, 2019.
[38] Y. Xie, B. Yu, S. Lv, C. Zhang, G. Wang, and M. Gong, "A survey on heterogeneous network representation learning," *Pattern Recognit.*, vol. 116, Aug. 2021, Art. no. 107936.
[39] L. R. Medsker and L. Jain, "Recurrent neural networks," *Des. Appl.*, vol. 5, pp. 64–67, Jan. 2001.
[40] A. Gillioz, J. Casas, E. Mugellini, and O. A. Khaled, "Overview of the transformer-based models for NLP tasks," in *Proc. 15th Conf. Comput. Sci. Inf. Syst. (FedCSIS)*, Sep. 2020, pp. 179–183.
[41] J. Zhou, G. Cui, and S. Hu, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, p. 5781, Dec. 2020.
[42] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021.
[43] I. Beltagy, K. Lo, and A. Cohan, "SciBERT: A pretrained language model for scientific text," 2019, *arXiv:1903.10676*.
[44] R. Trivedi et al., "DyRep: Learning representations over dynamic graphs," presented at the ICIR. [Online]. Avaliable: https://par.nsf.gov/biblio/10099025.
[45] G. Corso, L. Cavalleri, D. Beaini, P. Liò, and P. Veličković, "Principal neighbourhood aggregation for graph nets," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2020, pp. 13260–13271.
[46] S. Saha, S. Saint, and D. A. Christakis, "Impact factor: A valid measure of journal quality?" *J. Med. Library Assoc.*, vol. 91, no. 1, p. 42, 2003.
[47] L. Bornmann, "What is societal impact of research and how can it be assessed? A literature survey," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 64, no. 2, pp. 217–233, Feb. 2013.
[48] D. W. Aksnes, "Citation rates and perceptions of scientific contribution," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 57, no. 2, pp. 169–185, Jan. 2006.
[49] L. Waltman, "A review of the literature on citation impact indicators," *J. Informetrics*, vol. 10, no. 2, pp. 365–391, May 2016.



MINGHONG YANG received the B.Eng. degree in software engineering from Hunan University, in 2017. She is currently an Associate Professor of computer science with the School of Economics and Management, Hunan Applied Technology University. She has published nearly 20 articles and three textbooks. Her research interests include internet and data base.

ZHENHUA YANG received the B.Eng. degree in software engineering from Central South University, in 2017. He is currently an Associate Professor of computer science with the School of Information Engineering, Hunan Applied Technology University, specializing in teaching computer science courses. He has published 13 articles, four textbooks, and two patents in computer science.