Proceedings Paper:
Bai, Lu, Rossi, Luca, Cui, Lixin et al. (1 more author) (2017) A Nested Alignment Graph Kernel through the Dynamic Time Warping Framework. In: Graph-Based Representations in Pattern Recognition - 11th IAPR-TC-15 International Workshop, GbRPR 2017, Anacapri, Italy, May 16-18, 2017, Proceedings. Springer , pp. 59-69.

https://doi.org/10.1007/978-3-319-58961-9_6

A Nested Alignment Graph Kernel through the Dynamic Time Warping Framework*

Lu Bai[1], Luca Rossi[2], Lixin Cui[1], and Edwin R. Hancock[3]

[1]Central University of Finance and Economics, Beijing, China
[2]Aston University, Birmingham, UK
[3]University of York, York, UK

Abstract. In this paper, we propose a novel nested alignment graph kernel drawing on depth-based complexity traces and the dynamic time warping framework. Specifically, for a pair of graphs, we commence by computing the depth-based complexity traces rooted at the centroid vertices. The resulting kernel for the graphs is defined by measuring the global alignment kernel, which is developed through the dynamic time warping framework, between the complexity traces. We show that the proposed kernel simultaneously considers the local and global graph characteristics in terms of the complexity traces, but also provides richer statistic measures by incorporating the whole spectrum of alignment costs between these traces. Our experiments demonstrate the effectiveness and efficiency of the proposed kernel.

1 Introduction

In pattern recognition, graph kernels are powerful tools for applying standard machine learning techniques to graph datasets [24]. These kernels are typically used in conjuction with kernel methods such as Support Vector Machines (SVM) and kernel Principle Component Analysis (kPCA) for the purposes of classification or clustering [4, 21].

The idea underpinning most existing graph kernels is that of decomposing graphs into substructures and comparing pairs of specific isomorphic substructures. Some examples are graph kernels based on counting pairs of isomorphic a) walks [27], b) paths [1], and c) restricted subgraph or subtree substructures [14]. Other examples include the work of Bach [2], who proposed a family of kernels for comparing point clouds. These kernels are based on a local tree-walk kernel between subtrees, which is defined by a factorization on suitably defined graphical models of the subtrees. Wang and Sahbi [28], on the other hand, defined a graph kernel for action recognition. They first describe actions in the videos using directed acyclic graphs (DAGs). The resulting kernel is defined as an extending random walk kernel by counting the number of isomorphic walks of DAGs. Harchaoui and Bath [18] proposed a segmentation graph kernel for images by counting the inexact isomorphic subtree patterns between image segmentation graphs. Other state-of-the-art graph kernels include the subtree-based hypergraph kernel [7], the Lovász graph kernel [19], the aligned subgraph kernel [10], the

* Corresponding Authors: Lixin Cui (cuilixin@cufe.edu.cn)

subgraph matching kernel [21], the fast depth-based subgraph kernel [6], the optimal assignment kernel [22], and the aligned Jensen-Shannon subgraph kernel [11].

Unfortunately, all the aforementioned graph kernels tend to capture only local characteristics of graphs, since they usually use substructures of limited sizes. As a result, these kernels may fail to reflect global graph characteristics. To overcome this shortcoming, Johansson et al. [19] developed a family of global graph kernels using geometric embeddings. Specifically, they use the Lovász number and its associated orthonormal representation to capture global graph characteristics. Bai et al. and Rossi et al. [4, 9, 26, 25] developed a family of graph kernels based on the classical Jensen-Shannon divergence, as well as its quantum analogue. Specifically, they use either the classical or the quantum walk together with quantum information theoretical measures to probe the global structure of the graph.

The aim of this work is to overcome the gap between local kernels (i.e., kernels based on local substructures of limited sizes) and the global kernels (i.e., global kernels and quantum or classical Jensen-Shannon kernels), by proposing a novel nested alignment kernel for graphs based on their depth-based complexity traces [5] and the dynamic time warping framework [15]. For a pair of graphs, we commence by computing the depth-based complexity traces rooted at the centroid vertices. The resulting kernel is defined by measuring the global alignment kernel [15] between the complexity traces. Recall that the depth-based complexity trace of a graph is based on a family of expansion subgraphs that form a nested sequence which gradually expands from the centroid vertex to the global graph structure. As a consequence, this sequence of subgraphs can reflect both local and global structure information of a graph. Furthermore, we show that the associated global alignment kernel encapsulates the whole spectrum of the alignment cost between the complexity traces. As a result, the proposed kernel can not only simultaneously consider both local and global graph characteristics in terms of the nested depth-based complexity traces, but also provide richer statistic measures by incorporating the whole spectrum of alignment costs between these traces. Experiments demonstrate the effectiveness and efficiency of the proposed kernel.

The remainder of this paper is organized as follows. Section 2 reviews the preliminary concepts that will be used in this work. Specifically, we introduce the global alignment kernel through the dynamic time warping framework and the depth-based complexity trace. Section 3 defines the proposed nested alignment kernel. Section 4 provides the experimental evaluation. Section 6 concludes this work.

2 Preliminary Concepts

In this section, we review some preliminary concepts that will be used in this work. We commence by reviewing the dynamic time warping framework. Specifically, we introduce the global alignment kernel based on this framework. Finally, we review the concept of depth-based complexity trace of a graph.

2.1 Global Alignment Kernels from the Dynamic Time Warping Framework

In this subsection, we review the global alignment kernel based on the dynamic time warping framework proposed in [15]. Let $\mathbf{T}$ be a set of discrete time series that take

values in a space $\mathcal{X}$. For a pair of discrete time series $\mathbf{P} = (p_1, \ldots, p_m) \in \mathbf{T}$ and $\mathbf{Q} = (q_1, \ldots, q_n) \in \mathbf{T}$ with lengths m and n respectively, the alignment π between $\mathbf{P}$ and $\mathbf{Q}$ is defined as a pair of increasing integral vectors (π_p, π_q) of length $l \leq m+n-1$, where

$$1 = \pi_p(1) \leq \cdots \leq \pi_p(l) = m$$

and

$$1 = \pi_q(1) \leq \cdots \leq \pi_q(l) = n$$

such that (π_p, π_q) is defined to have unitary increments and no simultaneous repetitions. For any index $1 \leq i \leq l-1$, the increment vector of $\pi = (\pi_p, \pi_q)$ satisfies

$$\begin{pmatrix} \pi_p(i+1) - \pi_p(i) \\ \pi_q(i+1) - \pi_q(i) \end{pmatrix} \in \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}. \tag{1}$$

In the dynamic time warping framework [15], the coordinates π_p and π_q of the alignment π define the warping function. Let $\mathcal{A}(m, n)$ be the set of all possible alignments between $\mathbf{P}$ and $\mathbf{Q}$. The dynamic time warping distance between $\mathbf{P}$ and $\mathbf{Q}$ is defined as

$$\mathrm{DTW}(\mathbf{P}, \mathbf{Q}) = \min_{\pi \in \mathcal{A}(m,n)} D_{\mathbf{P},\mathbf{Q}}(\pi), \tag{2}$$

where the cost

$$D_{\mathbf{P},\mathbf{Q}}(\pi) = \sum_{i=1}^{|\pi|} \varphi(p_{\pi_p(i)}, q_{\pi_q(i)}), \tag{3}$$

is defined by a local divergence φ that measures the discrepancy between any pair of elements $p_i \in \mathbf{P}$ and $q_i \in \mathbf{Q}$. Generally, φ can be defined as the squared Euclidean distance, i.e., $\varphi(p, q) = \|p - q\|^2$.

Based on the dynamic time warping distance defined in Eq.(2), a dynamic time warping kernel k_{DTW} [17] between $\mathbf{P}$ and $\mathbf{Q}$ can be defined as

$$k_{\mathrm{DTW}}(\mathbf{P}, \mathbf{Q}) = e^{-\mathrm{DTW}(\mathbf{P},\mathbf{Q})}. \tag{4}$$

Unfortunately, this kernel is not positive definite. This is because the optimal alignment required by the dynamic time warping cannot guarantee transitivity. To overcome the shortcoming, Cuturi [15] considers all possible alignments in $\mathcal{A}(m, n)$ and proposes another dynamic time warping inspired kernel, i.e., the global alignment kernel, as

$$k_{\mathrm{GA}}(\mathbf{P}, \mathbf{Q}) = \sum_{\pi \in \mathcal{A}(m,n)} e^{-D_{\mathbf{P},\mathbf{Q}}(\pi)}, \tag{5}$$

where k_{GA} is positive definite, since it quantifies the quality of both the optimal alignment and all other alignments $\pi \in \mathcal{A}(m, n)$. The kernel k_{GA} elaborates on the dynamic time warping distance by considering the same set of elementary operations [16]. However k_{GA} not only generalizes the dynamic time warping kernel k_{DTW}, but also provides richer statistic measures by incorporating the whole spectrum of alignment costs $\{D_{\mathbf{P},\mathbf{Q}}(\pi), \pi \in \mathcal{A}(m, n)\}$.

Intuitively, the global alignment kernel k_{GA} allows one to define a new graph kernel, by measuring the warping alignment π between any types of graph characteristic

sequences (or graph embedding vectors [13])) that have certain element orders with increasing structural variables, e.g, the depth-based complexity traces [5] from expansion subgraphs of increasing sizes, or cycle characteristics with increasing lengths identified from the Ihara zeta function [23].

2.2 Centroid Depth-based Complexity Traces

We review the concept of the depth-based complexity trace of a graph rooted at the centroid vertex [5]. Let $G(V,E)$ be an undirected graph with vertex set V and edge set E. Based on Dijkstra's algorithm, we commence by computing the shortest path matrix S_G, where each element $S_G(v,u)$ of S_G represents the length of the shortest path between vertices $v \in V$ and $u \in V$. For each vertex $v \in V$, let $S(v)$ be the average length of the shortest paths from v to the remaining vertices, i.e.,

$$S(v) = \frac{1}{|V|}\sum_{u\in V} S_G(v,u). \tag{6}$$

As discussed in [5], the centroid vertex $\hat{v}_C$ of $G(V,E)$ can be identified by selecting the vertex that has the minimum variance of shortest path lengths to the remaining vertices, i.e., the index of $\hat{v}_C$ is

$$\hat{v}_C = \arg\min_{v} \sum_{u\in V} [S_G(v,u) - S_V(v)]^2. \tag{7}$$

Let $N_{\hat{v}_C}^K$ be a vertex subset of $G(V,E)$ satisfying

$$N_{\hat{v}_C}^K = \{u \in V \mid S_G(\hat{v}_C,u) \leq K\}. \tag{8}$$

For $G(V,E)$ and its centroid vertex $\hat{v}_C$, we construct a family of K-layer expansion subgraphs $\mathcal{G}_K(\mathcal{V}_K;\mathcal{E}_K)$ as

$$\begin{cases} \mathcal{V}_K &= \{u \in N_{\hat{v}_C}^K\}; \\ \mathcal{E}_K &= \{(u,v) \subset N_{\hat{v}_C}^K \times N_{\hat{v}_C}^K \mid (u,v) \in E\}. \end{cases} \tag{9}$$

Note that the number expansion subgraphs is equal to the greatest length L of the shortest paths from the centroid vertex to the remaining vertices of $G(V,E)$. Moreover, the L-layer expansion subgraph is the graph $G(V,E)$ itself. An example of constructing a K-layer subgraph is shown in Fig.1.

Definition (Depth-based complexity traces): For a sample undirected graph $G(V,E)$, let $\{\mathcal{G}_1,\cdots,\mathcal{G}_K,\cdots,\mathcal{G}_L\}$ be the family of K-layer expansion subgraphs rooted at the centroid vertex of $G(V,E)$. Then the depth-based complexity trace $DB(G)$ of $G(V,E)$ is computed by measuring the entropies of the subgraphs [5], i.e.,

$$\mathrm{DB}(G) = \{H_S(\mathcal{G}_1),\cdots,H_S(\mathcal{G}_K),\cdots,H_S(\mathcal{G}_L)\}, \tag{10}$$

where $\cdots, H_S(\mathcal{G}_K)$ is the Shannon entropy associated with the steady state random walk on the K-layer centroid expansion subgraph $\mathcal{G}_K$ [4]. □

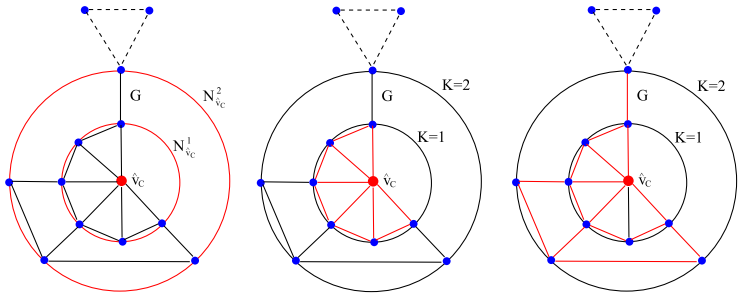

Fig. 1. The left-most figure shows the determination of K-layer centroid expansion subgraphs for a graph $G(V,E)$ which hold $|N^1_{\hat{v}_C}| = 6$ and $|N^2_{\hat{v}_C}| = 10$ vertices. While the middle and the right-most figure show the corresponding 1-layer and 2-layer subgraphs regarding the centroid vertex $\hat{v}_C$, and are depicted by red-colored edges. In this example, the vertices of different K-layer subgraphs regarding the centroid vertex $\hat{v}_C$ are calculated by Eq.(7), and pairwise vertices possess the same connection information in the original graph $G(V,E)$.

The depth-based complexity trace has a number of interesting properties [5]. First, it encapsulates the entropy-based information content flow through the family of K-layer expansion subgraphs rooted at the centroid vertex, and thus reflects rich intrinsic depth topology information of a graph. Second, it can be efficiently computed also on large graphs. This is because it is computed on a small set of expansion subgraphs rooted at the centroid vertex, and the computational complexity is polynomial. Furthermore, based on Eq.(9), we can also observe that the family of K-layer expansion subgraphs rooted at the centroid vertex $\hat{v}_C$ of the graph G constructs a nested sequence. This is because the family of the expansion subgraphs satisfies

$$\hat{v}_C \in \mathcal{G}_1 \cdots \subseteq \mathcal{G}_K \subseteq \cdots \subseteq \mathcal{G}_L \subseteq G.$$

In other words, it represents a sequence of subgraphs that gradually expand from the centroid vertex to the global graph. As a result of it nested nature, the depth-based complexity trace can reflecs both the local and global structure information of a graph. In summary, the depth-based complexity trace provides an elegant way of developing novel fast graph kernels that simultaneously consider local and global graph structures.

3 The Proposed Kernel

In this section, we introduce a novel nested alignment graph kernel through the dynamic time warping framework and the depth-based complexity trace.

3.1 A Nest Aligned Kernel from the Dynamic Time Warping Framework

Let $G_P(V_P,E_P)$ and $G_Q(V_Q,E_Q)$ be a pair of graphs, from a graph set $\mathbf{G}$. We commence by computing the depth-based complexity traces of G_P and G_Q as

$$\mathrm{DB}(G_P) = \{H_S(\mathcal{G}_{P;1}), \cdots, H_S(\mathcal{G}_{P;K}), \cdots, H_S(\mathcal{G}_{P;L^{\max}})\}$$

and

$$\mathrm{DB}(G_Q) = \{H_S(\mathcal{G}_{Q;1}), \cdots, H_S(\mathcal{G}_{Q;K}), \cdots, H_S(\mathcal{G}_{Q;L^{\max}})\},$$

respectively. Here $\mathcal{G}_{P;K}$ and $\mathcal{G}_{Q;K}$ are the K-layer expansion subgraphs rooted at the centroid vertices of G_P and G_Q, and $L^{\max}$ is the greatest length of the shortest paths rooted at the centroid vertices over all graphs in $\mathbf{G}$. Note that, for G_P and G_Q and the greatest lengths M and N of the shortest paths rooted at their centroid vertices, if $K \geq M$ and $K \geq M$ their K-layer expansion subgraphs are themselves, i.e., their global structures. Based on the global alignment kernel defined in Section 2.1, we develop a new nested alignment graph kernel k_{NA} between G_P and G_Q as

$$\begin{aligned} k_{\mathrm{NA}}(G_P, G_Q) &= k_{\mathrm{GA}}(\mathrm{DB}(G_P), \mathrm{DB}(G_Q)) \\ &= \sum_{\pi \in \mathcal{A}(L^{\max}, L^{\max})} e^{-D_{\mathbf{P},\mathbf{Q}}(\pi)}, \end{aligned} \tag{11}$$

where π denotes the warping alignment between $\mathrm{DB}(G_P)$ and $\mathrm{DB}(G_Q)$, $\mathcal{A}(L^{\max}, L^{\max})$ denotes all possible alignments, and $D_{\mathbf{P},\mathbf{Q}}(\pi)$ is the alignment cost defined in Eq.(3). Note that we cannot prove that the the proposed kernel k_{NA} is positive definite. Although our kernel is based on the global alignment kernel k_{GA}, which is a positive definite kernel, the time series compared by k_{NA} are not defined over the same underlying space but on two different graphs. Future work will explore the possibility of creating a positive definite kernel by computing the depth-based complexity traces over a common structure obtained by combining the input graphs.

As we have observed, the depth-based complexity trace reflects the nested entropy-based information and thus simultaneously considers the local and global graph structures. Furthermore, the proposed kernel $k_{\mathrm{NA}}(G_P, G_Q)$ is based on all possible warping alignments between depth-based complexity traces of the input graphs. As a result, $k_{\mathrm{NA}}(G_P, G_Q)$ can simultaneously capture richer local and global graph characteristics in terms of all possible alignments between the nested depth-based complexity traces.

3.2 Computational Analysis

For a pair of graphs both having n vertices, computing the nested alignment kernel k_{GA} has time complexity $O(n^3)$. This is because computing the depth-based complexity trace of a graph relies on the computation of the shortest path matrix and thus has time complexity $O(n^3)$. Furthermore, computing all possible alignments between the depth-based complexity traces has time complexity $O((L^{max})^2)$, where L^{max} is the greatest length of the shortest paths rooted at the centroid vertices of the two graphs and is lower than the vertex number n. As a result, the proposed kernel k_{GA} has polynomial time complexity $O(n^3)$.

4 Experimental Evaluations

4.1 Graph Datasets

We evaluate our kernels on standard graph datasets. These datasets include: MUTAG, PTC, COIL5, Shock and CATH2. Details of these datasets are shown in Table 1.

Table 1. Information on the selected graph based bioninformatics datasets

Datasets	MUTAG	PTC	COIL	Shock	CATH2
Max # vertices	28	109	241	33	568
Min # vertices	10	2	72	4	143
Mean # vertices	17.93	25.60	144.90	109.63	308.03
# graphs	188	344	360	150	190
# classes	2	2	5	5	2

MUTAG: The MUTAG dataset consists of graphs representing 188 chemical compounds labeled according to whether or not they affect the frequency of genetic mutations in the bacterium *Salmonella typhimuriums* and aims to predict whether each compound is associated with mutagenicity.
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 sparsem, 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: The COIL5 dataset is abstracted 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.
CATH2: The CATH2 dataset is harder to classify, since the proteins in the same topology class are structurally similar. The protein graphs are 10 times larger in size than chemical compounds, with 200 . 300 vertices. There is 190 testing graphs in the dataset.

5 Experiments on Standard Graph Datasets

We evaluate the performance of the nested alignment graph kernel (NAGK) on a number of graph classification tasks. Furthermore, we also compare our kernel with three state-of-the-art kernels, including 1) the Jensen-Shannon graph kernel (JSGK) [4], 2) the random walk graph kernel (RWGK) [20], 3) the unaligned quantum Jensen-Shannon graph kernel (QJSK) [9], and 4) the Lovász graph kernel (LGK) [19].

We compute the kernel matrix associated with each kernel on each dataset. We perform 10-fold cross-validation using a C-Support Vector Machine (C-SVM) to compute the classification accuracies, using LIBSVM software library [12]. We use nine samples for training and one for testing. The parameters of the C-SVMs are optimized on each training set using cross-validation. We report the average classification accuracy

and the runtime for each kernel in Table 2 and Table 3. The runtime is measured under Matlab R2015a running on a 2.5GHz Intel 2-Core processor (i.e., i5-3210m).

Table 2. Classification Accuracy (In $\%$ $\pm$ Standard Error) Runtime in Second.

Datasets	MUTAG	PTC	COIL5	Shock	CATH2
NAGK	**84.22** $\pm$ .50	**58.00** $\pm$ .64	69.75 $\pm$ .65	37.60 $\pm$ .62	**74.00** $\pm$ .83
JSGK	83.11 $\pm$ .80	57.29 $\pm$ .41	69.13 $\pm$ .79	21.73 $\pm$ .76	72.26 $\pm$ .76
RWGK	80.77 $\pm$ .75	53.97 $\pm$ .31	14.21 $\pm$ .65	0.33 $\pm$ .37	–
QJSK	82.72 $\pm$ .44	56.70 $\pm$ .49	**70.11** $\pm$ .61	**40.60** $\pm$ .92	71.11 $\pm$ .88
LGK	80.83 $\pm$ .43	56.29 $\pm$ .47	–	31.80 $\pm$ .89	–

Table 3. Runtime for Various Kernels.

Datasets	MUTAG	PTC	COIL5	Shock	CATH2
NAGK	$8.6 \cdot 10^2$	$2.3 \cdot 10^3$	$3.3 \cdot 10^3$	$3.8 \cdot 10^2$	$9.4 \cdot 10^2$
JSGK	$1.0 \cdot 10^0$	$1.0 \cdot 10^0$	$1.0 \cdot 10^0$	$1.0 \cdot 10^0$	$1.0 \cdot 10^0$
RWGK	$4.6 \cdot 10^1$	$6.7 \cdot 10^1$	$1.1 \cdot 10^3$	$2.3 \cdot 10^1$	–
QJSK	$2.0 \cdot 10^1$	$1.0 \cdot 10^2$	$1.0 \cdot 10^3$	$1.4 \cdot 10^1$	$4.4 \cdot 10^3$
LGK	$1.0 \cdot 10^3$	$7.4 \cdot 10^3$	–	$1.0 \cdot 10^3$	–

In terms of classification accuracy, Table 2 indicates that the proposed NAGK kernel can significantly outperform the alternative state-of-the-art graph kernels, excluding the QJSK kernel on the COIL5 and Shock datasets. However, the proposed NAGK kernel is still competitive to the QJSK kernel on the COIL5 dataset and outperforms the QJSK kernel on the MUTAG, PTC and CATH2 datasets. The reasons for this effectiveness are twofold. First, as we have stated, the depth-based complexity traces used by the proposed NAGK kernel encapsulate nested entropy-based information that extend from the centroid vertex to the global graph structure. As a consequence, the proposed NAGK kernel can simultaneously consider the local and global graph characteristics. By contrast, the the QJSK and JSGK kernels can only reflect global graph characteristics, whereas the LGK and RWGK can only reflect local graph characteristics. Second, the proposed NAGK kernel is based on all possible alignments between the complexity traces, and thus reflects rich statistic measures by incorporating the whole spectrum of alignment costs. On the other hand, we observe that the QJSK kernel based on the global von Neumann entropy from the continuous-time quantum walk is the most competitive kernel to the proposed NAGK kernel, though the QJSK kernel can only reflect global characteristics. This is because the entropy measure from the quantum walk can reflect richer intrinsic topology information than that from the classical steady state random walk (for the proposed NAGK kernel). This in turn suggest the possibility of further extending the NAGK kernel using quantum walks to extract an analogous of the depth-based complexity trace used in this study.

In terms of runtime, the proposed the NAGK kernel is not the fastest kernel, when compared to the other graph kernels. However, we can observe that the proposed NAGK kernel can always complete the computation of the kernel matrices, unlike some alternative graph kernels (e.g., the LGK and RWGK kernels), which failed complete the computation in a reasonable time.

6 Conclusion

In this paper, we have proposed a novel nested alignment graph kernel. The kernel is an adaptation of the dynamic time warping framework based kernel (i.e., the global alignment kernel) to graphs. To this end, we made use of the depth-based complexity traces of graphs, a powerful and fast to compute graph descriptor. Unlike most existing graph kernels that only probe local or global graph characteristics, the proposed kernel simultaneously considers local and global graph characteristics and thus reflects the presence of richer structural patterns. The experiments have demonstrated the effectiveness and efficiency of the proposed kernel.

Our future work is to extend the proposed kernel to attributed graphs that encapsulate vertex and edge labels. Moreover, we would also like to further develop novel graph kernels through the dynamic time warping framework associated with other types of (hyper)graph characteristic sequences, e.g., the cycle numbers identified by the Ihara zeta function, the time-varying entropies computed from the continuous-time or discrete-time quantum walk [9, 8], and the depth-based hypergraph complexity traces [3]. Finally, we are also interested in developing novel graph kernels for time-varying financial market networks [29], using the dynamic time warping framework.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant no. 61503422 and 61602535), the Open Projects Program of National Laboratory of Pattern Recognition, and the program for innovation research in Central University of Finance and Economics.

References

1. Marco A. Alvarez, Xiaojun Qi, and Changhui Yan. A shortest-path graph kernel for estimating gene product semantic similarity. *J. Biomedical Semantics*, 2:3, 2011.
2. Francis R Bach. Graph kernels between point clouds. In *Proceedings of ICML*, pages 25–32, 2008.
3. Lu Bai, Francisco Escolano, and Edwin R. Hancock. Depth-based hypergraph complexity traces from directed line graphs. *Pattern Recognition*, 54:229–240, 2016.
4. Lu Bai and Edwin R Hancock. Graph kernels from the jensen-shannon divergence. *Journal of mathematical imaging and vision*, 47(1-2):60–69, 2013.
5. Lu Bai and Edwin R Hancock. Depth-based complexity traces of graphs. *Pattern Recognition*, 47(3):1172–1186, 2014.
6. Lu Bai and Edwin R. Hancock. Fast depth-based subgraph kernels for unattributed graphs. *Pattern Recognition*, 50:233–245, 2016.
7. Lu Bai, Peng Ren, and Edwin R. Hancock. A hypergraph kernel from isomorphism tests. In *Proceddings of ICPR*, pages 3880–3885, 2014.
8. Lu Bai, Luca Rossi, Lixin Cui, Zhihong Zhang, Peng Ren, Xiao Bai, and Edwin R. Hancock. Quantum kernels for unattributed graphs using discrete-time quantum walks. *Pattern Recognition Letters*, 87:96–103, 2017.
9. Lu Bai, Luca Rossi, Andrea Torsello, and Edwin R. Hancock. A quantum jensen-shannon graph kernel for unattributed graphs. *Pattern Recognition*, 48(2):344–355, 2015.

10. Lu Bai, Luca Rossi, Zhihong Zhang, and Edwin R. Hancock. An aligned subtree kernel for weighted graphs. In *Proceedings of ICML*, pages 30–39, 2015.
11. Lu Bai, Zhihong Zhang, Chaoyan Wang, Xiao Bai, and Edwin R. Hancock. A graph kernel based on the jensen-shannon representation alignment. In *Proceedings of IJCAI*, pages 3322–3328, 2015.
12. Chih-Chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):27, 2011.
13. Donatello Conte, Jean-Yves Ramel, Nicolas Sidere, Muhammad Muzzamil Luqman, Benoit Gaüzère, Jaume Gibert, Luc Brun, and Mario Vento. A comparison of explicit and implicit graph embedding methods for pattern recognition. In *Proceedings of GbRPR*, pages 81–90, 2013.
14. Fabrizio Costa and Kurt De Grave. Fast neighborhood subgraph pairwise distance kernel. In *Proceedings ICML*, pages 255–262, 2010.
15. Marco Cuturi. Fast global alignment kernels. In *Proceedings of ICML*, pages 929–936, 2011.
16. Marco Cuturi, Jean-Philippe Vert, Øystein Birkenes, and Tomoko Matsui. A kernel for time series based on global alignments. In *Proceedings of ICASSP*, pages 413–416, 2007.
17. Bernard Haasdonk and Claus Bahlmann. Learning with distance substitution kernels. In *Proceedings of DAGM*, pages 220–227, 2004.
18. Zaïd Harchaoui and Francis Bach. Image classification with segmentation graph kernels. In *Proceedings of CVPR*, pages 1–8, 2007.
19. Fredrik D. Johansson, Vinay Jethava, Devdatt P. Dubhashi, and Chiranjib Bhattacharyya. Global graph kernels using geometric embeddings. In *Proceedings of ICML*, pages 694–702, 2014.
20. Hisashi Kashima, Koji Tsuda, and Akihiro Inokuchi. Marginalized kernels between labeled graphs. In *Proceedings of ICML*, pages 321–328, 2003.
21. Nils Kriege and Petra Mutzel. Subgraph matching kernels for attributed graphs. In *Proceedings of ICML*, 2012.
22. Nils M. Kriege, Pierre-Louis Giscard, and Richard C. Wilson. On valid optimal assignment kernels and applications to graph classification. In *Proceedings of NIPS*, pages 1615–1623, 2016.
23. Peng Ren, Tatjana Aleksić, Richard C Wilson, and Edwin R Hancock. A polynomial characterization of hypergraphs using the ihara zeta function. *Pattern Recognition*, 44(9):1941–1957, 2011.
24. Kaspar Riesen and Horst Bunke. *Graph classification and clustering based on vector space embedding*. World Scientific Publishing Co., Inc., 2010.
25. Luca Rossi, Andrea Torsello, and Edwin R Hancock. Measuring graph similarity through continuous-time quantum walks and the quantum jensen-shannon divergence. *Physical Review E*, 91(2):022815, 2015.
26. Luca Rossi, Andrea Torsello, Edwin R Hancock, and Richard C Wilson. Characterizing graph symmetries through quantum jensen-shannon divergence. *Physical Review E*, 88(3):032806, 2013.
27. Matthew Urry and Peter Sollich. Random walk kernels and learning curves for gaussian process regression on random graphs. *Journal of Machine Learning Research*, 14(1):1801–1835, 2013.
28. Ling Wang and Hichem Sahbi. Directed acyclic graph kernels for action recognition. In *Proceedings of ICCV*, pages 3168–3175, 2013.
29. Cheng Ye, Cesar H. Comin, Thomas K. DM. Peron, Filipi N. Silva, Francisco A. Rodrigues, Luciano da F. Costa, Andrea Tosello, and Edwin R. Hancock. Thermodynamic characterization of networks using graph polynomials. *Physical Review E*, 92(3):032810, 2015.