Joint Graph Embedding and Alignment with Spectral Pivot

Paris A. Karakasis*, Aritra Konar*, and Nicholas D. Sidiropoulos*

*Department of ECE
University of Virginia
USA

Table of Contents

Introduction

Problem Formulation

Algorithmic Approach

Experimental Evaluation

Conclusion & Future Work

Introduction

Problem Formulation

Algorithmic Approach

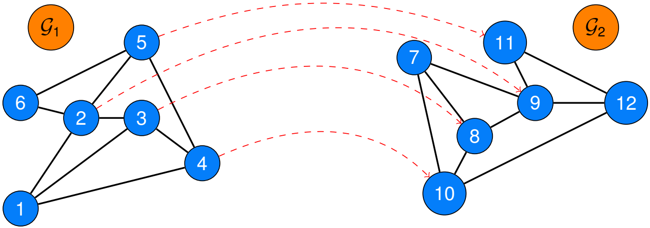
Experimental Evaluation

Conclusion & Future Work

Graph Alignment

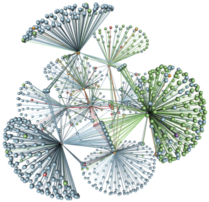
- Given $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$, find a bijection

$$f : \mathcal{V}_1 \to \mathcal{V}_2$$

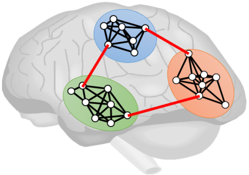


- Extensions: Directed & weighted graphs, labeled or attributed nodes, etc...
- Related problems: Subgraph matching, link prediction, and others...

Graph Alignment: Applications

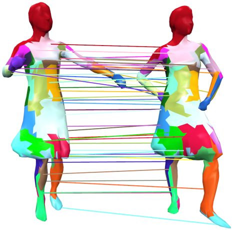


(a) PPI networks



(b) Brain networks

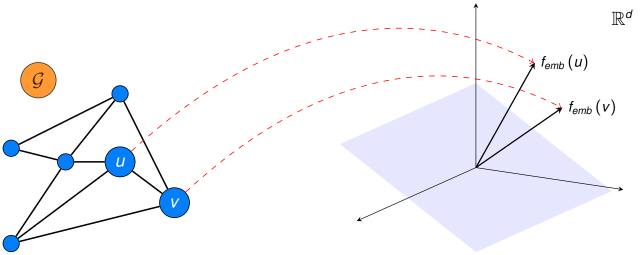
(c) Social networks



(d) Computer vision

Graph & Node Embeddings

- Find mappings of nodes to low-dimensional vector spaces



- Which mappings are useful?
- Graph Alignment: match nodes with similar embeddings!

Related Work

- Graph alignment as a QP problems ⇒ solved exactly or approximately
- Spectral methods ⇒ use the eigenvectors of the adjacency matrix
  - Umeyama's method, 1988
  - IsoRank, Singh et al. 2007
  - EigenAlign (EA) and LowRankAlign (LRA), Feizi et al. 2019
- Using node embeddings:
  - REGAL, Heimann et al. 2018
  - CONE-Align, Chencone et al. 2020
  - Gromov-Wasserstein discrepancy based learning, Xu et al. 2019

Introduction

Problem Formulation

Algorithmic Approach

Experimental Evaluation

Conclusion & Future Work

Problem Formulation of Graph Alignment

- Let $\mathcal{P}^n$ be the set of all $n \times n$ permutation matrices
- In terms of adjacency matrices $\mathbf{A}_1$ and $\mathbf{A}_2$

$$\min_{\mathbf{P}\in\mathcal{P}^n} \left\|\mathbf{A}_1 - \mathbf{P}\mathbf{A}_2\mathbf{P}^T\right\|_F^2 \Rightarrow \text{Quadratic Assignment Problem (QAP)}$$

$$\Rightarrow \text{NP-hard}$$

- In terms of node embeddings $\mathbf{E}_1$ and $\mathbf{E}_2$

$$\min_{\mathbf{P}\in\mathcal{P}^n} \|\mathbf{E}_1 - \mathbf{P}\mathbf{E}_2\|_F^2 \Rightarrow \text{Linear Assignment Problem (LAP)}$$

$$\Rightarrow \text{Hungarian algorithm } \mathcal{O}\left(n^3\right)$$

Using node embeddings for graph alignment has great potential!

Proposed Formulation for Graph Alignment

- Most approaches:

  Fix the nodes embeddings ($\mathbf{E}_1$ and $\mathbf{E}_2$) and solve the resulting LAP

- Proposed approach:
  - Consider embeddings of the form

$$\mathbf{E}_i = \mathbf{A}_i \mathbf{Q}_i \in \mathbf{R}^{n \times d}, \quad \text{for } i = 1, 2$$

  - Fix $\mathbf{E}_1 = \mathbf{A}_1 \mathbf{Q}_1 = d$ major principal components
  - Learn jointly $\mathbf{E}_2$ and $\mathbf{P}$ by solving

$$\begin{aligned} \min_{\mathbf{P},\ \mathbf{Q}_2} \quad & \|\mathbf{E}_1 - \mathbf{P}\mathbf{A}_2\mathbf{Q}_2\|_F^2 \quad + \quad \overbrace{\lambda \|\mathbf{Q}_1 - \mathbf{P}\mathbf{Q}_2\|_F^2}^{\text{isomorphic case} \Rightarrow \mathbf{Q}_1 = \mathbf{P}\mathbf{Q}_2} \quad , \\ \text{s.t.} \quad & \underbrace{\mathbf{P} \geq \mathbf{0}, \quad \mathbf{1}_n^T \mathbf{P} = \mathbf{1}_n^T,}_{\text{left stochasticity}} \; + \; \underbrace{\mathbf{P}^T\mathbf{P} = \mathbf{P}\mathbf{P}^T = \mathbf{I}_n}_{\text{orthonormality}} \quad (\equiv \mathbf{P} \in \mathcal{P}^n) \end{aligned} \tag{1}$$

Introduction

Problem Formulation

Algorithmic Approach

Experimental Evaluation

Conclusion & Future Work

Algorithmic Approach I

- Problem (1) is non-convex! The permutation constraints are tough!
- We propose the following "penalty" formulation

$$\min_{\mathbf{U},\,\mathbf{W},\,\mathbf{Q}_2} \;\; \|\mathbf{E}_1 - \mathbf{U}\mathbf{A}_2\mathbf{Q}_2\|_F^2 + \lambda\,\|\mathbf{Q}_1 - \mathbf{U}\mathbf{Q}_2\|_F^2 + \overbrace{\rho\|\mathbf{U}-\mathbf{W}\|_F^2}^{\rho\to\infty \;\Rightarrow\; \mathbf{U}=\mathbf{W}\,\left(=\mathbf{P}\in\mathcal{P}^n\right)} \quad , \qquad (2)$$

$$\text{s.t.} \qquad \mathbf{W} \geq \mathbf{0}, \quad \mathbf{1}_n^T\mathbf{W} = \mathbf{1}_n^T, \quad \mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I}_n,$$

- Block-separable problem $\Rightarrow$ Alternating Optimization!
- Given $\mathbf{U}^k, \mathbf{W}^k, \mathbf{Q}_2^k$, at the k-th iteration, we solve in a cyclic fashion:

$$\mathbf{Q}_2^{k+1} = \underset{\mathbf{Q}_2}{\text{argmin}} \left\|\mathbf{E}_1 - \mathbf{U}^k\mathbf{A}_2\mathbf{Q}_2\right\|_F^2 + \lambda\left\|\mathbf{Q}_1 - \mathbf{U}^k\mathbf{Q}_2\right\|_F^2,$$

$$\mathbf{W}^{k+1} = \underset{\mathbf{W}}{\text{argmin}} \left\|\mathbf{W} - \mathbf{U}^k\right\|_F^2, \;\; \text{s.t.} \;\; \mathbf{W} \geq \mathbf{0}, \;\; \mathbf{1}_n^T\mathbf{W} = \mathbf{1}_n^T,$$

$$\mathbf{U}^{k+1} \in \underset{\mathbf{U}:\mathbf{U}^T\mathbf{U}=\mathbf{U}\mathbf{U}^T=\mathbf{I}_n}{\text{argmin}} \left\|\mathbf{E}_1 - \mathbf{U}\mathbf{A}_2\mathbf{Q}_2^{k+1}\right\|_F^2 + \lambda\left\|\mathbf{Q}_1 - \mathbf{U}\mathbf{Q}_2^{k+1}\right\|_F^2 + \rho\left\|\mathbf{U} - \mathbf{W}^{k+1}\right\|_F^2$$

Algorithmic Approach II

- Convergence guarantees:
  i) the algorithm monotonically reduces the objective of (2)
  ii) every limit point of the proposed algorithm is a stationary point of (2)

- Algorithmic Complexity:
  a) Updating $\mathbf{Q}_2$: A special unconstrained least squares problem
     - Cholesky decomposition of the Gram matrix $\left(n^3/3 \text{ flops}\right)$ – **only once!**
     - Solving the linear system via forward-backward substitution $O(dn^2)$
  b) Updating $\mathbf{W}$: Euclidean projection of $\mathbf{U}^k$ onto left stochastic matrices $O(n^2 \log n)$
  c) Updating $\mathbf{U}$: An Orthogonal Procrustes problem $O(n^3)$

Overall complexity: $O(n^3)$

Introduction

Problem Formulation

Algorithmic Approach

Experimental Evaluation

Conclusion & Future Work

Experimental Evaluation

- Comparison against well known methods on real world graphs
  - **Umeyama's Method**, 1988: A spectral-embedding based method
  - **IsoRank**, Singh et al. 2017: A spectral method based on random walks
  - **Low-Rank Align**, Feizi et al. 2019: A spectral method for the QAP
  - **CONE-Align**, Chencone et al. 2020: Joint embedding and alignment
- Datasets: Real-world datasets from the KONECT Project and SNAP [1]

Network	number of vertices (n)	number of edges (m)	network type
C. ELEGANS	277	2,105	Interactome
ARENAS-EMAIL	1,133	5,451	Communications
POLBLOG	1,224	16,714	Social
AIRPORTS	1,574	17,215	Infrastructure
A. THALIANA	2,082	4,145	Interactome
JAPANESE BOOK	3,177	7,998	Word Adjacency
HOMOSAPIENS	3,890	38,292	Interactome
CA-GRQC	5,242	14,490	Co-authorship

[1] Kunegis, 2013 & Leskovec and Krevl, 2014, respectively

Experimental Setup I

- In our experiments we consider:
  i) a simple, undirected, unweighted graph as target graph $(\mathcal{G}_1)$
  ii) a “noisy" and permuted version of $\mathcal{G}_1$ as query graph $(\mathcal{G}_2)$
  iii) $\mathbb{E}\,[\text{extra edges in } \mathcal{G}_2]$ between 1% and 20% of $\#\mathcal{V}_1$
  iv) for each noise-level, averages over 20 Monte-Carlo runs
- Evaluation Metric:

$$\text{Edge Correctness} \;:= \frac{\text{number of edge overlaps induced by the algorithm}}{\text{the number of edges in } \mathcal{G}_1} \leq 1$$

Experimental Setup II

- Initialization: we use the output produced by CONE-Align
- Convergence criterion:

$$\frac{\left\|\mathbf{U}^{k}-\mathbf{U}^{k-1}\right\|_F}{\sqrt{n}} \leq 10^{(-2)} \text{ or number of iterations exceeds } K_{max} = 60$$

- Choice of parameters:
  i) the embedding dimension, $d \Rightarrow \nearrow d$ linearly with $n \rightarrow \nearrow$ performance
  ii) the level of non-isomorphism btw $\mathcal{G}_1$ & $\mathcal{G}_2$, $\lambda \geq 0 \Rightarrow$ trial-and-error
  iii) the penalty parameter ρ (violation of $\mathbf{U} = \mathbf{W}$) $\Rightarrow \rho \in [0.1, 0.3]$

Results I

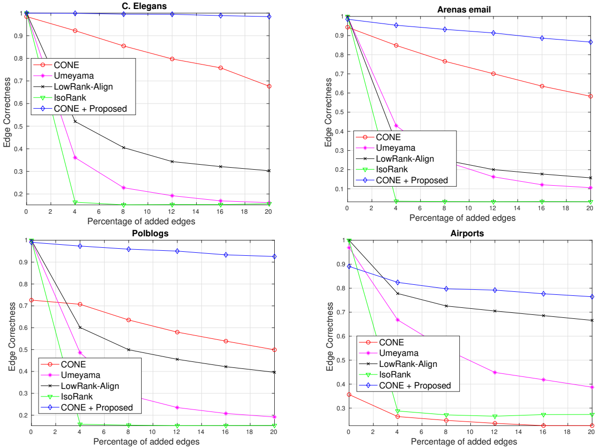


Figure: Edge correctness vs. noise level across networks. For each value of noise level (p_e), 20 different realizations of the graphs $\mathcal{G}_2$, with a certain percentage of additional edges and under a different and random permutation, were generated. The number of additional edges varied from 0 to 20% of the total number of edges of the fixed graph $\mathcal{G}_1$.

Results II

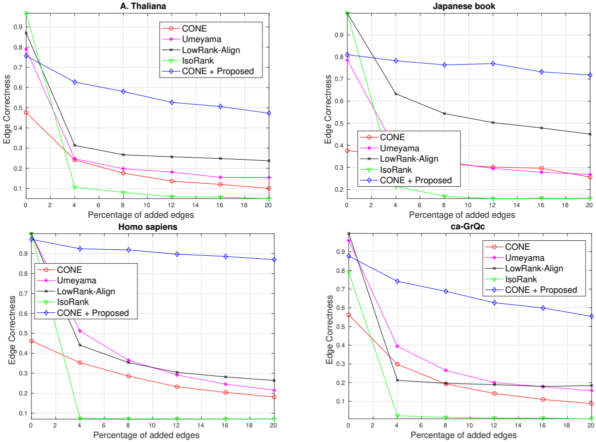


Figure: Edge correctness vs. noise level across networks. For each value of noise level (p_e), 20 different realizations of the graphs $\mathcal{G}_2$, with a certain percentage of additional edges and under a different and random permutation, were generated. The number of additional edges varied from 0 to 20$\%$ of the total number of edges of the fixed graph $\mathcal{G}_1$.

Results III

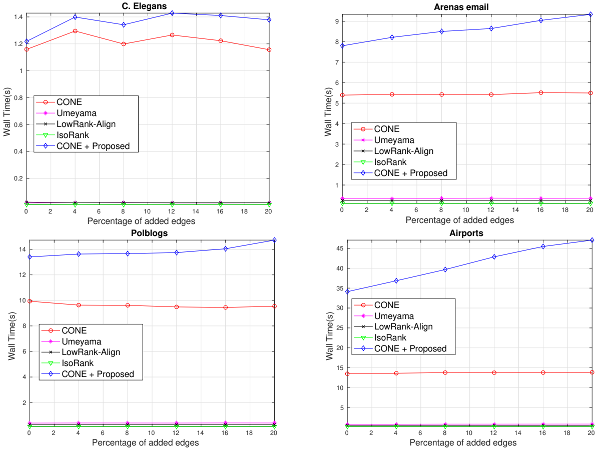


Figure: Wall time (in seconds) vs. noise level for different networks.

Results IV

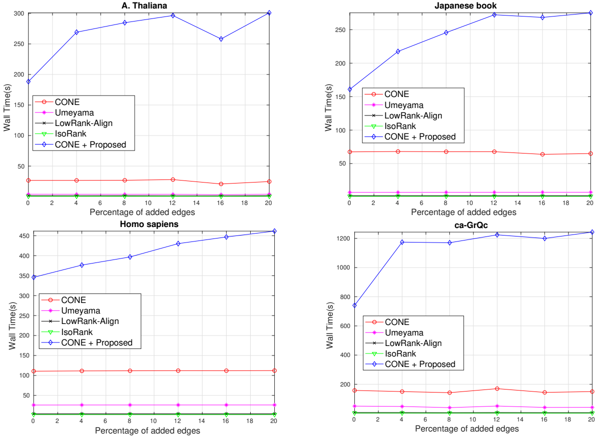


Figure: Wall time (in seconds) vs. noise level for different networks.

Introduction

Problem Formulation

Algorithmic Approach

Experimental Evaluation

Conclusion & Future Work

Conclusions & Future Work

- In this work, we:
  (i) proposed a novel formulation of graph alignment
  (ii) developed an optimization algorithm
  (iii) compared it against the *state-of-the-art*
- Our results indicate:
  (i) we achieve **much higher** alignment accuracy
  (ii) even in challenging problem instances
  (iii) there is a lot of room for improvement!
- Future work:
  (i) more efficient/scalable methods for the proposed formulation
  (ii) testing the embeddings for other tasks

Thank you!