A Quantum Jensen-Shannon Graph Kernel using Discrete-time Quantum Walks

Lu Bai[1,2], Luca Rossi[3], Peng Ren[4], Zhihong Zhang[5⋆], and Edwin R. Hancock[2]

[1]School of Information, Central University of Finance and Economics, Beijing, China
[2]Department of Computer Science, University of York, York, UK
[3]School of Computer Science, University of Birmingham, Birmingham, UK
[4]College of Information and Control Engineering, China University of Petroleum (Huadong), Shandong Province, Qingdao, P.R. China
[5]Software school, Xiamen University, Xiamen, Fujian, China

Abstract. In this paper, we develop a new graph kernel by using the quantum Jensen-Shannon divergence and the discrete-time quantum walk. To this end, we commence by performing a discrete-time quantum walk to compute a density matrix over each graph being compared. For a pair of graphs, we compare the mixed quantum states represented by their density matrices using the quantum Jensen-Shannon divergence. With the density matrices for a pair of graphs to hand, the quantum graph kernel between the pair of graphs is defined by exponentiating the negative quantum Jensen-Shannon divergence between the graph density matrices. We evaluate the performance of our kernel on several standard graph datasets, and demonstrate the effectiveness of the new kernel.

1 Introduction

Graph based representations are powerful tools for structural analysis in pattern recognition. One challenge of classifying graphs is how to convert the discrete graph structures into numeric features. One way is to use graph kernels. The main advantage of using graph kernels is that they characterize graph features in a high dimensional space and thus better preserve graph structures.

Generally speaking, a graph kernel is a similarity measure between a pair of graphs [1]. To extend the large spectrum of kernel methods from the general machine learning domain to the graph domain, Haussler [2] has proposed a generic way, namely the R-convolution, to define a graph kernel. For a pair of graphs, an R-convolution kernel is computed by decomposing each graph into smaller subgraphs and counting the number of isomorphic subgraph pairs between the two original graphs. Thus, a new type of decomposition of a graph usually results in a new graph kernel. Following this scenario, Kashima et al. [3] introduced the random walk kernel, which is based on the enumeration of common random walks between two graphs. Borgwardt et al. [4], on the other hand, proposed a shortest path kernel by counting the numbers of matching shortest paths over the graphs.

⋆ Corresponding author, email: zhihong@xmu.edu.cn.

Aziz et al. [5] introduced a backtrackless kernel using the cycles identified by the Ihara zeta function [6] in a pair of graphs. Shervashidze et al. [7] developed a fast subtree kernel by comparing pairs of subtrees identified by the Weisfeiler-Lehman (WL) algorithm. Some other alternative R-convolution kernels include a) the segmentation graph kernel developed by Harchaoui and Bach [8], b) the point cloud kernel developed by Bach [9], and c) the (hyper)graph kernel based on directed subtree isomorphism tests [10].

Recently, a number of alternative graph kernel measures have been introduced in the literature. These are based on the computation of the mutual information between two graphs in terms of the classical Jensen-Shannon divergence. In information theory, the classical Jensen-Shannon divergence is a dissimilarity measure between probability distributions. In [11–13], we have used the classical Jensen-Shannon divergence to define a Jensen-Shannon graph kernel. Unlike the R-convolution kernels that count the number of isomorphic substructure pairs, the Jensen-Shannon graph kernel is defined in terms of the entropy difference between a pair of graphs and their composite graph (e.g., the disjoint union graph or the product graph formed by the pair of graphs). Here, the entropy of a graph can be either the von Neumann entropy (associated with the graph spectrum information) or the Shannon entropy (associated with the steady state random walk or the information functional). Both the von Neumann entropy and the Shannon entropy of a graph can be directly computed without the need to decompose the graph into substructures. As a result, the Jensen-Shannon graph kernel avoids the computational burden of comparing all pairs of substructures for a pair of graphs. To develop this work further, in [14–16] we have introduced a new quantum Jensen-Shannon graph kernel using the quantum Jensen-Shannon divergence [17, 18] and the continuous-time quantum walk [19]. Here the basic idea is to associate with each graph a mixed quantum state representing the time evolution of a quantum walk. The kernel between a pair of graphs is then defined as the quantum Jensen-Shannon divergence between their corresponding density matrices [14].

The aim of this paper is to develop our previous quantum Jensen-Shannon kernels [14–16] one step further. We propose a new quantum Jensen-Shannon kernel for graphs using the discrete-time quantum walk. The discrete-time quantum walk is the quantum analogue of the discrete-time classical random walk [19]. Remarkably, the discrete-time quantum walk possesses a number of interesting properties that are not exhibited by its classical counterpart. In fact, the behaviour of the discrete-time quantum walk is governed by a unitary matrix rather than a stochastic matrix, as in the case of the classical random walk. As a consequence, its evolution is reversible and non-ergodic. However, unlike the continuous-time quantum walk, where the state space is the graph vertex set, the state space of the discrete-time quantum walk is the set of arcs residing on the graph edges. More specifically, given an undirected graph $G(V, E)$, each edge $\{u, v\} \in E$ is replaced by a pair of directed arcs (u, v) and (v, u), and the set of arcs is denoted by E_d. Then, the state space for the discrete-time quantum walk is the set of arcs E_d. We are interested in developing a new quantum Jensen-

Shannon kernel where the graph structure is probed by means of a discrete-time quantum walk. To commence, we perform a discrete-time quantum walk on each graph and we compute a mixed quantum state represented by a density matrix. With the density matrices for a pair of graphs to hand, the quantum graph kernel between the pair of graphs is defined by exponentiating the negative quantum Jensen-Shannon divergence between the graph density matrices. We evaluate the performance of our new kernel on several standard graph datasets from both bioinformatics and computer vision. The experimental results demonstrate the effectiveness of the proposed graph kernel. Our new kernel is shown to be competitive to state-of-the-art graph kernels.

2 Quantum Mechanical Background

In this section, we introduce the quantum mechanical formalism that will be used in this work. We commence by reviewing the concept of discrete-time quantum walk on a graph. Furthermore, we describe how to associate with each graph a density matrix describing the quantum walk evolution. Then, we show how to compute the von Neumann entropy of a graph through its density matrix. Finally, we discuss the relationship between the Perron-Frobenius operator [20] and the transition matrix of the discrete-time quantum walk, and thus explain the advantage of the discrete-time quantum walk over its continuous-time version.

2.1 Discrete-time Quantum Walks

The discrete-time quantum walk is the quantum counterpart of the discrete-time classical random walk [21]. To simulate the evolution of a discrete-time quantum walk on a graph $G(V,E)$, we first replace each edge $e(u,v) \in E$ with a pair of directed arcs $e_d(u,v)$ and $e_d(v,u)$. This in turn ensures the reversibility of the quantum process. Let us denote the new set of arcs as E_d. Then, the state space for the discrete-time quantum walk is E_d, and we denote the state corresponding to the walker being on the arc $e_d(u,v)$ as $|uv\rangle$. A general state of the walk is

$$|\psi\rangle = \sum_{e_d(u,v)\in E_d} \alpha_{uv}|uv\rangle, \tag{1}$$

where the quantum amplitudes α_{uv} are complex, i.e., $\alpha_{uv} \in \mathbb{C}$. The probability that the walk is in the state $|uv\rangle$ is given by $\Pr(|uv\rangle) = \alpha_{uv}\alpha_{uv}^*$, where α_{uv}^* is the complex conjugate of α_{uv}.

The evolution of the state vector between the steps t and $t+1$ is determined by the transition matrix $\boldsymbol{U}$. The entries of $\boldsymbol{U}$ determine the transition probabilities between states, i.e., $|\psi_{t+1}\rangle = \boldsymbol{U}|\psi_t\rangle$. Since the evolution of the walk is linear and conserves probability, the matrix $\boldsymbol{U}$ must be unitary, i.e., $\boldsymbol{U}^{-1} = \boldsymbol{U}^{\dagger}$, where $\boldsymbol{U}^{\dagger}$ denotes the Hermitian transpose of $\boldsymbol{U}$.

It is usual to adopt the Grover diffusion matrix [22] as the transition matrix. Using the Grover diffusion matrices, the transition matrix $\boldsymbol{U}$ has entries

$$U_{(u,v),(w,x)} = \begin{cases} \frac{2}{d_x} - \delta_{ux}, & v = w; \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where $U_{(u,v),(w,x)}$ gives the quantum amplitude for the transition $e_d(u,v) \rightarrow e_d(w,x)$ and δ_{ux} is the Kronecker delta, i.e., $\delta_{ux} = 1$ if $u = x$ and 0 otherwise. Given a state $|u_1 v\rangle$, the Grover matrix assigns the same amplitudes to all transitions $|u_1 v\rangle \rightarrow |v u_i\rangle$, and a different amplitude to the transition $|u_1 v\rangle \rightarrow |v u_1\rangle$, where u_i denotes a neighbour of v. Finally, note that although the entries of $\boldsymbol{U}$ are real, they can be negative as well as positive. It is important to stress that, as a consequence of this, negative quantum amplitudes can arise during the evolution of the walk. In other words, the definition in Eq.(2) allows *destructive interference* to take place.

2.2 A Density Matrix from the Mixed State

In quantum mechanics, a pure state can be described as a single ket vector. A quantum system, however, can also be in a mixed state, i.e., a statistical ensemble of pure quantum states $|\psi_i\rangle$, each with probability p_i. The density matrix (or density operator) of such a system is defined as

$$\rho = \sum_i p_i \left|\psi_i\right\rangle \left\langle\psi_i\right| \tag{3}$$

Assume a sample graph $G(V,E)$. Let $|\psi_t\rangle$ denote the state corresponding to a discrete-time quantum walk that has evolved from the step $t = 0$ to the step $t = T$. We define the time-averaged density matrix ρ_G^T for $G(V,E)$ as

$$\rho_G^T = \frac{1}{T+1} \sum_{t=0}^{T} \left|\psi_t\right\rangle \left\langle\psi_t\right| . \tag{4}$$

Since $|\psi_t\rangle = \boldsymbol{U}^t |\psi_0\rangle$, where $\boldsymbol{U}$ is the transition matrix of the discrete-time quantum walk, Eq.(4) can be re-written in terms of the initial state $|\psi_0\rangle$ as

$$\rho_G^T = \frac{1}{T+1} \sum_{t=0}^{T} (\boldsymbol{U})^t \left|\psi_0\right\rangle \left\langle\psi_0\right| (\boldsymbol{U}^\top)^t. \tag{5}$$

As a result, the density matrix ρ_G^T describes a quantum system that has an equal probability of being in each of the pure states defined by the evolution of the discrete-time quantum walk from step $t = 0$ to step $t = T$.

2.3 The von Neumann Entropy of a Graph

In quantum mechanics, the *von Neumann entropy* [23] H_N of a density matrix ρ is defined as $H_N = -\mathrm{tr}(\rho \log \rho) = -\sum_i \xi_i \ln \xi_i$, where $\xi_1, \ldots, \xi_n$ denote the eigenvalues of ρ. Note that if the quantum system is in a pure state $|\psi_i\rangle$ with probability $p_i = 1$, then the Von Neumann entropy $H_N(\rho) = -\mathrm{tr}(\rho \log \rho)$ is zero. On the other hand, a mixed state generally has a non-zero Von Neumann entropy associated with its density matrix. Here we propose to compute the von

Neumann entropy for each graph using the density matrix defined in Eq.(5). Consider a graph $G(V,E)$, the von Neumann entropy of $G(V,E)$ is defined as

$$H_N(\rho_G) = -\mathrm{tr}(\rho_G^T \log \rho_G^T) = -\sum_j^{|V|} \lambda_j^G \log \lambda_j^G, \tag{6}$$

where $\lambda_1^G, \ldots, \lambda_j^G, \ldots, \lambda_{|V|}^G$ are the eigenvalues of ρ_G^T.

2.4 Relation to the Perron-Frobenius Operator

In [20], Ren et al. have demonstrated that the Perron-Frobenius operator can be represented in terms of the transition matrix of discrete-time quantum walks. To show this connection, we first introduce the definitions of the directed line graph and the positive support of a matrix.

Definition 1 For a sample graph $G(V,E)$, the directed line graph $OLG(V_L, E_{dL})$ is a dual representation of $G(V,E)$. To obtain $OLG(V_L, E_{dL})$, we first construct the associated symmetric digraph $SDG(V, E_d)$ of $G(V,E)$, by replacing every edge $e(u,w) \in E(G)$ by a pair of reverse arcs, i.e., directed edges $e_d(u,w) \in E_d(G)$ and $e_d(w,u) \in E_d(G)$ for $u, w \in V$. The directed line graph $OLG(V_L, E_{dL})$ is the directed graph with vertex set V_L and arc set E_{dL} defined as follows

$$\begin{aligned} V_L &= E_d(SDG), \\ E_{dL} &= \{(e_d(u,v), e_d(v,w)) \in E_d(SDG) \times E_d(SDG) \mid u,v,w \in V, u \neq w\}. \end{aligned} \tag{7}$$

The Perron-Frobenius operator $\boldsymbol{T} = [T_{i,j}]_{|V_L|\times|V_L|}$ of $G(V,E)$ is the adjacency matrix of the associated directed line graph $OLG(V_L, E_{dL})$. □

Definition 2 The positive support $\mathrm{S}^+(\boldsymbol{M}) = [s_{i,j}]_{m\times n}$ of the matrix $\boldsymbol{M} = [M_{i,j}]_{m\times n}$ is defined to be a matrix with entries

$$s_{i,j} = \begin{cases} 1, & M_{i,j} > 0, \\ 0, & \text{otherwise}, \end{cases} \tag{8}$$

where $1 \leq i \leq m, 1 \leq j \leq n$. □

Based on the definition in [20], we can re-define the Perron-Frobenius operator in terms of the unitary matrix of the discrete-time quantum walk. Let $G(V,E)$ be a sample graph and $\boldsymbol{U}$ be the unitary matrix associated with the discrete-time quantum walk on $G(V,E)$. The Perron-Frobenius operator $\boldsymbol{U}$ of $G(V,E)$ is

$$\boldsymbol{T} = \boldsymbol{S}^+(\boldsymbol{U}^\top). \tag{9}$$

Def.1, Def.2 and Eq.(9) show us how the discrete-time quantum walk and the Perron-Frobenius operator (i.e., the directed line graph) are co-related. For a graph $G(V,E)$ and its directed line graph $OLG(V_L, E_{dL})$, V_L is just the state space of the discrete-time quantum walk on $G(V,E)$, i.e., each vertex in

$OLG(V_L, E_{dL})$ corresponds to a unique directed arc residing on the corresponding edge in $G(V, E)$. Moreover, if there is a directed edge from a vertex $v_L \in V_L$ to a vertex $u_L \in V_L$, the transition of the quantum walk on $G(V, E)$ is allowed from the arc corresponding to v_L to the arc corresponding to u_L, and vice versa. As a result, the discrete-time quantum walk on a graph can also be seen as a walk performed on its directed line graph. The state space of the walk is the vertex set of the line graph, and the transition of the walk relies on the connections between pairs of vertices in the line graph.

Furthermore, in [10, 20], we have observed that the directed line graph of a graph possesses some interesting properties that are not available on the original graph. For instance, compared to the original graph the line graph spans a higher dimensional feature space and thus exposes richer graph characteristics. This is because the cardinality of the vertex set for the line graph is much greater than, or at least equal to, that of the original graph. This property suggests that the discrete-time quantum walk may reflect richer graph characteristics than the continuous-time quantum walk on the original graph.

Finally, since the discrete-time quantum walk can be seen as a walk on the line graph and the state space of the walk is the vertex set of the line graph, we propose to use the rooting of the in-degree distribution of the line graph as the initial state of the discrete-time quantum walk.

3 A Quantum Jensen-Shannon Graph Kernel

In this section, we develop a new quantum Jensen-Shannon kernel for graphs by using the quantum Jensen-Shannon divergence and the discrete-time quantum walk. We commence by reviewing the concept of the classical and quantum Jensen-Shannon divergence. Finally, we give the definition of the new kernel.

3.1 Classical and Quantum Jensen-Shannon Divergence

The classical Jensen-Shannon divergence is a non-extensive mutual information measure defined between probability distributions. Consider two (discrete) probability distributions $\mathcal{P} = (p_1, \ldots, p_a, \ldots, p_A)$ and $\mathcal{Q} = (q_1, \ldots, q_b, \ldots, q_B)$, then the classical Jensen-Shannon divergence between $\mathcal{P}$ and $\mathcal{Q}$ is defined as

$$D_{JS}(\mathcal{P}, \mathcal{Q}) = H_S\Big(\frac{\mathcal{P}+\mathcal{Q}}{2}\Big) - \frac{1}{2}H_S(\mathcal{P}) - \frac{1}{2}H_S(\mathcal{Q}), \tag{10}$$

where $H_S(\mathcal{P}) = \sum_{a=1}^{A} p_a \log p_a$ is the Shannon entropy of distribution $\mathcal{P}$. D_{JS} is always well defined, symmetric, negative definite and bounded, i.e., $0 \leq D_{JS} \leq 1$.

The quantum Jensen-Shannon divergence has recently been developed as a generalization of the classical Jensen-Shannon divergence to quantum states by Lamberti et al. [17]. Given two density operators ρ and σ, the quantum Jensen-Shannon divergence between them is defined as

$$D_{QJS}(\rho, \sigma) = H_N\Big(\frac{\rho+\sigma}{2}\Big) - \frac{1}{2}H_N(\rho) - \frac{1}{2}H_N(\sigma). \tag{11}$$

D_{QJS} is always well defined, symmetric, negative definite and bounded, i.e., $0 \leq D_{QJS} \leq 1$ [17].

3.2 A Quantum Kernel Using the Discrete-time Quantum Walk

We propose a novel quantum Jensen-Shannon kernel for graphs using the quantum Jensen-Shannon divergence associated with the discrete-time quantum walk. Given a set of graphs $\{G_1, \ldots, G_a, \ldots, G_b, \ldots, G_N\}$, we simulate a discrete-time quantum walks on each $G_a(V_a, E_a)$ and $G_b(V_b, E_b)$ for $t = 0, 1, \ldots, T)$. Then, the density matrices $\rho^S_{G;a}$ and $\sigma^T_{G;b}$ of $G_a(V_a, E_a)$ and $G_b(V_b, E_b)$ can be computed using Eq.(5). With the density matrices $\rho^T_{G;a}$ and $\sigma^T_{G;b}$ to hand, the quantum Jensen-Shannon divergence $D_{QJS}(\rho_{G;a}, \sigma_{G;b})$ is computed as in Eq.(11). Finally, the quantum Jensen-Shannon kernel $k_{QJS}(G_a, G_b)$ between the pair of graphs $G_a(V_a, E_a)$ and $G_b(V_b, E_b)$ is defined as

$$\begin{aligned} k_{QJS}(G_a, G_b) &= \exp(-\alpha D_{QJS}(\rho^T_{G;a}, \sigma^T_{G;b})) \\ &= \exp\{-\alpha H_N\Big(\frac{\rho^T_{G;a} + \sigma^T_{G;b}}{2}\Big) + \alpha\frac{1}{2}H_N(\rho^T_{G;a}) + \alpha\frac{1}{2}H_N(\sigma^T_{G;b})\}. \end{aligned} \tag{12}$$

where $\frac{\rho^T_{G;a}+\sigma^T_{G;b}}{2}$ is a mixed state, α is a decay factor satisfying $0 \leq \alpha \leq 1$, and $H_N(\cdot)$ is the von Neumann entropy defined in Eq.(6). For simplification, in this work we set α as 1.

Lemma *The quantum Jensen-Shannon kernel k_{QJS} is positive definite **pd**.*

Proof This follows the definitions in [17, 18]. The quantum Jensen-Shannon divergence between a pair of density operators $\rho^T_{G;a}$ and $\sigma^T_{G;b}$ is symmetric and is a dissimilarity measure. Thus, the proposed quantum kernel k_{QJS} that is computed by exponentiating the negative divergence measure is **pd**. ■

For a pair of graphs, each of which has n vertices and m edges, the quantum kernel k_{QJS} requires time complexity $O(m^3)$. This is because the state space of the discrete-time quantum walk for a graph corresponds to the vertex set of its line graph. The number of the vertex in the line graph is double in the number of the edges of the original graph, i.e., the size of the unitary matrix or the density matrix for a graph is $m \times m$. The von Neumann entropy relies on the eigen decomposition of the density matrix, and thus requires time complexity $O(m^3)$. As a result, the whole time complexity of the kernel k_{QJS} is $O(m^3)$.

4 Experimental Evaluations

4.1 Graph Datasets

We explore our new kernel on five standard graph datasets from bioinformatics and computer vision. These datasets include: MUTAG, PPIs, PTC(MR), COIL5 and Shock. Some statistic concerning the datasets are given in Table 1.

MUTAG: The MUTAG dataset consists of graphs representing 188 chemical compounds, and aims to predict whether each compound possesses mutagenicity.
PPIs: The PPIs dataset consists of protein-protein interaction networks (PPIs). The graphs describe the interaction relationships between histidine kinase in different species of bacteria. There are 219 PPIs in this dataset and they are collected from 5 different kinds of bacteria. Here we select two kinds of bacteria, i.e., Proteobacteria40 PPIs and Acidobacteria46 PPIs, as the testing graphs.
PTC: The PTC (The Predictive Toxicology Challenge) dataset records the carcinogenicity of several hundred chemical compounds for male rats (MR), female rats (FR), male mice (MM) and female mice (FM). These graphs are very small (i.e., $20 - 30$ vertices), and sparse (i.e., $25 - 40$ edges. We select the graphs of male rats (MR) for evaluation. There are 344 test graphs in the MR class.
COIL5: We create a dataset referred to as COIL5 from the COIL image database. The COIL database consists of images of 100 3D objects. In our experiments, we use the images for the first five objects. For each of these objects we employ 72 images captured from different viewpoints. For each image we first extract corner points using the Harris detector, and then establish Delaunay graphs based on the corner points as vertices. Each vertex is used as the seed of a Voronoi region, which expands radially with a constant speed. The linear collision fronts of the regions delineate the image plane into polygons, and the Delaunay graph is the region adjacency graph for the Voronoi polygons.
Shock: The Shock dataset consists of graphs from the Shock 2D shape database. Each graph is a skeletal-based representation of the differential structure of the boundary of a 2D shape. There are 150 graphs divided into 10 classes.

4.2 Experiments on Standard Graph Datasets from Bioinformatics

Experimental Setup: We compare the performance of our new quantum Jensen-Shannon kernel (QJSD) with that of several alternative state-of-the-art graph kernels. These kernels include 1) the unaligned quantum Jensen-Shannon kernel (UQJS) associated with the continuous-time quantum walk [14], 2) the Weisfeiler-Lehman subtree kernel (WL) [7], 3) the shortest path graph kernel (SPGK) [4], 4) the Jensen-Shannon graph kernel associated with the steady state random walk (JSGK) [11], 5) the backtrackless random walk kernel using the Ihara zeta function based cycles (BRWK) [5], and 6) the random-walk graph kernel [3]. For our QJSD kernel, we let $T = 40$. In fact, as we let $T \geqslant 30$ we observe that the von Neumann entropy of the density matrices reaches an asymptote. While the optimal procedure would be that of selecting the value of T through cross-validation, the computational complexity of the kernel makes it unfeasible to do so. Moreover, previous work has shown that letting $T \to \infty$ allows us to achieve a good trade-off in terms of accuracy and computational effort [16]. For the Weisfeiler-Lehman subtree kernel, we set the dimension of the

Table 1. Information of the Graph based Datasets.

Datasets	MUTAG	PPIs	PTC	COIL5	Shock
Max # vertices	28	232	109	241	33
Min # vertices	10	3	2	72	4
Ave # vertices	17.93	109.60	25.60	144.90	13.16
# graphs	188	86	344	360	150
# classes	2	2	2	5	10

Table 2. Accuracy Comparisons (In % ± Standard Errors) on Graph Datasets.

Datasets	MUTAG	PPIs	PTC(MR)	COIL5	Shock
QJSD	83.16 ± .86	70.57 ± 1.20	**58.23** ± .80	69.78 ± .37	**44.86** ± .64
QJSU	82.72 ± .44	69.50 ± 1.20	56.70 ± .49	**70.11** ± .61	40.60 ± .92
WL	82.05 ± .57	**78.50** ± 1.40	56.05 ± .51	33.16 ± 1.01	36.40 ± 1.00
SPGK	**83.38** ± .81	61.12 ± 1.09	56.55 ± .53	69.66 ± .52	37.88 ± .93
JSGK	83.11 ± .80	57.87 ± 1.36	57.29 ± .41	69.13 ± .79	21.73 ± .76
BRWK	77.50 ± .75	53.50 ± 1.47	53.97 ± .31	14.63 ± .21	0.33 ± .37
RWGK	80.77 ± .72	55.00 ± .88	55.91 ± .37	20.80 ± .47	2.26 ± 1.01

Weisfeiler-Lehman isomorphism as 10. Based on the definition in [7], this means that we compute 10 different Weisfeiler-Lehman subtree kernel matrices (i.e., $k(1), k(2), \ldots, k(10)$) with different subtree heights $h(h = 1, 2, \ldots, 10)$, respectively. Note that, the WL and SPGK kernels are able to accommodate attributed graphs. In our experiments, we use the vertex degree as a vertex label for the WL and SPGK kernels.

Given these datasets and kernels, we perform a 10-fold cross-validation using a C-Support Vector Machine (C-SVM) to evaluate the classification accuracies of the different kernels. More specifically, we use the C-SVM implementation of LIBSVM. For each class, we use 90% of the samples for training and the remaining 10% for testing. The parameters of the C-SVMs are optimized separately for each dataset. We report the average classification accuracies (± standard error) of each kernel in Table 2. **Results:** Overall, in terms of classification accuracies our QJSK kernel outperforms or is competitive with the state-of-the-art kernels. In particular, the classification accuracies of our quantum kernel are significantly better than those of the graph kernels based on the classical random walk and the backtrackless random walk, over all the datasets. This suggests that our kernel using discrete-time quantum walks has better ability of capturing the graph characteristics. With respect to the quantum Jensen-Shannon kernel based on the continuous-time quantum walk, we observe a significant improvement on the PTC and Shock datasets. This is because the discrete-time quantum walk can be seen as a walk on line graphs, and reflects richer graph characteristics than the continuous-time quantum walk on the original graphs.

5 Conclusion

In this paper, we develop a new quantum Jensen-Shannon kernel for graphs by using the quantum Jensen-Shannon divergence and the discrete-time quantum walk. Our new quantum kernel can reflect richer graph characteristics than our previous quantum Jensen-Shannon kernel using the continuous-time quantum walk. Experiments demonstrate the effectiveness of the new quantum kernel.

Acknowledgments

Lu Bai and Zhihong Zhang are supported by National Natural Science Foundation of China (Grant No.61402389). Edwin R. Hancock is supported by a Royal Society Wolfson Research Merit Award.

References

1. Schölkopf, B., Smola, A.: Learning with Kernels. MIT Press (2002)
2. Haussler, D.: Convolution kernels on discrete structures. In: Technical Report UCS-CRL-99-10, Santa Cruz, CA, USA (1999)
3. Kashima, H., Tsuda, K., Inokuchi, A.: Marginalized kernels between labeled graphs. In: Proceedings of ICML. (2003) 321–328
4. Borgwardt, K.M., Kriegel, H.P.: Shortest-path kernels on graphs. In: Proceedings of the IEEE International Conference on Data Mining. (2005) 74–81
5. Aziz, F., Wilson, R.C., Hancock, E.R.: Backtrackless walks on a graph. IEEE Transactions on Neural Networks and Learning Systems **24** (2013) 977–989
6. Ren, P., Wilson, R.C., Hancock, E.R.: Graph characterization via ihara coefficients. IEEE Transactions on Neural Networks **22** (2011) 233–245
7. Shervashidze, N., Schweitzer, P., van Leeuwen, E.J., Mehlhorn, K., Borgwardt, K.M.: Weisfeiler-lehman graph kernels. Journal of Machine Learning Research **1** (2010) 1–48
8. Harchaoui, Z., Bach, F.: Image classification with segmentation graph kernels. In: Proceedings of CVPR. (2007)
9. Bach, F.R.: Graph kernels between point clouds. In: Proceedings of ICML. (2008) 25–32
10. Bai, L., Ren, P., Hancock, E.R.: A hypergraph kernel from isomorphism tests. In: Proceedings of ICPR. (2014) 3880–3885
11. Bai, L., Hancock, E.R.: Graph kernels from the jensen-shannon divergence. Journal of Mathematical Imaging and Vision **47** (2013) 60–69
12. Bai, L., Hancock, E.R.: Graph clustering using the jensen-shannon kernel. In: Proceedings of CAIP. (2011) 394–401
13. Bai, L., Hancock, E.R., Ren, P.: Jensen-shannon graph kernel using information functionals. In: Proceedings of ICPR. (2012) 2877–2880
14. Bai, L., Rossi, L., Torsello, A., Hancock, E.R.: A quantum jensen-shannon graph kernel for unattributed graphs. Pattern Recognition **48** (2015) 344–355
15. Bai, L., Hancock, E.R., Torsello, A., Rossi, L.: A quantum jensen-shannon graph kernel using the continuous-time quantum walk. In: Proceedings of GbRPR. (2013) 121–131
16. Rossi, L., Torsello, A., Hancock, E.R.: A continuous-time quantum walk kernel for unattributed graphs. In: Proceedings of GbRPR. (2013) 101–110
17. Lamberti, P., Majtey, A., Borras, A., Casas, M., Plastino, A.: Metric character of the quantum jensen-shannon divergence. Physical Review A **77** (2008) 052311
18. Majtey, A., Lamberti, P., Prato, D.: Jensen-shannon divergence as a measure of distinguishability between mixed quantum states. Physical Review A **72** (2005) 052310
19. Farhi, E., Gutmann, S.: Quantum computation and decision trees. Physical Review A **58** (1998) 915
20. Ren, P., Aleksic, T., Emms, D., Wilson, R.C., Hancock, E.R.: Quantum walks, ihara zeta functions and cospectrality in regular graphs. Quantum Information Process **10** (2011) 405–417
21. Emms, D., Severini, S., Wilson, R.C., Hancock, E.R.: Coined quantum walks lift the cospectrality of graphs and trees. Pattern Recognition **42** (2009) 1988–2002
22. L, G.: A fast quantum mechanical algorithm for database search. In: Proceedings of ACM Symposium on the Theory of Computation. (1996) 212–219
23. Nielsen, M., Chuang, I.: Quantum computation and quantum information. Cambridge university press (2010)