39th International Conference on Machine Learning

A Hierarchical Transitive-Aligned Graph Kernel for Un-attributed Graphs

Lu Bai[1,2], Lixin Cui[2,*], Edwin R. Hancock[3]

[1] School of Artificial Intelligence, Beijing Normal University, Beijing, China

[2] Central University of Finance and Economics, Beijing, China

[3] Department of Computer Science, University of York, UK

* Corresponding Author: cuilixin@cufe.edu.cn

Outline

1 Background

2 Our Methods

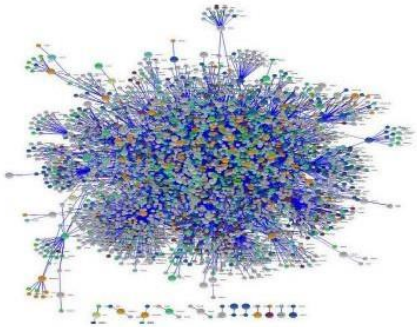
3 Experiments

Why Graphs ?

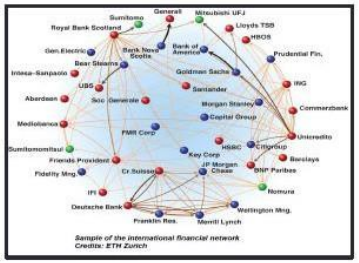
- **Graph-based representations are powerful tools to represent structure data that is described with pairwise relationships between components.**

Social Networks

Biology Networks



Finance Networks



Internet of Things

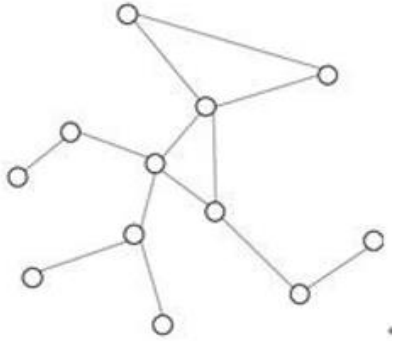
Information Networks

Logistic Networks

Graphs are Everywhere

Challenges and Methods

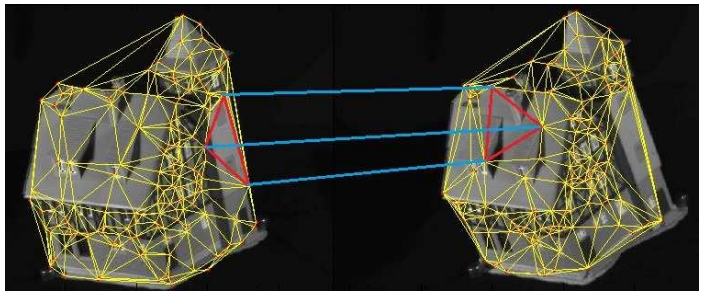
- **Challenges:** Graphs are typical non-Euclidean data, thus it is hard to learn effective numeric features for graphs



- **State-of-The-Art Methods:**
  - **Graph Embeddings:** Represent the characteristics of graphs in a low dimensional vector space
  - **Graph Kernels: Compute the similarity between graphs in a high dimensional Hilbert space, better preserve structure information**
  - **Graph Convolutional Networks:** Generalize the convolution operation of CNNs to graphs, provide end-to-end learning framework

Graph Kernels

- **R-convolution:** The widely used framework to define graph kernels, defined by decomposing graphs into subsructures and then measuring the isomorphism between them
- **Drawbacks of R-convolution Graph Kernels:** Ignore the correspondence information between substructures, thus cannot reflect precise similarity measures



- **Alignment or Matching Graph Kernels:** Integrate the vertex/edge correspondence information into the kernel computation, the correspondence is usually not transitive, thus are not Positive Definite (pd)

Outline

1 Background

2 Our Methods

3 Experiments

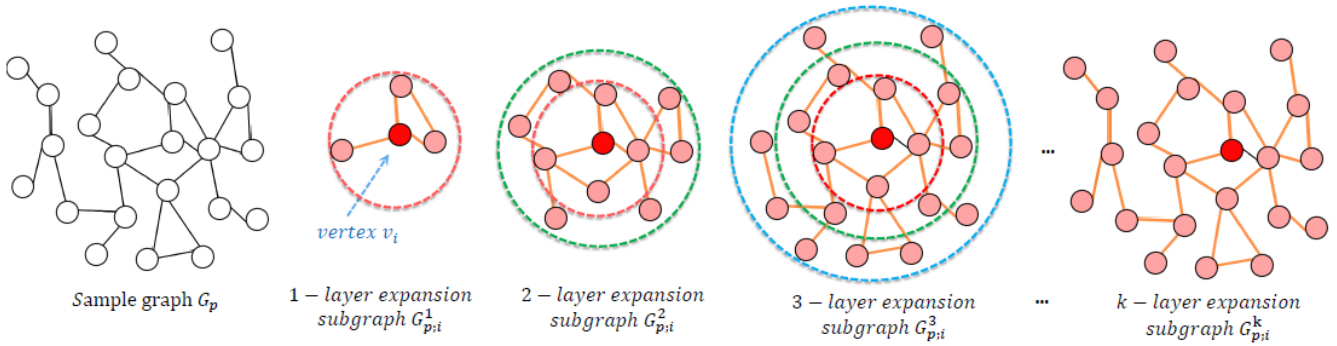
The Objective of This Work

- **Develop New Alignment Kernels:**
  - **Overcome the Drawback of Ignoring Correspondence Information**
  - **Guarantee the Transitivity between Aligned Vertices**
  - **Guarantee the Positive Definite**

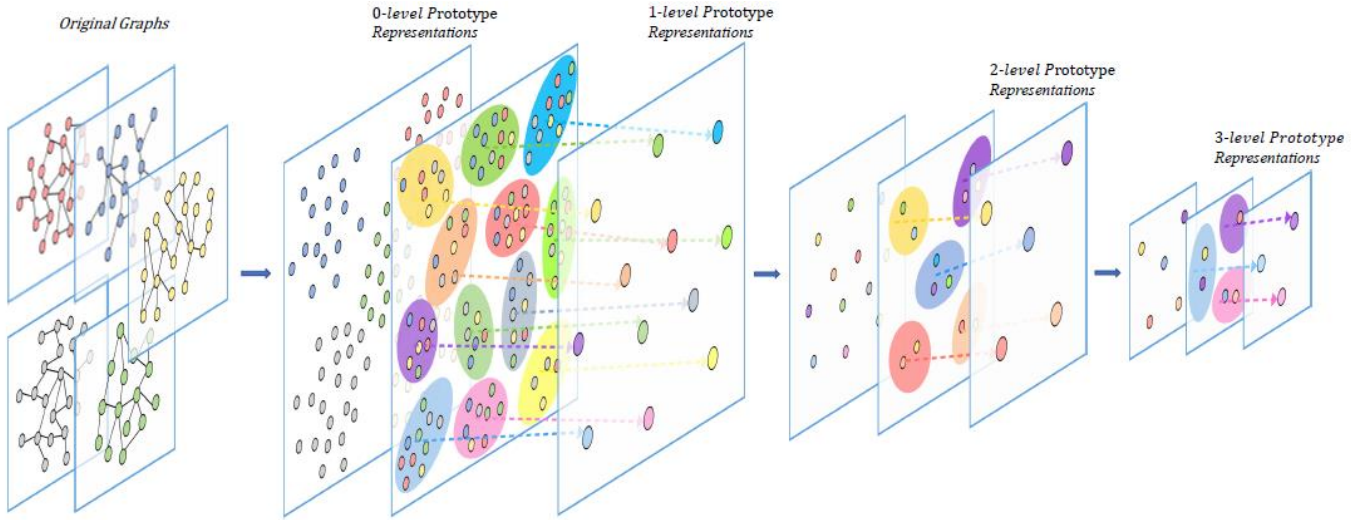
- **The Ideas:**
  - **Construct A Family of Hierarchical Prototype Representations**
  - **Align Each Individual Graph Structure to the Same Prototype Representation Sets of Different Levels**
  - **Define the Kernels by Counting the Number of the Hierarchically Aligned Vertex Pairs**

The Prototype Representations

- **Initialize The Vectorial Representations of Vertices:** Depth-based (DB) Complexity Traces of Vertices



- **Hierarchical Prototype Representations:** Hierarchically employ the k-means clustering method on the DB complexity traces



The Proposed Kernel

- **The Correspondence Matrix:** Record the correspondence information of each graph G_p to the k-level prototype representations

$$M_p^{h,k}(i,n) = \begin{cases} 1 & \text{if } R_p^{h,k}(i,n) \text{ is the smallest} \\ & \text{element in row } n, \text{ and } |\mathcal{S}_i^k| \neq 0; \\ 0 & \text{otherwise.} \end{cases}$$

- **The Proposed Kernel:** Counting the numbers of the transitive aligned vertex pairs between a pair of graphs G_p and G_q

$$k_{\mathrm{HTAK}}^{(H)}(G_p, G_q) = \sum_{h=1}^{H} \sum_{k=1}^{K} \sum_{i=1}^{|V_p|} \sum_{j=1}^{|V_q|} \mathcal{M}_{p;q}^{(h,k)}(i,j)$$

Outline

1 Background

2 Our Methods

3 Experiments

Classification Evaluations on Benchmark Datasets

The Benchmark Datasets:

Table 1. Information of the graph based computer vision (CV), bioinformatics (Bio), and social network (SN) datasets.

Datasets	BAR31	BSPHERE31	GEOD31	MUTAG	NCI1	CATH2	COLLAB	IMDB-B	IMDB-M
Max # vertices	220	227	380	28	111	568	492	136	89
Mean # vertices	95.42	99.83	57.42	17.93	29.87	308.03	74.49	19.77	13.00
# graphs	300	300	300	188	4110	190	5000	1000	1500
# classes	15	15	15	2	2	2	3	2	3
Description	CV	CV	CV	Bio	Bio	Bio	SN	SN	SN

Comparisons with Graph Kernels:

Table 2. Classification Accuracy (In % ± Standard Error) for Comparisons with Graph Kernels.

Datasets	BAR31	BSPHERE31	GEOD31	MUTAG	NCI1	CATH2	COLLAB	IMDB-B	IMDB-M
HTAK	71.00±.45	**62.90**±.65	**47.80**±.49	87.32±.60	79.01±.14	**87.89**±.71	**79.87**±.15	**72.89**±.56	50.23±.18
ASK	**73.10**±.67	60.30±.44	46.21±.69	87.50±.65	78.47±.12	78.52±.67	77.53±.31	70.38±.72	50.12±.51
WLSK	58.53±.53	42.10±.68	38.20±.68	82.88±.57	84.77±.13	67.36±.63	77.39±.35	71.88±.77	49.50±.49
SPGK	55.73±.44	48.20±.76	38.40±.65	83.38±.81	74.21±.30	81.89±.63	58.80±.20	71.26±1.04	**51.33**±.57
CORE SP	–	–	–	**88.29**±1.55	73.46±.32	–	–	72.62±.59	49.43±.42
GCGK	23.40±.60	18.80±.50	22.36±.55	82.04±.39	63.72±.12	73.68±1.09	–	–	–
JTQK	60.56±.35	46.93±.61	40.10±.46	85.50±.55	**85.32**±.14	68.70±.69	76.85±.40	72.45±.81	50.33±.49
PMGK	–	–	–	80.66±.90	72.27±.59	–	–	68.53±.61	45.75±.66
CORE PM	–	–	–	87.19±1.47	74.90±.45	–	–	71.04±.64	48.30±1.01
RetGK(MC)	–	–	–	–	–	–	73.60 ± .30	71.00±.60	46.70±.60

The symbol - means that some approaches were not evaluated by the original authors.

Comparisons with Deep Learning Methods:

Table 3. Classification Accuracy (In % ± Standard Error) for Comparisons with Deep Learning Methods.

Datasets	MUTAG	NCI1	COLLAB	IMDB-B	IMDB-M
HTAK	87.32±.60	**79.01**±.14	**79.87**±.15	**72.89**±.56	**50.23**±.18
DGCNN	85.83±1.66	74.44±.47	73.76±.49	70.03±.86	47.83±.85
PSGCNN	**88.95**±4.37	76.34±1.68	72.60±2.15	71.00±2.29	45.23±2.84
DCNN	66.98	56.61±1.04	52.11±.71	49.06±1.37	33.49±1.42
DGK	82.66±1.45	62.48±.25	73.09±.25	66.96±.56	44.55±.52

Thank you !

This work is supported by the National Natural Science Foundation of China under Grants T2122020, 61976235, and 61602535.