Learning Graph Convolutional Networks based on Quantum Vertex Information Propagation

(Extended Abstract)

Lu Bai*, Yuhang Jiao*, Lixin Cui*, Luca Rossi†, Yue Wang*, Philip S. Yu‡, Edwin R. Hancock§
*Engineering Research Center of State Financial Security, Ministry of Education,
Central University of Finance and Economics, Beijing, China.
†School of Electronic Engineering and Computer Science, Queen Mary University of London, UK.
‡Department of Computer Science, University of Illinois at Chicago, US.
§Department of Computer Science, University of York, York, UK.

***Abstract*—This paper proposes a novel Quantum Spatial Graph Convolutional Neural Network (QSGCNN) model that can directly learn a classification function for graphs of arbitrary sizes. The main idea is to define a new quantum-inspired spatial graph convolution associated with pre-transformed fixed-sized aligned grid structures of graphs, in terms of quantum information propagation between grid vertices of each graph. We show that the proposed QSGCNN model can significantly reduce either the information loss or the notorious tottering problem arising in existing spatially-based Graph Convolutional Network (GCN) models. Experiments on benchmark graph datasets demonstrate the effectiveness of the proposed QSGCNN model.**

I. Introduction

GCNs [1] have proven to be effective tools to extract meaningful features for graph analysis. One prevalent way of defining novel GCN models is to generalize the graph convolution operation to the spatial structure of a graph, by directly defining an operation on neighboring vertices [2]. These so-called spatially-based GCN models are not restricted to the same-sized graph structures, and have been widely employed for graph classification tasks.

Unfortunately, for the local-level vertex features extracted from the convolution operation, most existing spatially-based GCN models tend to directly sum them up through a SumPooling layer [3] or only preserve the vertices with high ranks through a SortPooling layer [2]. Thus, these GCN models tend to suffer from drawbacks of information loss, and have fairly poor performance on graph classification.

This paper aims to address the aforementioned problems by developing a novel QSGCNN model [4]. Specifically, we make the following contributions. **First**, we convert graphs of arbitrary sizes into fixed-sized aligned vertex grid structures, through a transitive graph alignment method. **Second**, we define a new spatial graph convolution operation associated with the grid structure of each graph, by propagating vertex information through quantum walks. We show that the quantum-inspired convolution simultaneously reduces the information loss and tottering problems arising in classical spatial graph convolution operation. **Third**, we empirically demonstrate the effectiveness of the QSGCNN model.

II. Formulations and Preliminary Concepts

We aim to develop a new spatial graph convolution layer by propagating information between vertices. To this end, we employ the vertex information propagation process of the continuous-time quantum walk (CTQW), that is the quantum analogue of the classical continuous-time random walk (CTRW) [5]. Since the evolution of the CTQW is not dominated by the low frequency components of the Laplacian spectrum, it can not only reduce the tottering problem arising in classical CTRW, but also better discriminate between different graphs.

Specifically, we propose to employ the average mixing matrix (AMM) to capture the time-averaged behaviour of the CTQW being transmitted between the graph vertices. For a sample graph $G(V,E)$ with vertex set V and edge set E, we adopt the adjacency matrix as the Hamiltonian. Based on [6], the behaviour of a CTQW over $G(V,E)$ at time t can be summarized using the mixing matrix

$$Q_M(t) = U(t) \circ U(-t) = e^{i\mathcal{H}t} \circ e^{-i\mathcal{H}t}, \tag{1}$$

where the operation symbol $\circ$ represents the Schur-Hadamard product of $e^{i\mathcal{H}t}$ and $e^{-i\mathcal{H}t}$. Because U is unitary, $Q_M(t)$ is a doubly stochastic matrix and each entry $Q_M(t)_{uv}$ indicates the probability of the CTQW visiting vertex v at time t when the walk initially starts from vertex u. However, $Q_M(t)$ cannot converge. Thus, we enforce convergence by taking a time average, and we take the Cesàro mean to define the AMM

$$Q = \lim_{T\to\infty} \int_0^T Q_M(t)dt, \tag{2}$$

where each entry $Q_{v_iv_j}$ of the AMM Q represents the average probability for a CTQW to visit vertex v_j starting from vertex v_i, and Q is still a doubly stochastic matrix. Note that, since the entries of Q are rational numbers, one can easily compute Q from the spectrum of the Hamiltonian.

III. The Proposed QSGCNN Model

The architecture. The architecture and definition of the proposed QSGCNN model is exhibited in Fig.1 [4]. Specifically, the architecture is composed of three sequential stages, i.e., 1) the grid structure construction and input layer, 2) the

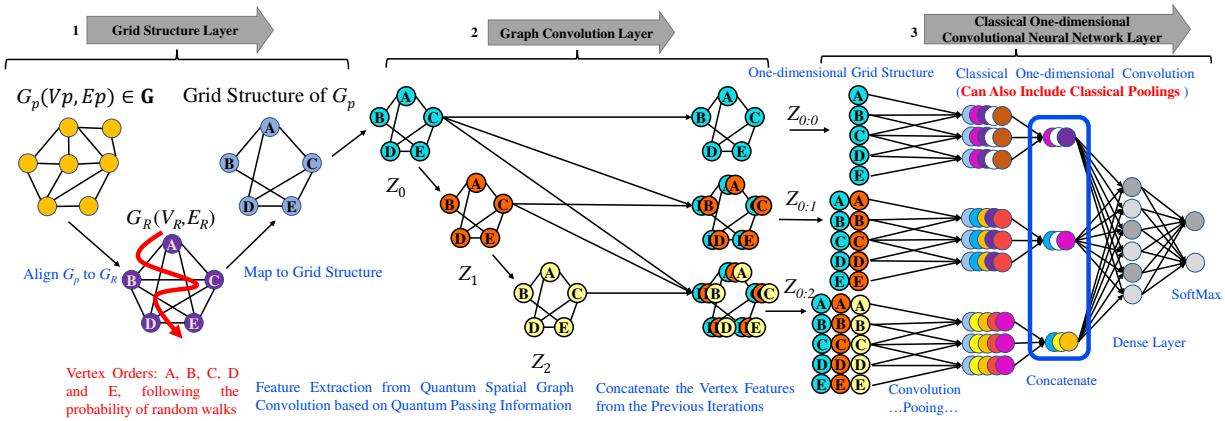

Fig. 1. **The architecture of the proposed QSGCNN model**. **(1)** An prototype graph $G_R(V_R, E_R)$ is first constructed by locating M ($M = |V_R|$) centroids as the prototype vertex representations from the vectorial vertex signatures over all graphs, through the classical k-means method. Then, each input graph $G_p(V_p, E_p) \in \mathbf{G}$ of arbitrary size is aligned to G_R, by identifying the structure correspondence information between the vertices of G_p and G_R. Then, the vertices of G_p aligned to the same vertex of G_R will be mapped into the same aligned vertex, where each aligned vertex follows the same vertex order of the corresponding vertex of G_R, i.e., these new aligned vertices follow the same vertex spatial positions of G_R. Here, the red curved arrow on the graph G_R indicates the predetermined spatial orders of its vertices. This process in turn forms a nature fixed-sized aligned vertex grid structure, where a standard CNN can be directly performed. Since the above construction process will not discard any original vertex of G_p, the resulting aligned vertex grid structure can reduce the problem of information loss that arises in existing graph convolutional neural network models associated with the SortPolling operation. **(2)** The grid structure of G_p is passed through multiple quantum spatial graph convolution layers to extract multi-scale vertex features, where the vertex information is propagated between specified vertices following the average mixing matrix. **(3)** Since the graph convolution layers preserve the original vertex orders of the input grid structure, the concatenated vertex features through the graph convolution layers form a new vertex grid structure for G_p. This vertex grid structure is then passed to a traditional CNN layer to learn a classification function. Note, vertex features are visualized as different colors.

quantum spatial graph convolution layer, and 3) the traditional convolutional neural network and Softmax layers.

The quantum convolution. For each graph G_p, we commence by converting its vertex feature matrix X_p and its associated vertex adjacency matrix A_p into the fixed-sized aligned vertex grid structure $\widehat{X}_p \in \mathbb{R}^{M\times c}$ (i.e., the aligned grid vertex feature matrix) and the associated aligned grid vertex adjacency matrix $\widehat{A}_p \in \mathbb{R}^{M\times M}$. The quantum spatial graph convolution operation of the QSGCNN model is defined as

$$Z = \mathrm{Relu}(Q\widehat{X}_pW), \tag{3}$$

where Relu is a nonlinear activation function, Q is the AMM of the CTQW on $\widehat{A}_p$ of G_p defined in Section II, $W \in \mathbb{R}^{c\times c'}$ is the matrix of trainable parameters of the proposed graph convolutional layer, and $Z \in \mathbb{R}^{M\times c'}$ is the output matrix.

Discussions. The proposed QSGCNN model has a number of significant theoretical differences with existing methods, explaining its effectiveness. **First**, unlike the classical spatially-based GCN models DGCNN [2] and FAGCN [1], the QSGCNN model employs the AMM of CTQW rather than the original vertex adjacency matrix to determine how to pass the information among the vertices. Since the CTQW has a better ability to distinguish different graph structures, the QSGCNN model can extract more discriminative vertex features. **Second**, in order to maintain the scale of the extracted vertex features, the convolution operations of the DGCNN and FAGCNN models need to perform a multiplication by the inverse of the vertex degree matrix. This can be seen as a normalizing process and assigns equal weights between neighbor vertices. By contrast, the convolution operation of the QSGCNN model assigns an average visiting probability distribution of the CTQW to specified vertices with each vertex having a different weight. Therefore, the QSGCNN model can better discriminate the mutual-influences between specified vertices. **Third**, the QSGCNN models propagate vertex information in terms of CTQW, addressing the tottering problem arising in most existing spatially-based GCN models that are theoretical related to the classical Weisfeiler-Lehman (WL) algorithm [2]. **Forth**, the QSGCNN model is defined on aligned grid structures that are transformed from original graph structures without discarding any vertex, reducing the information loss arising in existing spatially-based GCN models associated with a SumPooling or a SortPooling layer.

IV. Experimental Evaluation

Datasets. We empirically evaluate the classification performance of the proposed QSGCNN model on nine benchmark graph datasets from bioinformatics and social networks. These datasets include: MUTAG, PTC, NCI1, PROTEINS, D&D, COLLAB, IMDB-B, IMDB-M and RED-B.

Result. Experiments in [4] indicates that the proposed QSGCNN model outperforms state-of-the-art graph kernels and GCNs on benchmark datasets, demonstrating the effectiveness.

References

[1] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *CoRR*, vol. abs/1609.02907, 2016.

[2] M. Zhang, Z. Cui, M. Neumann, and Y. Chen, "An end-to-end deep learning architecture for graph classification," in *Proceedings of AAAI*, 2018.

[3] J. Atwood and D. Towsley, "Diffusion-convolutional neural networks," in *Proceedings of NIPS*, 2016, pp. 1993–2001.

[4] L. Bai, Y. Jiao, L. Cui, L. Rossi, Y. Wang, P. S. Yu, and E. R. Hancock, "Learning graph convolutional networks based on quantum vertex information propagation," *IEEE Trans. Knowl. Data Eng.*, vol. early access, no. 10.1109/TKDE.2021.3106804, 2021.

[5] E. Farhi and S. Gutmann, "Quantum computation and decision trees," *Physical Review A*, vol. 58, p. 915, 1998.

[6] C. Godsil, "Average mixing of continuous quantum walks," *Journal of Combinatorial Theory, Series A*, vol. 120, no. 7, pp. 1649–1662, 2013.