



Role-Based Network Embedding via Quantum Walk with Weighted Features Fusion

Mingqiang Zhou*, Mengjiao Li, Zhiyuan Qian and Kunpeng Li

College of Computer Science, Chongqing University, Chongqing, 400044, China

*Corresponding Author: Mingqiang Zhou. Email: zmq@cqu.edu.cn

Received: 24 December 2022; Accepted: 20 February 2023; Published: 30 August 2023

Abstract: Role-based network embedding aims to embed role-similar nodes into a similar embedding space, which is widely used in graph mining tasks such as role classification and detection. Roles are sets of nodes in graph networks with similar structural patterns and functions. However, the role-similar nodes may be far away or even disconnected from each other. Meanwhile, the neighborhood node features and noise also affect the result of the role-based network embedding, which are also challenges of current network embedding work. In this paper, we propose a Role-based network Embedding via Quantum walk with weighted Features fusion (REQF), which simultaneously considers the influence of global and local role information, node features, and noise. Firstly, we capture the global role information of nodes via quantum walk based on its superposition property which emphasizes the local role information via biased quantum walk. Secondly, we utilize the quantum walk weighted characteristic function to extract and fuse features of nodes and their neighborhood by different distributions which contain role information implicitly. Finally, we leverage the Variational Auto-Encoder (VAE) to reduce the effect of noise. We conduct extensive experiments on seven real-world datasets, and the results show that REQF is more effective at capturing role information in the network, which outperforms the best baseline by up to 14.6% in role classification, and 23% in role detection on average.

Keywords: Role-based network embedding; quantum walk; quantum walk weighted characteristic function; complex networks

1 Introduction

Roles are sets of nodes with similar structural patterns and functions. Role-based network embedding aims to project role-similar nodes into a compact low-dimensional vector space. It is widely used in various downstream tasks such as role classification, etc. The first appearance of the concept of roles is in sociology [1] for mining potential social relationships, and it is gradually applied in complex networks like traffic network congestion [2]. By measuring the structural similarity [3], role-similar nodes can be found in distant locations, even in different subgraphs of disconnected networks.



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Most previous studies mainly focus on the local context such as node structure, while rarely measuring their relevance globally, especially when nodes are far away so that ignoring their similarity in the global context. For example, Role-based graph embedding (Role2vec) [4] uses the classical random walk or its variants to extract structural information and aggregate role-based embedding. It selectively extracts the structural information from low-order neighborhoods, which ignores global information. However, the situation that role-similar nodes are far away from each other in the graph while showing high relevance in terms of node structural similarity may exist. As shown in Fig. 1, in the world-air-traffic network, landmarks with the same color indicate the same hub role, and landmarks with the same role may be far away. For example, two distant landmarks in Beijing and Washington D.C. airports. They are both capitals of their country and have the same status and role, which is reflected by global information. More recently, Role Embedding via Discrete-time quantum walk (RED) [5] considers the global role information and highlights the global role relevance of nodes, but ignores the effect of node features.

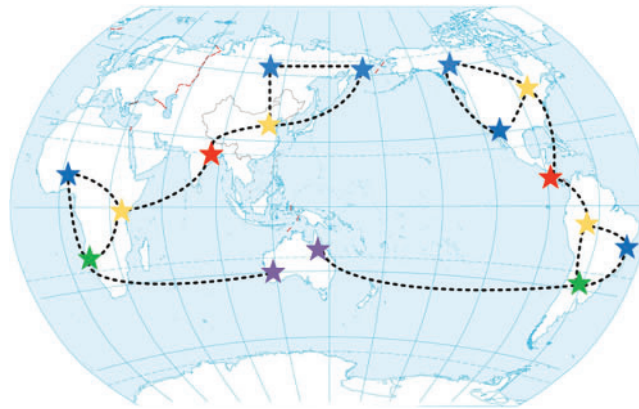


Figure 1: World-air-traffic network

In addition to the features of the node itself, its neighborhood nodes will also indeed affect the embedding of role-based networks. For example, a music teacher in the social network is interested in both music and photography, the role of which is more influenced by music than photography and its neighborhoods may be more music lovers than photography lovers. Most of the current methods based on features focus on linear aggregation [6] or manual [7] and recursive [8] feature extraction, which cannot represent the diverse information of neighborhood features. Characteristic functions on graphs: birds of a Feather (FEATHER) [9] considers this problem and proposes a characteristic function to make meaningful combinations of multiple node and neighborhood features.

Nevertheless, the extraction or aggregation of multiple features is extremely sensitive to noise that may interfere with the final embedding. VAE [10] assumes that the input follows an ideal data distribution by adding constraints to the encoder. Therefore, even if the noise disturbs networks to different degrees, a corresponding output can also be generated by using the coding within the disturbed range. Role-based network Embedding via Structural features reconstruction with Degree-regularized constraint (RESN) [11] uses VAE to reduce the effect of noise on the role-based network. It also proves the adverse effect of noise on embedding experimentally.

Above all, there are mainly three challenges in current role-based network embedding. 1) Role-similar nodes may be far away so we cannot capture global role information. 2) The varied distributions of nodes and their features are always ignored or simplified. 3) The noise may have an adverse influence

on the embedding. To address these challenges, we propose a method called Role-based network Embedding via Quantum walk with weighted Features fusion (REQF).

Firstly, we superimpose all walk paths based on the superposition of the quantum walk to capture the global role information by multi-step evolution sequences and use the biased quantum walk to emphasize the local structure and learn role closeness. Secondly, we design a quantum walk weighted characteristic function to fuse features of nodes and their neighborhoods by different distributions. Finally, we use VAE to reduce the effects of noise and generate role-based network embedding. Extensive experiments on real-world networks demonstrate the effectiveness and stability of the proposed REQF.

Our main contributions are listed as follows:

- We propose a novel method REQF, which considers the influence of global and local role information, node features, and the effect of noise simultaneously for generating role-based network embedding.
- We utilize quantum walk to capture role information from global and local perspectives for solving the node distance problem, and we construct a quantum walk weighted characteristic function, that uses the quantum walk as probability weights of the characteristic function, to take node features into consideration for feature fusion.
- We use VAE to reduce the effect of noise in the network and get the optimal role-based network embedding.

The rest of the paper is organized as follows: In [Section 2](#), we briefly review the related work. In [Section 3](#), we introduce the proposed REQF in detail. In [Section 4](#), we report experimental results on real-world networks. Finally, in [Section 5](#), we conclude the paper.

2 Related Work

In graph networks, roles are sets of nodes with similar structural patterns and functions. It is first defined as classes of structurally equivalent nodes, and the structural equivalence is measured by the structural similarity of nodes [12]. Role-based network embedding aims to embed role-similar nodes into a compact embedding space.

Role-based networks are more concerned with the structural similarity of nodes and are independent of distance. To represent role information, Role-based graph embedding (Role2vec) [4] uses an attribute-based random walk to capture the role information of nodes. Learning structural Node Embeddings on Graphs via diffusion Wavelets (GraphWave) [13] learns structural node embedding by diffusion wavelets. Fast structural node embedding via role identification (RiWalk) [14] uses random walk to approximate graph kernels for role identification. Graph neural network with Local Structural Patterns (GraLSP) [15] incorporates local structure into sets of neighborhoods by anonymous walk. These methods based on the random walk are selective in obtaining the structural information and do not consider the global role information of nodes, which results in information loss. Thus, role-based network embedding needs a technique to overcome the limitations of distance and global information loss.

Quantum walk [16] is the quantization of walk based on the principle of random motion of particles. The generation of a walking path is determined by the walk position indicating the current location of the node and the walk state indicating the transfer process of walking to the next node. The classical random walk only generates one path for each walk, and no superposition is allowed. In contrast, the quantum walk has the property of coherent superposition that achieve randomness due

to its superposition [17]. It can be described as the superposition of all walk states and it walks to all paths simultaneously. As shown in Fig. 2, quantum walk allows all nodes to interact with the whole network simultaneously.

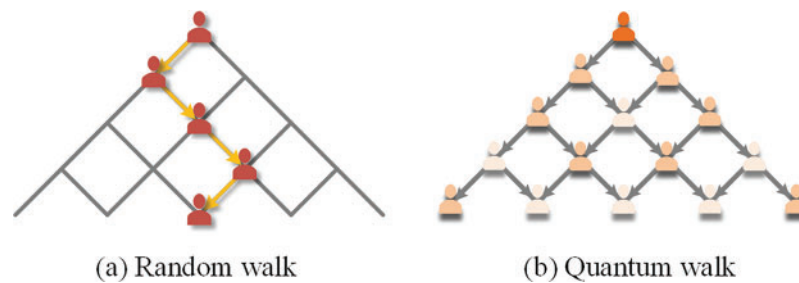


Figure 2: Illustration of comparison between random walk and quantum walk

Presently, the quantum walk has considerable research prospects in the field of role embedding. Most quantum walk of graphs is implemented by calculating the superposition of probability distributions. Fast Quantum Walk Kernel (FQWK) [18] uses quantum walk as the convolution kernel to measure the similarity of neighborhood substructures between node pairs. However, the method does not consider the structural influence of nodes from the global perspective. RED [5] uses quantum walk to capture the role information of nodes from global and local perspectives, but it ignores the feature information of nodes. It also proposes that role-similar nodes have similar evolution sequences, and the discrete-time quantum walk (DTQW) models are highly correlated with the structural information of nodes. Motivated by it, we introduce the DTQW model to learn the global and local structural patterns of nodes.

On the other side, most of the current feature-based network embedding methods generate embedding by simple extraction or linear aggregation of features. For example, Role eXtraction (RoIX) [6] aggregates node features based on the matrix decomposition. Graph Auto-encoder guided by Structural information (GAS) [7] uses the manually extracted features as guidance to train the graph auto-encoder. Role Discovery-guided network embedding based on Autoencoder and Attention mechanism (RDAA) [8] uses recursive feature extraction and incorporates attention mechanisms to guide role discovery in auto-encoders. These methods view all neighborhood features of nodes as uniform distributions, ignoring the fact that different distributions of neighborhood features may have different effects on the central node. Thus, FEATHER [9] uses a complex-valued characteristic function to make meaningful combinations of multiple features and proves that characteristic functions can obtain effective neighborhood features at discrete evaluation points. But its probability weights are simply described by the power of the adjacency matrix, which cannot represent the correlated information well. Therefore, we try to use the quantum walk to appropriately represent the probability weights of the characteristic function that can implicitly highlight the role information and fuse various features in role-based network embedding.

Furthermore, noise affects the generation of high-quality network embedding. Due to the over-reliance on neighborhood information, the feature-based methods are sensitive to noise, causing side effects on the generation of embedding. RESD [11] uses VAE to reduce the effect of noise after extracting node features and preserve structural identity with degree regularization. Adversarial Variational Autoencoder Embedding (AVAE) [19] utilizes VAE to avoid noise from obstructing the convergence of generated data. Dual Adversarial Variational Embedding (DAVE) [20] proposes a model based on VAE for recommendations providing personalized noise reduction for different users

and items. Purified Graph Generation (PGG) [21] also uses a variant of VAE to filter noise during graph generation. These methods indicate that the influence of noise can be reduced by using the approach of an encoder. Therefore, we attempt to use VAE to obtain a more accurate role-based embedding.

In this paper, we try to overcome the limitations of role-based embedding in terms of node distance and features via quantum walk and weighted characteristic function to capture the global role information and fuse node features, and we use VAE to effectively reduce the effect of noise, and finally obtain a more compact role-based network embedding.

3 The Proposed Method

In this section, we introduce the notation and the proposed method REQF. We first start with an overview and then present the detailed designs, including role information representation, feature information fusion, and noise effect reduction. Finally, we analyze the computational complexity of REQF.

For clear clarification, the symbols and their definitions are listed in [Table 1](#).

Table 1: Symbols and definitions

Symbols	Definitions
$G = (V, E)$	The graph network, V is a set of nodes, E is a set of edges.
$ \psi\rangle$	The state of quantum walk
$ v\rangle$	The superposition of quantum states at the node v
U	The evolution operator of quantum walk
α_v	The complex-valued probability amplitude of node v
p'_v	The probability of quantum walk being at a node v and evolution time t
\mathbf{M}_t	The probability distribution of all nodes at the t^{th} evolution
$R(v_i), R(v_j)$	The role of nodes v_i, v_j
$\mathbf{q}(v_i), \mathbf{q}(v_j)$	The evolution sequence of nodes v_i, v_j
$f(\cdot)$	The summation function by row
$S(\cdot)$	The similarity measure function
$\mathcal{N}(v)$	The neighborhood nodes of node v
\mathbf{A}	The adjacency matrix
$d(v_i)$	The degree of a node v_i
p^{biased}	The biased probability
t_g, t_l	The evolution times of the global part and local part
$\text{sort}(\cdot)$	The proximity ranking function by row
\mathbf{x}_v	The feature vector of node v
θ	The evaluate points
$\mathbf{Z}_{\text{global}}, \mathbf{Z}_{\text{local}}, \mathbf{Z}_{\text{feature}}, \mathbf{Z}$	The node representation of the global role, local role, feature, and VAE
\mathbf{Z}	Role-based network embedding

3.1 Notation and Problem Definition

Definition 1 Quantum Walk: Given an undirected and unweighted graph $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_n\}$ is the set of n nodes. Quantum walk consists of **the state of quantum walk** $|\psi\rangle$ and **the state evolution** $|\psi'\rangle \xrightarrow{U} |\psi\rangle$. The state of quantum walk is described by a coordinate wave function $|\psi\rangle = \sum_v \alpha_v |v\rangle$, where $|v\rangle$ denotes the superposition of quantum states at the node v , which is an n -dimensional column vector, and α_v is the corresponding complex-valued probability amplitude. The state evolution process is usually described by $|\psi'\rangle \xrightarrow{U} |\psi\rangle$, where the evolution operator U is used to represent the position and direction of the quantum walk.

We define the probability of quantum walk being at a node v and evolution time t as $p'_v = \sum_v |\alpha_v(t)|^2$, where $|\alpha_v(t)|^2$ is the complex-valued probability amplitude. The result of the quantum walk \mathbf{M}_t is a $n \times n$ matrix, representing the probability distribution of all nodes at the t^{th} evolution, that is $\mathbf{M}_t = \{p_1^t, p_2^t, \dots, p_n^t\}^T$, $t = 1, 2, \dots, n$. The state evolution is defined as:

$$\mathbf{M}_t = \mathbf{U}\mathbf{M}_{t-1} \quad (1)$$

Definition 2 Role Similarity: Roles are sets of nodes with similar structural patterns and functions in graph networks. The role similarity of nodes is defined by the evolution sequence that is formed by summing the probability distribution matrix of the quantum walk.

$$\mathbf{q} = \{f(\mathbf{M}_1), f(\mathbf{M}_2), \dots, f(\mathbf{M}_t)\} \quad (2)$$

$$S(R(v_i), R(v_j)) \rightarrow 1 \Leftrightarrow S(\mathbf{q}(v_i), \mathbf{q}(v_j)) \rightarrow 1 \quad (3)$$

where \mathbf{q} is the evolution sequence of quantum walks, $R(v_i), R(v_j)$ represent the role of node v_i, v_j , $f(\cdot)$ is the summation function by row, and $S(\cdot)$ is the similarity measure function. Roles will get more similar if the value of $S(\cdot)$ reaches closer to 1.

Definition 3 Role-based Network Embedding: For a graph G , the neighborhood nodes of a node v are defined as $\mathcal{N}(v) = \{u \in V | (u, v) \in E\}$. Role-based network embedding $\mathbf{Z} \in \mathbb{R}^{n \times d}$ aims to generate the embedding that maximum reserves the role information contained in it, where d is the embedding dimension. $\mathbf{z}_v \in \mathbb{R}^d$ is the map for each node v into a low-dimensional vector space that contains role information.

$$\mathbf{Z} = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n\}^T \quad (4)$$

Role-similar nodes would be embedded into similar representations.

3.2 Overview of REQF

The illustration of our framework REQF is shown in Fig. 3. REQF consists of three parts: quantum walk, feature fusion, and VAE.

First of all, we utilize quantum walk to capture the global role information $\mathbf{z}_{\text{global}}$ (red arrow flow) and local role proximity $\mathbf{z}_{\text{local}}$ (blue arrow flow) of nodes by the evolution sequence \mathbf{q} . The block graphs represent the process of the quantum walk evolution as an operator U . The dotted arrows around nodes indicate that all nodes are involved in the evolution at the same time, and the thick dotted arrows indicate the biased process of the quantum walk with probability p_{biased} to make the node obtain a higher initial probability. Then, we use the quantum walk weighted characteristic function for feature fusion (yellow arrow flow), which uses \mathbf{M}_t to weight characteristic function and is calculated at evaluation points θ to obtain the node neighborhood feature distribution $\mathbf{z}_{\text{feature}}$ with structural information. Finally, we leverage VAE to reduce the effect of noise \mathbf{z} and generate the embedding \mathbf{Z} .

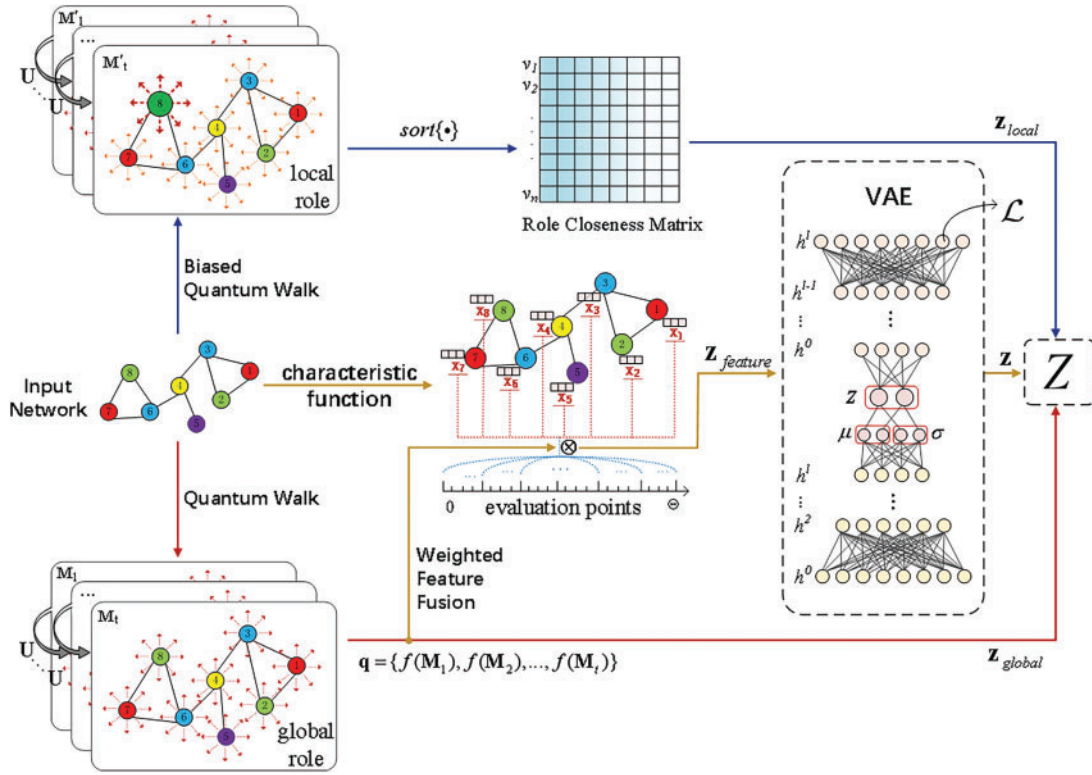


Figure 3: Illustration of our proposed framework REQF

3.3 Role Information Representation

REQF captures role information via quantum walk. It captures global role information and local role information by controlling the initial state probability. The global role information is the role relevance between nodes from the whole network. The local role information is the role proximity between nodes by the local structure of nodes, and it emphasizes more on the correlation between nodes and their neighborhoods.

3.3.1 Global Role Information

When all nodes have the same initial state probability, they evolve simultaneously over the entire network, and REQF simulates the global evolution by the probability distribution matrix M_t of quantum walk to represent the global role relevance of nodes and generate the global role representation z_{global} . The formulated Grover operator [22] is shown in Eq. (5).

$$M_t(v_i) = \left(\frac{2}{d(v_i)} - 1 \right) M_{t-1}(v_i) + \frac{2}{d(v_i)} \sum_{\forall v_j \in V, v_j \neq v_i} M_{t-1}(v_j) \quad (5)$$

$$M_0(v_i) = \alpha_{v_i}(0) = \frac{A}{\sqrt{n \cdot d(v_i)}} \quad (6)$$

$$z_{global} = \{f(M_1), f(M_2), \dots, f(M_{t_g})\} \quad (7)$$

where $d(v_i)$ is the degree of a node v_i , \mathbf{A} is the adjacency matrix, $f(\cdot)$ is the summation function by row, and t_g is the evolution times of quantum walk.

3.3.2 Local Role Information

If the initial state probability of a particular node is higher, the evolution of the quantum walk should be biased to emphasize the similarity of the local structure of the node [23], and the probability distribution obtained can be viewed as the correlation between the node and the neighborhood. Thus, REQF use emphasizes the local structure of the node by biased quantum walk \mathbf{M}_t to learn the role similarity of the node and generate a local role representation \mathbf{z}_{local} .

$$\mathbf{M}'_0(v_i) = \alpha_{v_i}(0) = \frac{\mathbf{A} \cdot p_{biased}}{\sqrt{n \cdot d(v_i)}} \quad (8)$$

$$\mathbf{z}_{local} = \text{sort}(f(\mathbf{M}'_1), f(\mathbf{M}'_2), \dots, f(\mathbf{M}'_{t_l})) \quad (9)$$

where p_{biased} is the biased probability, t_l is the evolution times of biased quantum walk, and $\text{sort}(\cdot)$ is the proximity ranking function by row to find other nodes with high role similarity to the central node.

3.4 Feature Information Fusion

Different distributions of neighborhood features may have different effects on nodes in real networks. FEATHER has theoretically demonstrated that graphs with the same structure have the same Characteristic Functions (CF). Therefore, considering the varied distributions of node neighborhood features, we use CF on the graph to describe the neighborhood feature distribution of the nodes, instead of fusing feature information in a simple linear aggregation way. The probability weights are defined by quantum walk to weight the CF which is to obtain feature information that highlights the structural property of the nodes.

According to the definition of the CF, we define the probability weights as the probability value of the quantum walk distribution matrix \mathbf{M}_t between the source node v_i and target node v_j . REQF uses the quantum walk weighted CF to fuse features and obtain a better and richer node representation $\mathbf{z}_{feature}$.

$$\mathbf{z}_{feature} = \sum_{v_i, v_j \in V} \mathbf{M}_t(v_i, v_j) [\cos(\theta \mathbf{x}_{v_j}) | \sin(\theta \mathbf{x}_{v_j})] \quad (10)$$

where \mathbf{x}_{v_j} is the feature vector of the target node v_j , θ is the value of evaluation points.

3.5 Noise Reduction and Embedding Generation

In reality, a lot of noise exists in extracting the role and feature information of nodes, and some node features unrelated to roles may interfere with the final role-based network embedding. To reduce the effect of noise, we use VAE to encode node role and feature information into the node embedding representation.

We use a Multi-Layer Perceptron (MLP) as our decoder and define it as follows:

$$\mathbf{h}'_l = \tanh(\mathbf{W}^l \mathbf{h}'_{l-1} + \mathbf{b}^l), l = 1, 2, \dots, L \quad (11)$$

where \mathbf{W}^l and \mathbf{b}^l represent the weight vector and bias vector of the l^{th} layer, L is the number of hidden layers in the encoder, and $\tanh(\cdot)$ is the hyperbolic tangent activation function. The vector $\mathbf{z}_{feature}(v_i)$ contains the role and feature information of node v_i , $\mathbf{h}^0_{v_i} = \mathbf{z}_{feature}(v_i)$ is the input of the encoder.

The node embedding representation \mathbf{z}_{v_i} of node v_i follows a Gaussian distribution, that is, $\mathbb{P}(\mathbf{z}_{v_i}) \sim \mathcal{N}(\mu_{v_i}, \text{diag}(\sigma_{v_i}^2))$. It can be obtained through a reparameterization trick [11] for learning the parameter μ_{v_i} and σ_{v_i} .

$$\boldsymbol{\mu}_{v_i} = \mathbf{W}^\mu \mathbf{h}_{v_i}^L + \mathbf{b}^\mu \quad (12)$$

$$\log \boldsymbol{\sigma}_{v_i} = \mathbf{W}^\sigma \mathbf{h}_{v_i}^L + \mathbf{b}^\sigma \quad (13)$$

$$\mathbf{z}_{v_i} = \boldsymbol{\mu}_{v_i} + \boldsymbol{\sigma}_{v_i} \odot \boldsymbol{\iota} \quad (14)$$

where \mathbf{W}^μ , \mathbf{W}^σ , \mathbf{b}^μ and \mathbf{b}^σ are respectively the weight and bias of the last layer of the encoder. \odot is the element-wise product, and $\boldsymbol{\iota}$ follows a standard Gaussian distribution, $\boldsymbol{\iota} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.

Equally, we use an MLP as the decoder:

$$\hat{\mathbf{h}}_i^l = \tanh(\hat{\mathbf{W}}^l \hat{\mathbf{h}}_i^{l-1} + \hat{\mathbf{b}}^l), l = 1, 2, \dots, L \quad (15)$$

where $\hat{\mathbf{W}}^l$ and $\hat{\mathbf{b}}^l$ are the weight and bias vector of the l^{th} hidden layer in the decoder. $\hat{\mathbf{h}}_i^0 = \mathbf{z}_{v_i}$ is the input of the decoder and $\hat{\mathbf{h}}_i^L$ is the reconstructed vector.

Finally, we define the loss function and the final embedding is obtained as follows:

$$\mathcal{L} = \left\| \mathbf{h}^0 - \hat{\mathbf{h}}^L \right\|_2^2 = \sum_{i=1}^n \left\| \mathbf{h}_{v_i}^0 - \hat{\mathbf{h}}_{v_i}^L \right\|_2^2 \quad (16)$$

$$\mathbf{Z} = \{\mathbf{z}_{\text{global}}, \mathbf{z}_{\text{local}}, \mathbf{z}\} \quad (17)$$

The whole process of REQF is described in Algorithm 1.

Algorithm 1: Process of REQF

Input: A network $G = (V, E)$,

a set of node feature vectors $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$,

quantum walk evolution times t_g and biased evolution times t_l ;

output: Role-based node embedding \mathbf{Z}

- 1: //Capture node role information
 - 2: Initialize the probability distribution matrix \mathbf{M}_0 by Eq. (6);
 - 3: **for** each evolution time t_g **do**
 - 4: Compute the global probability distribution matrix \mathbf{M}_l by Eq. (5);
 - 5: Obtain global role representation $\mathbf{z}_{\text{global}}$ by Eq. (7);
 - 6: **end for**
 - 7: Initialize the probability distribution matrix \mathbf{M}'_0 by Eq. (8);
 - 8: **for** each biased evolution time t_l **do**
 - 9: Compute the local probability distribution matrix \mathbf{M}'_l by Eq. (5);
 - 10: Obtain local role representation $\mathbf{z}_{\text{local}}$ by Eq. (9);
 - 11: **end for**
 - 12: //Fuse node feature information
 - 13: Obtain feature-fused representation $\mathbf{z}_{\text{feature}}$ by Eq. (10);
 - 14: //Variational auto-encoder
 - 15: **for** each epoch **do**
 - 16: Encode node information \mathbf{z} by Eq. (14);
-

(Continued)

Algorithm 1 (continued)

```

17:   end for
18:   //Get embedding
19:   return embedding  $\mathbf{Z} = \{\mathbf{z}_{\text{global}}, \mathbf{z}_{\text{local}}, \mathbf{z}\}$ 

```

3.6 Computational Complexity Analysis

The computational complexity of the proposed REQF mainly depends on the quantum walk and VAE, which are highly related to the number of nodes and feature dimensions.

Given a network $G = (V, E)$, V is the set of nodes. REQF captures the global (lines 2–6) and local (lines 7–11) role information of nodes through the quantum walk, which takes $O(t_g + t_l \cdot |V|)$ time. In the quantum walk module, we simulate the quantum walk on classical computers, and the complexity is $O(|V|^3)$. The process of feature fusion (line 13) takes $O(mt_g \cdot |V|)$, and m denote the dimension of the node feature. In the VAE module, the computational complexity is $O(r^2d)$, and r denotes the dimension of the input features, d denotes the output dimension. Therefore, the whole computational complexity of REQF is $O(t_l \cdot |V|^4 + r^2d)$.

The computational complexity of the REQF algorithm here is a bit high because the quantum walk is simulated with a conventional computer. But the complexity of the actual quantum walk is only $O(1)$, so the theoretical complexity of REQF is only $O(mt_g \cdot |V| + r^2d)$, which indicates that the simulation of the quantum walk can be replaced by directly conducted in the future technologies, and make the present high complexity greatly reduced certainly.

4 Experiments

In this section, we evaluate the performance of REQF on real-world networks from the tasks of role classification, role detection and visualization, parameter sensitivity, and ablation study.

4.1 Datasets

We conduct experiments on seven real-world networks with unweighted undirected edges, the detailed statistics of the datasets are shown in [Table 2](#).

Table 2: Statistics of datasets

Datasets	Nodes	Edges	Features	Classes
Brazil flights	131	1,003	248	4
Europe flights	399	5,993	652	4
USA flights	1,190	13,599	1,160	4
Cora	2,708	5,429	300	7
Actor	7,758	26,646	2,836	4
LastFM Asia	7,624	27,806	7,842	18
Film	27,312	122,706	1,292	4

Air-traffic networks [11] is about air-traffic information network in different countries including America, Brazil, and Europe, Cora [15] is a citation network related to science, Actor [3] is a social

network of American actors, LastFM Asia [9] is a social network of the LastFM music platform composed of Asian users, and Film [3] is English-language film network.

4.2 Experiment Setting

We compare REQF with classic and advanced role-based network embedding methods including Role2vec [4], FEATHER [9], RED [5], and RESD [11]. All the experimental parameters are set to the default values provided in their papers.

For the REQF, we set the final parameters in the parameter sensitivity experiment. We set quantum walk evolution times $t_g = 16$, biased evolution times $t_l = 4$, biased probability $p_{biased} = 0.9$, and evaluation points $\theta = 5$. As for the datasets with missing features, like the Film dataset, we use ReFex [24] to extract the features of nodes. In VAE, we use a 1-layer MLP model with the ReLU activation function and set the number of hidden layers to 2, and set the number of epochs to 50.

4.3 Role Classification

We conduct the task of role-based node classification on seven real-world networks. In this experiment, we input the embedding into a linear logistic regression classifier, 70% of the data is randomly selected for training, and the rest is used for testing. We run all the methods 20 times, and compute their average F1 score to measure the accuracy of the classification, and the area under the curve (AUC) score to measure the quality of it. The results are shown in Tables 3–5.

Table 3: Macro-F1 score in role classification, with the best score in bold

Datasets	Role2vec	FEATHER	RED	RESD	REQF
Brazil flights	0.348 ± 0.069	0.436 ± 0.062	0.226 ± 0.073	0.742 ± 0.062	0.816 ± 0.043
Europe flights	0.303 ± 0.047	0.508 ± 0.045	0.350 ± 0.022	0.478 ± 0.056	0.587 ± 0.027
USA flights	0.409 ± 0.033	0.564 ± 0.022	0.268 ± 0.021	0.556 ± 0.013	0.592 ± 0.026
Cora	0.752 ± 0.016	0.791 ± 0.016	0.113 ± 0.017	0.145 ± 0.014	0.581 ± 0.012
Actor	0.307 ± 0.011	0.401 ± 0.006	0.225 ± 0.018	0.459 ± 0.014	0.482 ± 0.008
LastFM Asia	0.730 ± 0.012	0.711 ± 0.009	0.112 ± 0.001	0.043 ± 0.003	0.563 ± 0.006
Film	0.308 ± 0.005	0.356 ± 0.005	0.437 ± 0.005	0.433 ± 0.004	0.444 ± 0.003

The results of role classification indicate that the proposed REQF has achieved the best performance in five networks, with 14.6% higher than the best baseline method. However, the performance of REQF on Cora and LastFM Asia networks is lower than FEATHER and Role2Vec separately. The reason may be that the users of these two social networks are mostly mutual strangers, with relatively equal interconnection and less obvious role characteristics, resulting in poor performance in role-based classification. It can be proved that both datasets also show low performance in the RED and RESD methods that are based on role embedding. So, the results demonstrate the effectiveness of REQF in the role classification task.

Table 4: Micro-F1 score in role classification, with the best score in bold

Datasets	Role2vec	FEATHER	RED	RESD	REQF
Brazil flights	0.365 ± 0.076	0.448 ± 0.056	0.279 ± 0.095	0.745 ± 0.057	0.823 ± 0.042
Europe flights	0.304 ± 0.044	0.524 ± 0.050	0.415 ± 0.024	0.488 ± 0.049	0.601 ± 0.032
USA flights	0.412 ± 0.031	0.568 ± 0.019	0.297 ± 0.019	0.564 ± 0.015	0.599 ± 0.028
Cora	0.760 ± 0.013	0.808 ± 0.016	0.142 ± 0.016	0.168 ± 0.012	0.633 ± 0.014
Actor	0.315 ± 0.012	0.424 ± 0.008	0.250 ± 0.008	0.478 ± 0.014	0.486 ± 0.006
LastFM Asia	0.819 ± 0.009	0.809 ± 0.005	0.135 ± 0.002	0.043 ± 0.004	0.716 ± 0.005
Film	0.346 ± 0.005	0.417 ± 0.007	0.757 ± 0.004	0.737 ± 0.006	0.771 ± 0.004

Table 5: AUC scores in role classification, with the best score in bold

Datasets	Role2vec	FEATHER	RED	RESD	REQF
Brazil flights	0.577 ± 0.050	0.632 ± 0.038	0.519 ± 0.063	0.830 ± 0.038	0.882 ± 0.028
Europe flights	0.536 ± 0.029	0.683 ± 0.031	0.610 ± 0.016	0.659 ± 0.033	0.734 ± 0.021
USA flights	0.608 ± 0.021	0.702 ± 0.013	0.531 ± 0.013	0.709 ± 0.010	0.733 ± 0.018
Cora	0.860 ± 0.008	0.875 ± 0.009	0.450 ± 0.009	0.514 ± 0.007	0.828 ± 0.008
Actor	0.543 ± 0.008	0.616 ± 0.005	0.450 ± 0.005	0.652 ± 0.009	0.658 ± 0.004
LastFM Asia	0.904 ± 0.005	0.854 ± 0.003	0.496 ± 0.001	0.036 ± 0.002	0.754 ± 0.003
Film	0.564 ± 0.004	0.611 ± 0.005	0.838 ± 0.003	0.825 ± 0.004	0.848 ± 0.003

4.4 Role Detection

Role detection is one of the important tasks of role-based network embedding. It clusters the embedding into different existing classes to represent different roles in the real network. The experiment is conducted to K-means clustering method using Brazil flights dataset to observe the performances.

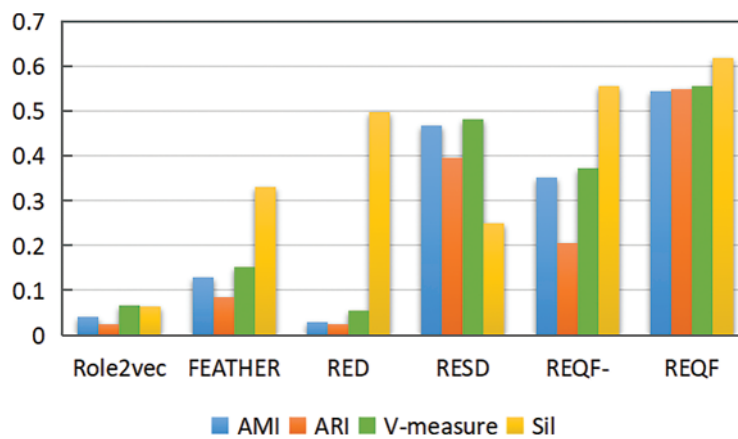
We evaluate the clustering extensively using four popular metrics: Adjusted Mutual Information (AMI) [25] to assess the closeness between the clustering results and the true classes of the samples, Adjusted Rand Index (ARI) [26] to measure the degree of fit between the two data distributions, V-measure [27] to indicate the closeness of the group divisions, and Silhouette score (Sil) [28] to indicate the closeness of each group class. The values of all metrics range from 0 to 1, and a value closer to 1 means a better effect.

The results are shown in Table 6, and REQF-means REQF without VAE to reduce the effect of noise. REQF significantly outperformed all other models in role detection due to considering node role information, features, and the effect of noise at the same time, which other methods ignore.

Compared with the best baseline, REQF is improved by 16.5% in AMI, 38.9% in ARI, 15.5% in V-measure, and 24.1% in Sil. An intuitive comparison is shown in Fig. 4. It is obvious that our REQF shows the best performance in role detection.

Table 6: Results of role detection

Algorithm	AMI	ARI	V-measure	Sil
Role2vec	0.0403	0.0236	0.0659	0.0634
FEATHER	0.1284	0.0859	0.1518	0.3311
RED	0.0284	0.0242	0.0557	0.4981
RES D	0.4668	0.3953	0.4809	0.2489
REQF-	0.3509	0.2062	0.3728	0.5559
REQF	0.5438	0.5489	0.5555	0.6182

**Figure 4:** Histogram of role detection. REQF-means REQF without VAE

4.5 Parameter Sensitivity

For parameters of REQF, we conduct experiments to analyze the sensitivity of all parameters on the Brazil flights network. Fig. 5 plots the micro-F1 score of the role classification task with varying parameters.

- *Effect of t_g .* As shown in Fig. 5a, when evolution times $t_g < 16$, the micro-F1 score fluctuates slightly, and with the increase of the evolution times, the model gradually tends to be stable. However, when $t_g > 16$, the model performance decreases significantly. Therefore, we can get the best model performance when $t_g = 16$.
- *Effect of t_l .* Fig. 5b shows how the biased evolution times t_l affects the performance. We can observe that REQF gets the best results with $t_l = 4$. When $t_l = 2$, although the average score is higher, it fluctuates more making the model more unstable, especially when $t_l = 8$ with the longest box. Thus, the performance does not become better as the biased evolution times increase, which indicates that quantum walk can capture network information and achieve better performance in short evolution times.
- *Effect of p_{biased} .* The results of varying biased probability p_{biased} are shown in Fig. 5c. We can get the optimal mean value when $p_{biased} = 0.2$ and the second-best when $p_{biased} = 0.9$, but it is more stable at $p_{biased} = 0.9$. It confirms that the performance of our model is optimal when the biased probability is 0.9.

- *Effect of θ .* Fig. 5d shows the influence of evaluation points on REQF performance. When $\theta = 1$, we get the highest score but the length of the box is long, which means it fluctuates greatly. While when $\theta = 7$, the length of the box is the shortest and the model is the most stable. On the whole, varying θ has little effect on the mean value, so the appropriate parameter value can be set according to the demand for downstream tasks.

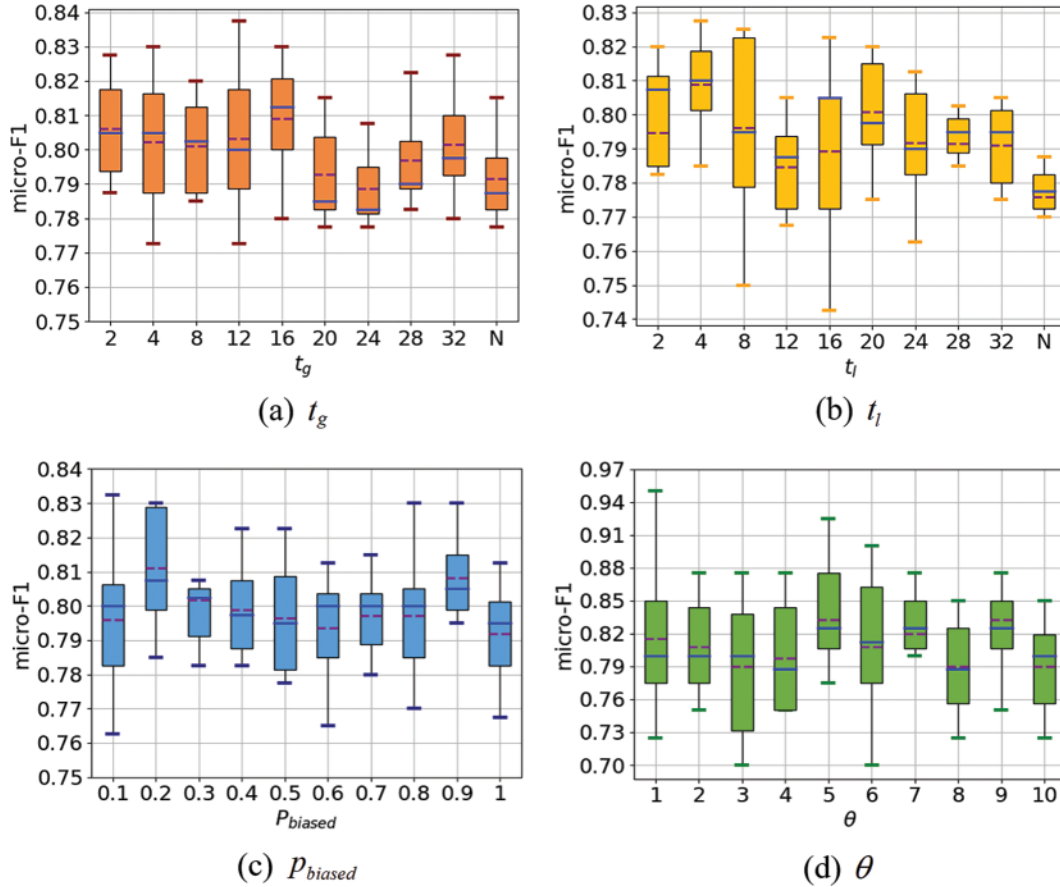


Figure 5: Parameters sensitivity on Brazil Network with varying (a) evolution times, (b) biased evolution times, (c) biased probability, and (d) evaluation points. The solid line denotes the median and the dotted line denotes the average

4.6 Ablation Study

To explore the influence of quantum walk, CF, and VAE modules on the REQF, we compare the performance between REQF and REQF without the quantum walk (REF), feature fusion (REQ), and VAE (REQF-). The results are shown in Fig. 6.

It is obvious that REQF significantly outperforms the other models, showing the best performance. REF is second-best, while it is still inferior to REQF, indicating the necessity of global role information and the effectiveness of the quantum walk module in capturing it. The results of REQF- and REF are close, which proves that noise reduction of VAE plays an important role same as the quantum walk. REQ has the worst results, especially the Sil metric, which is significantly lower than all

the others. It indicates that the CF greatly affects the performance of REQF, especially in the clustering tasks, and it is necessary to consider the different distributions of node features and neighborhoods. Hence, all modules indeed play an important role in REQF, and all of them effectively improve the overall performance of REQF.

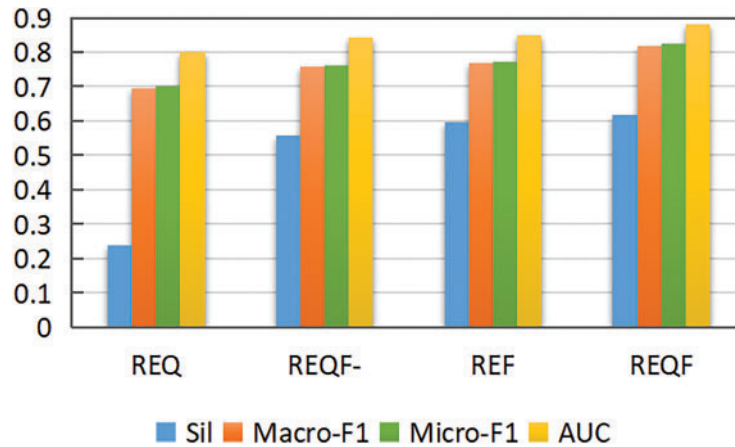


Figure 6: Ablation of REQF, REF, REQ, REQF-, which denote REQF without quantum walk (REF), feature fusion (REQ), and variational auto-encoder (REQF-), respectively

4.7 Visualization

In order to visualize the effect of role clustering, we compare REQF with baselines for visualization on the Brazil flights dataset due to its proper size and uniform classes of node roles. We use t-SNE [29] to get the dimensional reduction visualization, which is sensitive to the local structure of nodes and exhibits more stability than other visualization algorithms [30]. The roles are mapped into the color of points, points with the same color should be close to each other, while points with different colors should be far apart.

As shown in Fig. 7, each node is represented as a point with different colors, and the color indicates the node's role. Points with the same role should be clustered together, that is, points with the same color should be close to each other, while points with different colors should be far apart. The visualization results show that the embedding of the REQF (Fig. 7f) can be well clustered in general, although there are a few individual nodes with incorrect clustering results. This is because REQF considers global and local role information, node features, and noise simultaneously. The points in different colors of Role2vec (Fig. 7a) are mixed up as it only considers the local structure of nodes and points cannot be classified well. RED (Fig. 7b) is better than Role2vec as it considers global role information. FEATHER (Fig. 7c) considers node features so it clusters closer while it still cannot effectively achieve role clustering. RESD (Fig. 7d) is more effective as the points of the same color are quite close to each other, but there is no obvious division of the different roles of nodes as it ignores global information. In addition, we conduct the experiment of REQF-(Fig. 7e) that REQF without VAE, and its visualization illustrates that the VAE indeed reduces the effect of noise and enables the model to have better role clustering ability. Therefore, REQF has the best visualization effect and the model can divide the roles of nodes well in the real network.

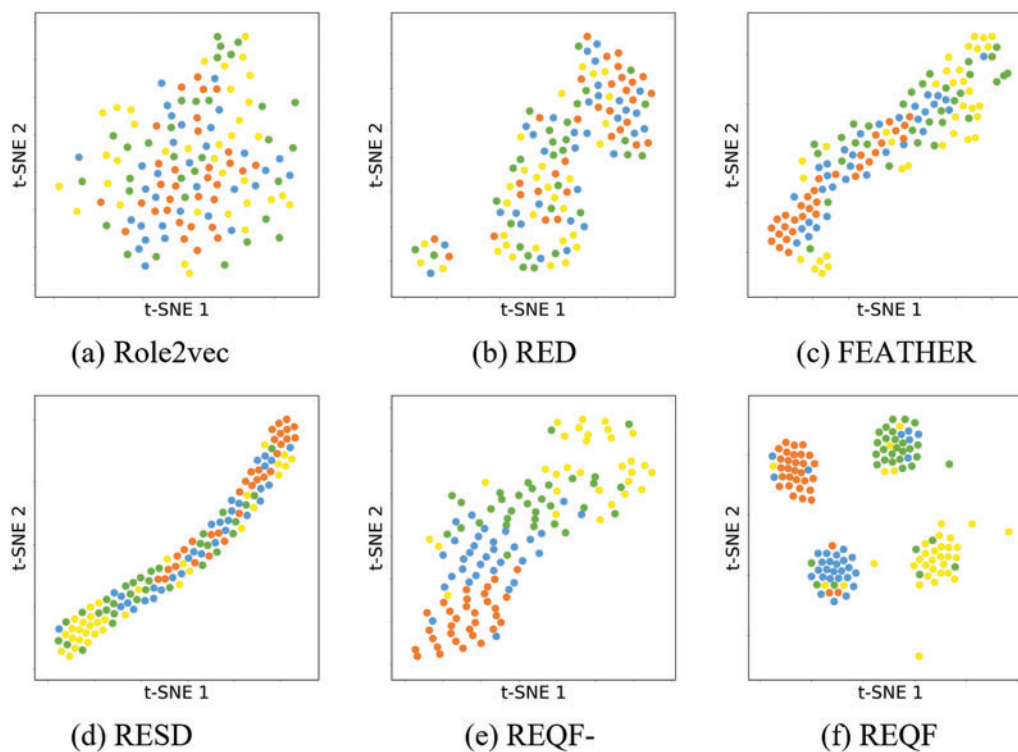


Figure 7: Visualization of different methods on Brazil network in 2D space

5 Conclusion

In this paper, we propose REQF to generate role-based network embedding via quantum walk and its weighted feature fusion, which simultaneously considers node role information, node features, and noise. REQF utilizes quantum walk to capture the global and local role information of nodes, leverage its weighted characteristic function for feature fusion, and finally use VAE to reduce the effect of noise. The experimental results demonstrate the effectiveness and stability of the REQF on real-world datasets for downstream tasks. We also demonstrate the importance of each module and explore optimal values of parameters. For the limitation of our work, on the one hand, we only use a simple multi-layer perceptron, which may not be able to reduce the effect of noise sufficiently. On the other hand, our proposed method only focuses on homogeneous networks, and it cannot be expanded to heterogeneous graph networks. For future work, we can try a more efficient deep-learning framework to obtain a better representation of the role-based network. We can also expand to explore the field of heterogeneous role-based graph networks.

Funding Statement: This work was supported in part by the National Nature Science Foundation of China (Grant 62172065) and the Natural Science Foundation of Chongqing (Grant cstc2020jcyj-msxmX0137).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] R. A. Rossi and N. K. Ahmed, "Role discovery in networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 4, pp. 1112–1131, 2015.
- [2] J. Ma, M. Li and H. J. Li, "Traffic dynamics on multilayer networks with different speeds," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 3, pp. 1697–1701, 2022.
- [3] P. Jiao, X. Guo, T. Pan, W. Zhang, Y. Pei *et al.*, "A survey on role-oriented network embedding," *IEEE Transactions on Big Data*, vol. 8, no. 4, pp. 933–952, 2022.
- [4] N. K. Ahmed, R. A. Rossi, J. B. Lee, T. L. Willke, R. Zhou *et al.*, "Role-based graph embeddings," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 5, pp. 2401–2415, 2022.
- [5] X. Wang, S. Jian, K. Lu, Y. Zhang and K. Liu, "RED: Learning the role embedding in networks via discrete-time quantum walk," *Applied Intelligence*, vol. 52, no. 2, pp. 1493–1507, 2022.
- [6] K. Henderson, B. Gallagher, T. Eliassi-Rad, H. Tong, S. Basu *et al.*, "RoLX: Structural role extraction & mining in large graphs," in *Proc. KDD*, New York, NY, USA, pp. 1231–1239, 2012.
- [7] X. Guo, W. Zhang, W. Wang, Y. Yu, Y. Wang *et al.*, "Role-oriented graph auto-encoder guided by structural information," in *Proc. DASFAA*, Jeju, South Korea, pp. 466–481, 2020.
- [8] P. Jiao, Q. Tian, W. Zhang, X. Guo, D. Jin *et al.*, "Role discovery-guided network embedding based on autoencoder and attention mechanism," *IEEE Transactions on Cybernetics*, vol. 53, no. 1, pp. 365–378, 2023.
- [9] B. Rozemberczki and R. Sarkar, "Characteristic functions on graphs: Birds of a feather, from statistical descriptors to parametric models," in *Proc. CIKM*, New York, NY, USA, pp. 1325–1334, 2020.
- [10] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proc. ICLR*, Banff, AB, Canada, pp. 14–16, 2014.
- [11] W. Zhang, X. Guo, W. Wang, Q. Tian, L. Pan *et al.*, "Role-based network embedding via structural features reconstruction with degree-regularized constraint," *Knowledge-Based Systems*, vol. 218, no. 3, pp. 106872, 2021.
- [12] R. A. Rossi, D. Jin, S. Kim, N. K. Ahmed, D. Koutra *et al.*, "On proximity and structural role-based embeddings in networks: Misconceptions, techniques, and applications," *ACM Transactions on Knowledge Discovery from Data*, vol. 14, no. 5, pp. 1–37, 2020.
- [13] C. Donnat, M. Zitnik, D. Hallac and J. Leskovec, "Learning structural node embeddings via diffusion wavelets," in *Proc. KDD*, New York, NY, USA, pp. 1320–1329, 2018.
- [14] X. Ma, G. Qin, Z. Qiu, M. Zheng and Z. Wang, "Riwalk: Fast structural node embedding via role identification," in *Proc. ICDM*, Beijing, China, pp. 478–487, 2019.
- [15] Y. Jin, G. Song and C. Shi, "GraLSP: Graph neural networks with local structural patterns," in *Proc. AAAI*, New York, NY, USA, pp. 4361–4368, 2020.
- [16] S. E. Venegas-Andraca, "Quantum walks: A comprehensive review," *Quantum Information Processing*, vol. 11, no. 5, pp. 1015–1106, 2012.
- [17] M. Islam Kamran, M. A. Khan, S. A. Alsuhibany, Y. Yasin Ghadi, Arshad *et al.*, "A highly secured image encryption scheme using quantum walk and chaos," *Computers, Materials & Continua*, vol. 73, no. 1, pp. 657–672, 2022.
- [18] Y. Zhang, L. Wang, R. C. Wilson and K. Liu, "An r-convolution graph kernel based on fast discrete-time quantum walk," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 1, pp. 292–303, 2022.
- [19] X. Zhang, L. Yao and F. Yuan, "Adversarial variational embedding for robust semi-supervised learning," in *Proc. KDD*, New York, NY, USA, pp. 139–147, 2019.
- [20] Q. Yi, N. Yang and P. Yu, "Dual adversarial variational embedding for robust recommendation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 2, pp. 1421–1433, 2023.
- [21] Y. He, Y. Dong, P. Cui, Y. Jiao, X. Wang *et al.*, "Purify and generate: Learning faithful item-to-item graph from noisy user-item interaction behaviors," in *Proc. KDD*, New York, NY, USA, pp. 3002–3010, 2021.
- [22] V. Chauhan, S. Negi, D. Jain, P. Singh, A. K. Sagar *et al.*, "Quantum computers: A review on how quantum computing can boom AI," in *Proc. ICACITE*, Greater Noida, India, pp. 559–563, 2022.

- [23] X. Wang, K. Lu, Y. Zhang and K. Liu, "QSIM: A novel approach to node proximity estimation based on discrete-time quantum walk," *Applied Intelligence*, vol. 51, no. 4, pp. 2574–2588, 2021.
- [24] K. Henderson, B. Gallagher, L. Li, L. Akoglu, T. Eliassi-Rad *et al.*, "It's who you know: Graph mining using recursive structural features," in *Proc. KDD*, New York, NY, USA, pp. 663–671, 2011.
- [25] M. E. J. Newman, G. T. Cantwell and J. G. Young, "Improved mutual information measure for clustering, classification, and community detection," *Physical Review E*, vol. 101, no. 4, pp. 042304, 2020.
- [26] M. J. Warrens and H. van der Hoef, "Understanding the adjusted rand index and other partition comparison indices based on counting object pairs," *Journal of Classification*, vol. 39, no. 3, pp. 487–509, 2022.
- [27] J. Nowosad and T. F. Stepinski, "Spatial association between regionalizations using the information-theoretical v-measure," *International Journal of Geographical Information Science*, vol. 32, no. 12, pp. 2386–2401, 2018.
- [28] K. R. Shahapure and C. Nicholas, "Cluster quality analysis using silhouette score," in *Proc. DSAA*, Sydney, NSW, Australia, pp. 747–748, 2020.
- [29] F. Anowar, S. Sadaoui and B. Selim, "Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE)," *Computer Science Review*, vol. 40, pp. 100378, 2021.
- [30] S. Kang and S. Kyun Kim, "Outlier behavior detection for indoor environment based on t-SNE clustering," *Computers, Materials & Continua*, vol. 68, no. 3, pp. 3725–3736, 2021.