Multilevel Network Alignment

Presented by Si Zhang (ASU)

Si Zhang  Hanghang Tong  Ross Maciejewski  Tina Eliassi-Rad
Multiple Networks Are Prevalent!

**Scenarios**
- social networks
- transaction networks
- PPI networks
- knowledge graphs

- yeast
- elegans
- fly
- mouse

**Tasks**
- graph level
- subgraph level
- node level

- similar?
- matched subgraphs?
- node correspondence
What Is Network Alignment?

- Find node correspondence across networks
Other Applications

- Knowledge completion
Other Applications

- Fraud detection

Unsuspicious patterns become suspicious!

**Question**: How to identify the correspondences across networks?
Network Alignment: How to

- **Topological alignment**
  - If two nodes are aligned, their neighbors are likely to be aligned

- **Attributed alignment [Zhang’16]**
  - Consider both topological and attribute consistency

- **Embedding-based alignment [Liu’16]**
  - Aligned nodes are closed in the embedding space
Network Alignment: How to (con’t)

- Topological alignment: FINAL-P [Zhang’16]
  - if two nodes are likely to be aligned (i.e., similar)
  - their close neighbors are likely to be aligned (similar)

- Optimization formulation
  $$\min_{s_1} \alpha s_1^T (I - A_1 \otimes B_1) s_1 + (1 - \alpha) \|s_1 - h_1\|^2_2$$
  - $A_1, B_1$ are symmetrically normalized adjacency matrices
  - $s_1, h_1$ are the vectorization of alignment $S_1$ and preference $H_1$
  - convex optimization $\rightarrow$ global optimal solution

- Optimization algorithm
  - fixed point solution: $S_1 = \alpha B_1 S_1 A_1 + (1 - \alpha) H_1$
Network Alignment: Limitations

- **Existing methods**
  - Align networks at node level (and cluster level)
  - Have an at least **quadratic** computational complexity

- **Rich patterns in networks**
  - E.g., hierarchical cluster-within-clusters structure

- **Question:** how to align networks at different granularities?
Challenge #1: Alignment Accuracy

- Error propagation through different levels

- If soccer in $G_1$ is aligned with basketball in $G_2$
- Next cluster level: in $G_1$ cannot be aligned with in $G_2$
- Node level: nodes in cluster in $G_1$ can’t be aligned with nodes in cluster in $G_2$

**Question:** How to mitigate error propagation?
Challenge #2: Scalability

- **Time complexity**
  - At least $O(n^2)$ due to dense matrix multiplication

- **Space complexity**
  - At least $O(n^2)$ to store the dense alignment matrix

**Question**: How can we reduce the complexity?
Outline

- Motivations
- Q1: Moana Formulation
- Q2: Moana Algorithm
- Experimental Results
- Conclusions
Prob. Def: Multilevel Network Alignment

- **Given:**
  - (1) adjacency matrices $\overline{A}_1, \overline{B}_1$ of two undirected networks;
  - (2) a sparse prior alignment preference $H_1$;
  - (3) the number of levels $L \geq 2$ of interests.

- **Find:** a set of alignment matrices $S_l$ at level-$l$, $l = 1, \ldots, L$
  - where $S_1$ indicates the alignment at the node level

- **An illustrative example**
Moana Formulation #1: Multilevel Optimization

- Generic strategy
  - coarsening $\rightarrow$ alignment $\rightarrow$ interpolation

- Alignment interpolations
  - bilinear interpolations by $P_l \in R^{p_l \times n_1}, Q_l \in R^{q_l \times n_2}$ ($p_l \leq n_1, q_l \leq n_2$)
  - w.l.o.g., $S_1 = Q_1^T S_2 P_1$ between level-1 & level-2

- Multilevel alignment formulation
  - Level-1: $\min_{s_1} \alpha s_1^T (I - A_1 \otimes B_1)s_1 + (1 - \alpha)\|s_1 - h_1\|_2^2$
  - final $P$ at node level
  - Level-2: $\min_{s_2} \alpha s_2^T (I - A_2 \otimes B_2)s_2 + (1 - \alpha)\|s_2 - h_2\|_2^2$
  - where $A_2 = P_1 A_1 P_1^T, B_2 = Q_1 B_1 Q_1^T$ and $H_2 = Q_1 H_1 P_1^T$
  - same properties (e.g., convexity) and algorithm as final $P$
  - ‘good’ (semi-) orthogonal $P_1, Q_1$ can make $A_2, B_2$ well-represented
Moana Formulation #2: Perfect Interpolation

- Alignment error propagation
  - imperfect interpolations bring errors to $S_l$ even from optimal $S_{l+1}^*$
  - mathematically, $S_l^* \neq Q_l^T S_{l+1}^* P_l$ if $P_l, Q_l$ are not well-chosen
  - errors can be propagated or diverged to level-(l-1)

- Perfect interpolation
  - if $P_l, Q_l$ ($l = 1, \cdots, L - 1$) are orthogonal
  - then $S_l^* = Q_l^T S_{l+1}^* P_l$ where $S_l^*, S_{l+1}^*$ are optimal solutions at level-$l$ and level-(l+1)
  - proof in the paper
Outline

- Motivations
- Q1: Moana Formulation
- Q2: Moana Algorithm
- Experimental Results
- Conclusions
Moana Algorithm #1: Coarsening

- **Generic strategy**
  - coarsening → alignment → interpolation

- **Network coarsening by** $P_l, Q_l$
  - $A_{l+1} = P_l A_l P_l^T, B_{l+1} = Q_l B_l Q_l^T$

- **Requirements on** $P_l, Q_l$
  - **perfect interpolation**: they are orthogonal matrix
  - **efficient computation**: they are sparse matrix
  - **informative coarsening**: they can uncover hierarchical cluster-within-clusters structures
Moana Algorithm #1: Coarsening (Con’t)

- **Multiresolution matrix factorization** [Kondor’14]
  - $\Pi$ is to reorder for visualization (no need to calculate)
  - $P_l$ contains: (1) a rotation matrix block, (2) an identity matrix block
  - active set $S_l$ indicates nodes at the $l$-th granularity (i.e., clusters)

- **Coarsening procedure**
  - $P_{L-1} \cdots P_2 P_1 A_1 P_1^T P_2^T \cdots P_{L-1}^T = A_L \rightarrow \tilde{A}_L$
  - $Q_{L-1} \cdots Q_2 Q_1 B_1 Q_1^T Q_2^T \cdots Q_{L-1}^T = B_L \rightarrow \tilde{B}_L$

- **Remark:** $S(S_{B_l}, S_{A_l})$ indicates the alignment among clusters at the $l$-th granularity
Moana Algorithm #2: Alignment

- Generic strategy
  - coarsening → **alignment** → interpolation

- Alignment across the coarsest networks

\[
\tilde{S}_L = \alpha \begin{bmatrix} \tilde{B}_{L1} & 0 \\ 0 & \tilde{B}_{L2} \end{bmatrix} \begin{bmatrix} \tilde{S}_{L1} \\ \tilde{S}_{L3} \end{bmatrix} \begin{bmatrix} \tilde{A}_{L1} \\ 0 \end{bmatrix} + (1 - \alpha) \begin{bmatrix} \tilde{H}_{L1} \\ \tilde{H}_{L3} \end{bmatrix} 
\]

- block-wise computation

\[
\tilde{S}_{L1} = \alpha \tilde{B}_{L1} \tilde{S}_{L1} \tilde{A}_{L1} + (1 - \alpha) \tilde{H}_{L1} \\
\tilde{S}_{L2} = \alpha \tilde{B}_{L1} \tilde{S}_{L2} \tilde{A}_{L2} + (1 - \alpha) \tilde{H}_{L2} \\
\tilde{S}_{L3} = \alpha \tilde{B}_{L2} \tilde{S}_{L3} \tilde{A}_{L1} + (1 - \alpha) \tilde{H}_{L3} \\
\tilde{s}_{L4} = (1 - \alpha) (I - \alpha \tilde{A}_{L2} \otimes \tilde{B}_{L3})^{-1} \tilde{h}_{L4}
\]

- matrix composition: e.g., \( S_L(S_{BL}, S_{AL}) = \tilde{S}_{L1}, \ S_L(\tilde{S}_{BL}, \tilde{S}_{AL}) = \tilde{S}_{L4} \)

- Alignment at finer levels

  - perfect interpolations: \( S_l = Q_l^T S_{l+1} P_l \)
Moana Algorithm: Analysis

- **Alignment error bound**
  \[
  \frac{\|S_l^* - S_l\|_F}{\|S_l\|_F} \leq \frac{2\epsilon \kappa}{1 - \epsilon \kappa}, \forall l = 1, \ldots, L
  \]
  where \( \epsilon = \sqrt{\frac{\alpha}{2n}} (\delta_1 r_2 + \delta_2 r_1 + \delta_1 \delta_2) \),
  \( \delta_1 = \|A_L - \tilde{A}_L\|_F, \delta_2 = \|B_L - \tilde{B}_L\|_F \),
  \( \kappa \) is condition number, \( r_1, r_2 \) are ranks

- **Complexity analysis**
  - linear time and space complexity
Outline

- Motivations ✅
- Q1: Moana Formulation ✅
- Q2: Moana Algorithm ✅
- Experimental Results
- Conclusions
Experimental Setup

- **Datasets**
  - Gr-Qc network vs. its permutation (nodes: 5,241 vs. 5,241)
  - Google+ network vs. its permutation (nodes: 23,628 vs. 23,628)
  - Amazon co-purchasing networks (nodes: 74,596 vs. 66,951)
  - ACM vs DBLP coauthor networks (nodes: 9,872 vs. 9,916)

- **Evaluation objectives**
  - Effectiveness: how accurate is our algorithm in aligning networks?
  - Efficiency: how fast and scalable is our algorithm?

- **Comparison methods**

<table>
<thead>
<tr>
<th></th>
<th>Moana</th>
<th>AMG-FINAL</th>
<th>Umeyama</th>
<th>PriorSim</th>
</tr>
</thead>
<tbody>
<tr>
<td>FINAL-P</td>
<td></td>
<td>HubAlign</td>
<td>ModuleAlign</td>
<td>iNeat</td>
</tr>
</tbody>
</table>
R1: Effectiveness Results

- Effectiveness in node-level alignment

**Observations:**
1. the performance of Moana is close to FINAL-P;
2. Moana outperforms all other methods.
R2: Effectiveness Results

- Effectiveness in cluster-level alignment

Observations: Moana achieves a good performance in cluster alignment at different levels.
R3: Case Study on Multilevel Alignment

- A case study on Zachary’s Karate networks

**Observations:** Moana can unveil meaningful alignment of clusters at different granularities.
R4: Quality-Speed Balance

Observations: Moana can achieve a better quality-speed balance.
R5: Scalability

Observations:
(1) Moana scales linearly w.r.t. the number of edges;
(2) Moana scales linearly w.r.t. the number of nonzero elements in $H_1$. 
Outline

- Motivations ✓
- Q1: Moana Formulation ✓
- Q2: Moana Algorithm ✓
- Experimental Results ✓
- Conclusions
Conclusions

▪ Multilevel network alignment
  – Q1: Formulation
  – A1: Multilevel optimization + perfect interpolation
  – Q2: Scalability
  – A2: Moana algorithm

▪ Results
  – Moana outperforms most baseline methods in node alignment
  – Moana achieves good performance in cluster alignment
  – Moana has linear complexity

▪ More in paper
  – Proof of algorithm analysis & more experimental results