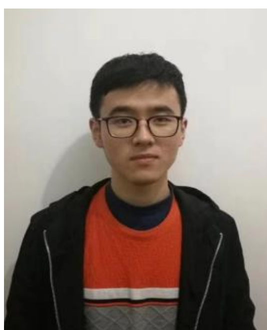
BRIGHT: A Bridging Algorithm for Network Alignment

Yuchen Yan
(UIUC)

Si Zhang
(UIUC)

Hanghang Tong
(UIUC)



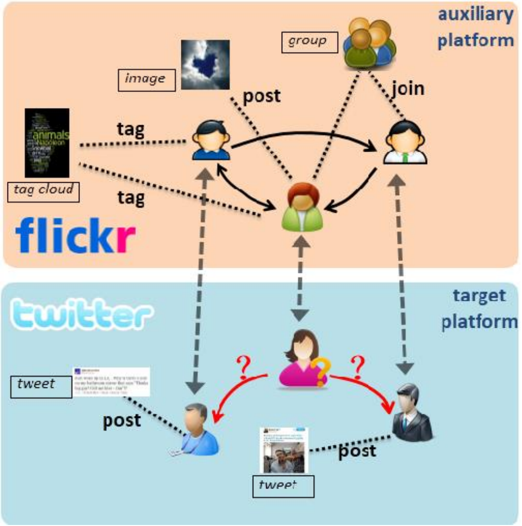
Outline

- Introduction
- Theoretical Analysis
- Proposed Model
- Experimental Results
- Conclusions

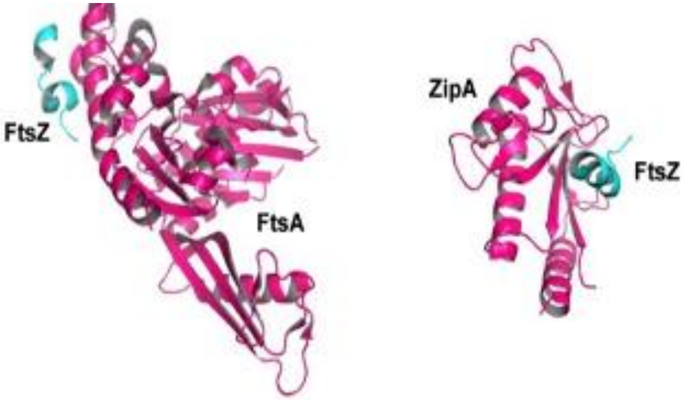
Network Alignment

- Networks are often multi-sourced
- To find node correspondence across networks

Friend recommendation

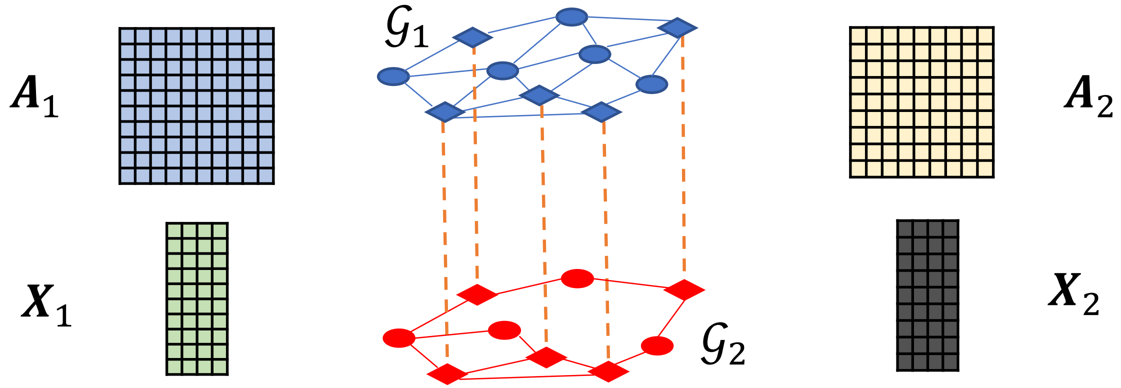


Drug design



Prob. Def.: Semi-Supervised Attributed Network Alignment

- **Given:** (1) two attributed networks $\mathcal{G}_1 = \{\boldsymbol{A}_1, \boldsymbol{X}_1\}$, $\mathcal{G}_2 = \{\boldsymbol{A}_2, \boldsymbol{X}_2\}$; (2) a set of anchor node pairs $\boldsymbol{L}$.
- **Output:** an $n_2 \times n_1$ alignment/similarity $\boldsymbol{S}$.
- Scenario variants:
  - Semi-supervised plain network alignment (without $\boldsymbol{X}_1$, $\boldsymbol{X}_2$)
  - Unsupervised attributed network alignment (without $\boldsymbol{L}$)

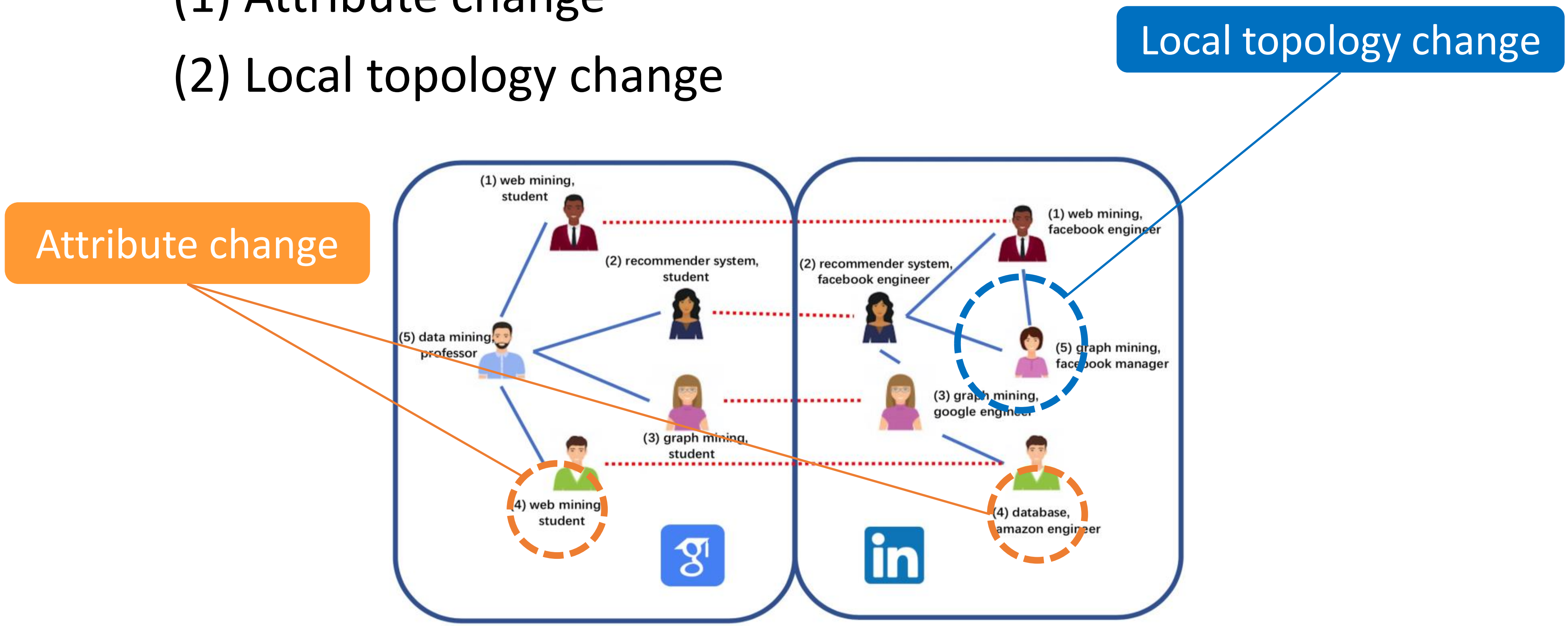


Existing Methods: Limitation #1

- Consistency optimization based methods
  - Consistency assumption violation:

    (1) Attribute change

    (2) Local topology change

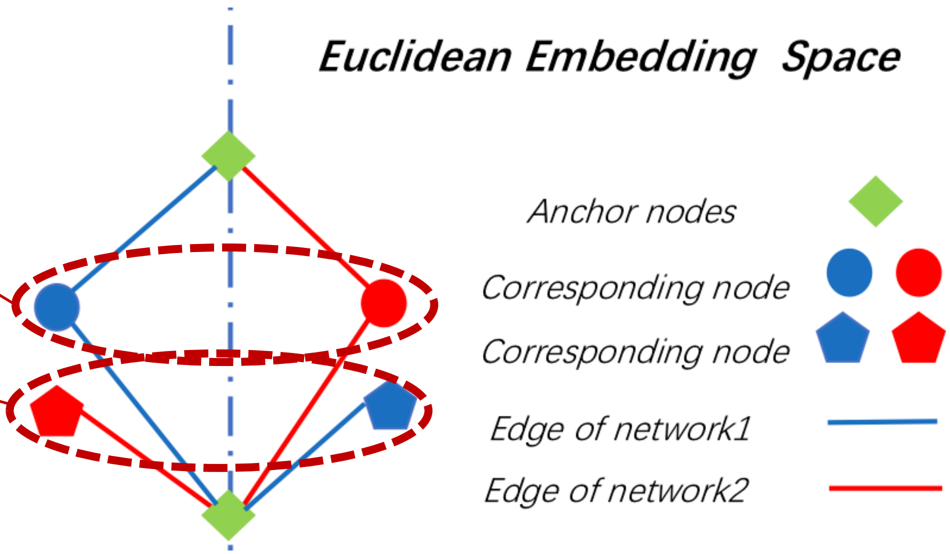


[1] Danai Koutra, Hanghang Tong, and David Lubensky. 2013. Big-align: Fast bipartite graph alignment. In 2013 IEEE 13th International Conference on Data Mining. IEEE, 389–398.
[2] Rohit Singh, Jinbo Xu, and Bonnie Berger. 2008. Global alignment of multiple protein interaction networks with application to functional orthology detection. Proceedings of the National Academy of Sciences 105, 35 (2008), 12763–12768.

Existing Methods: Limitation #2

- Embedding based methods
  —Introduce the space disparity issue

Corresponding nodes lie symmetrically in the embedding space.



Outline

- Introduction ✓
- Theoretical Analysis
- Proposed Model
- Experimental Results
- Conclusions

Theoretical Analysis #1

Main Claim/Insight: Consistency optimization based methods are essentially random walk propagation of anchor links.

- Drawbacks

—Exactly same steps

—Equal weight for all anchor links

$$O_c(\mathbf{S}) = \alpha \sum_{a,b,x,y} \left[\frac{\mathbf{S}(x,a)}{\sqrt{\mathbf{d}_2(x)\mathbf{d}_1(a)}} - \frac{\mathbf{S}(y,b)}{\sqrt{\mathbf{d}_2(y)\mathbf{d}_1(b)}}\right]^2 \mathbf{A}_1(a,b)\mathbf{A}_2(x,y) + (1-\alpha)\|\mathbf{S}-\mathbf{H}\|_F^2$$

1. Topology Consistency

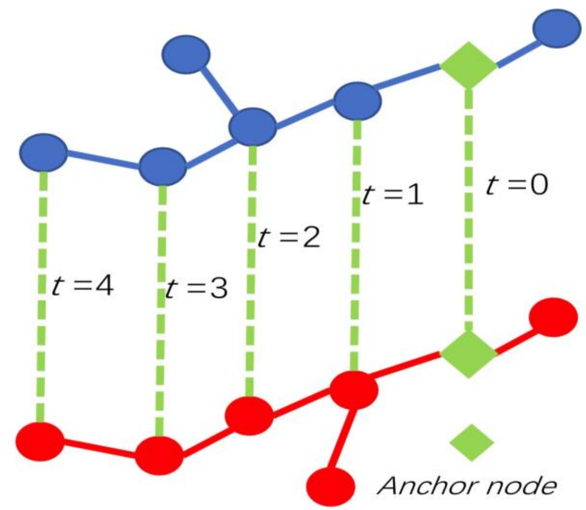
2. Known Anchor Links

Objective function solutions

$$\mathbf{s} = (1-\alpha)(\mathbf{I} - \alpha\hat{\mathbf{W}})^{-1}\mathbf{h}$$

$$\mathbf{W} = \mathbf{A}_1 \otimes \mathbf{A}_2$$

$$\mathbf{s}(i) = (1-\alpha)\sum_{j=0}^{n_2\times n_1-1}\sum_{t=0}^{\infty}\alpha^t \hat{\mathbf{W}}^t(i,j)\mathbb{1}(\mathbf{h}(j))$$



[1] Rohit Singh, Jinbo Xu, and Bonnie Berger. 2008. Global alignment of multiple protein interaction networks with application to functional orthology detection. Proceedings of the National Academy of Sciences 105, 35 (2008), 12763–12768.
[2] Si Zhang and Hanghang Tong. 2016. Final: Fast attributed network alignment. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 1345–1354.

Theoretical Analysis #2

Main Claim/Insight: Embedding based methods relax the objective function of consistency optimization based methods.

- Drawback

—Space disparity

$$O_1^{in} = \sum_a (d(a,b) - d(a,c))$$

$$O^{cross} = \sum_{(l_1,l_2)\in\mathcal{L}} d(l_1, l_2)$$

Two parts of loss

In a simple *K*-regular graph:

Consistency method loss

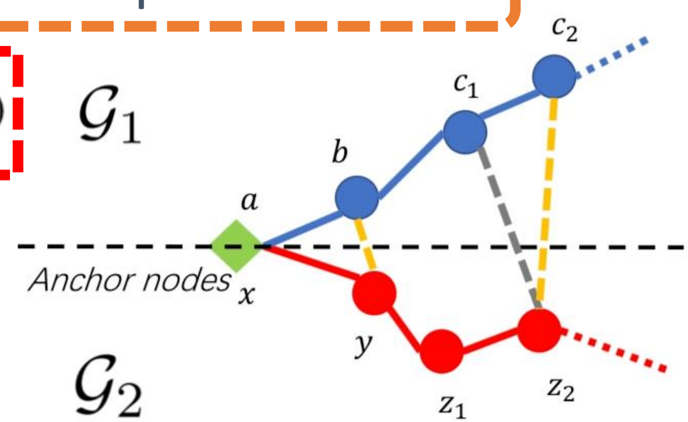
$$O_c(\mathrm{S}) = \frac{1}{K^2}(1 - e^{-d(y,b)})^2 \mathrm{A}_1(a,b)\mathrm{A}_2(x,y)$$

Embedding method loss

$$O = d(b,a) + d(a,x) + d(x,y)$$

Relaxation relation

$$d(b,y) \leq d(b,a) + d(a,x) + d(x,y)$$



[1] Xiaokai Chu, Xinxin Fan, Di Yao, Zhihua Zhu, Jianhui Huang, and Jingping Bi. 2019. Cross-Network Embedding for Multi-Network Alignment. In The World Wide Web Conference (WWW ’19). ACM, New York, NY, USA, 273–284. https://doi.org/10.1145/3308558.3313499

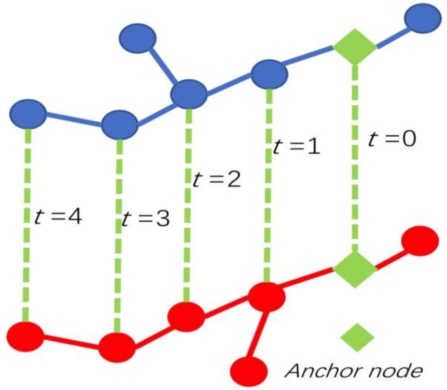
[2] Li Liu, William K Cheung, Xin Li, and Lejian Liao. [n.d.]. Aligning Users across Social Networks Using Network Embedding.

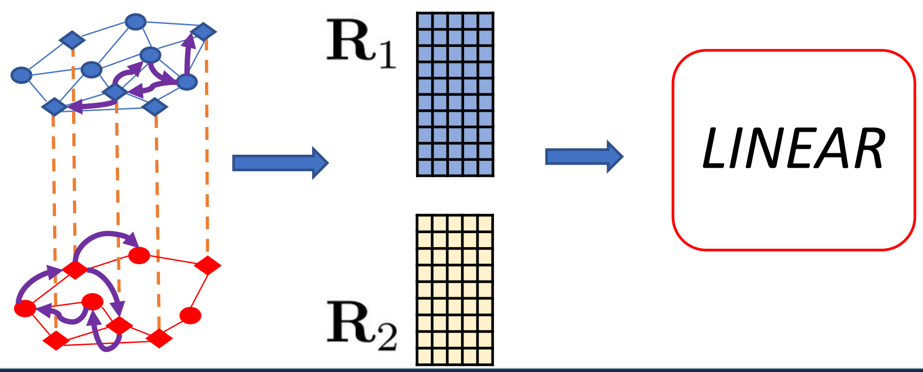
Outline

- Introduction
- Theoretical Analysis ✓
- Proposed Model
- Experimental Results
- Conclusions

Key Idea #1: RWR for Flexible Propagation

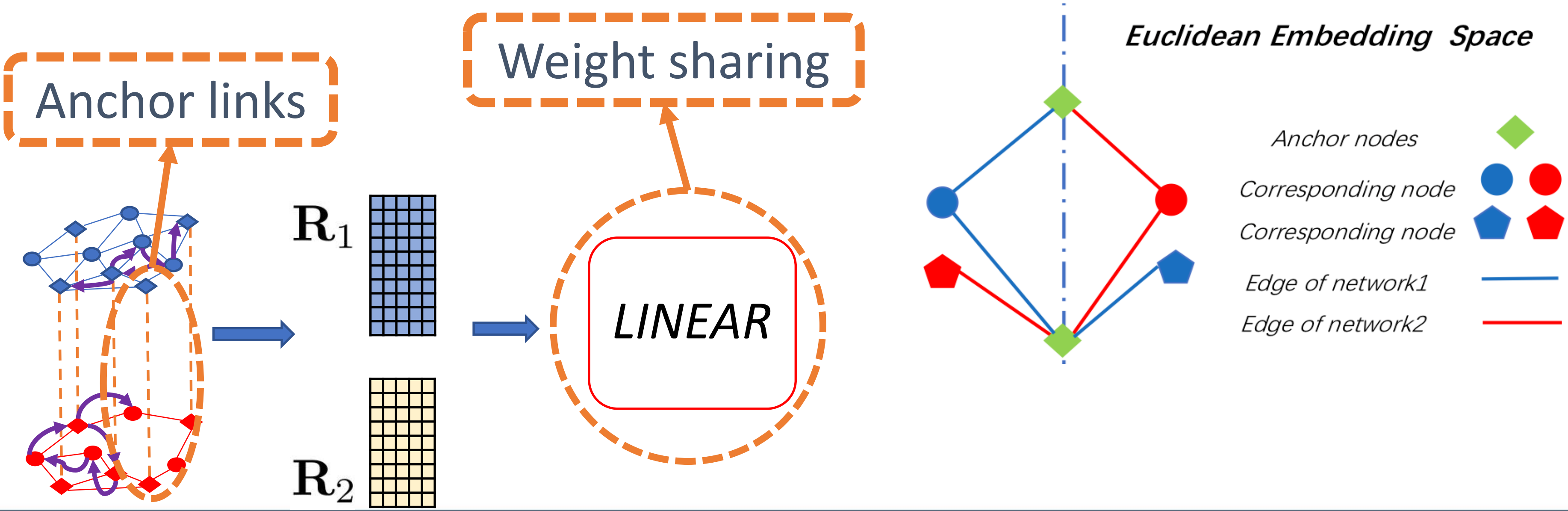
- Drawbacks for consistency optimization based methods
  — Exactly same steps
  Random walk with restart allows **restart**.
  — Equal weight for all anchor links
  Linear layer trains **different weights**.





Key Idea #2: Build a Unified Space

- Drawbacks for embedding based methods
  —Space disparity
  (1) Anchor links as basis
  (2) Weight sharing



Part #1: BRIGHT-U (Plain Network)

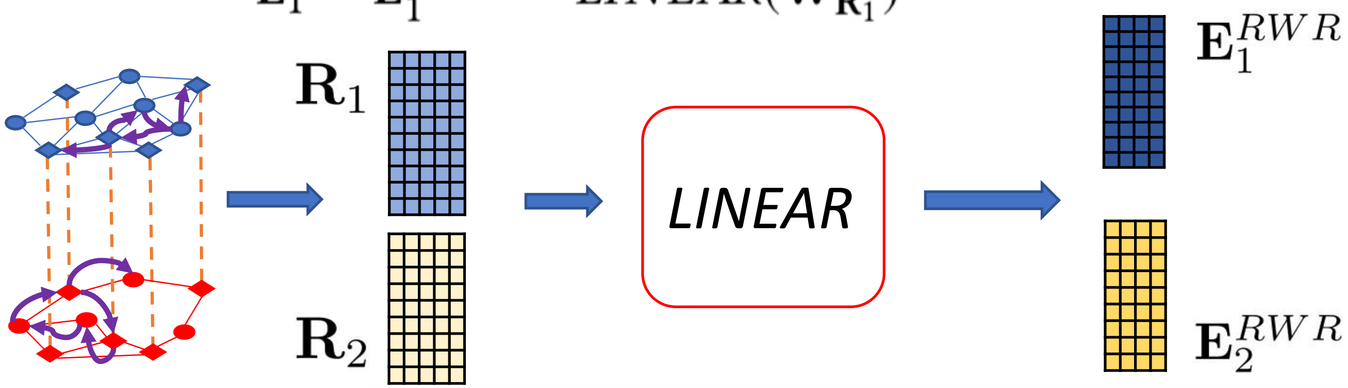
- RWR from anchor links

$$\mathbf{r}_{l_1} = (1-\beta)\hat{\mathbf{W}}_1\mathbf{r}_{l_1} + \beta\mathbf{e}_{l_1} \qquad \hat{\mathbf{W}}_1 = (\mathbf{D}^{-1}\mathbf{A}_1)^T$$

$$\mathbf{r}_{l_1} = \beta(\mathbf{I} - (1-\beta)\hat{\mathbf{W}}_1)^{-1}\mathbf{e}_{l_1}$$

- Put all r_{l_i} together as RWR embedding matrix $\hat{\mathbf{W}}_{\mathbf{R}_1}$
- Use a shared linear layer to adjust anchor link weights

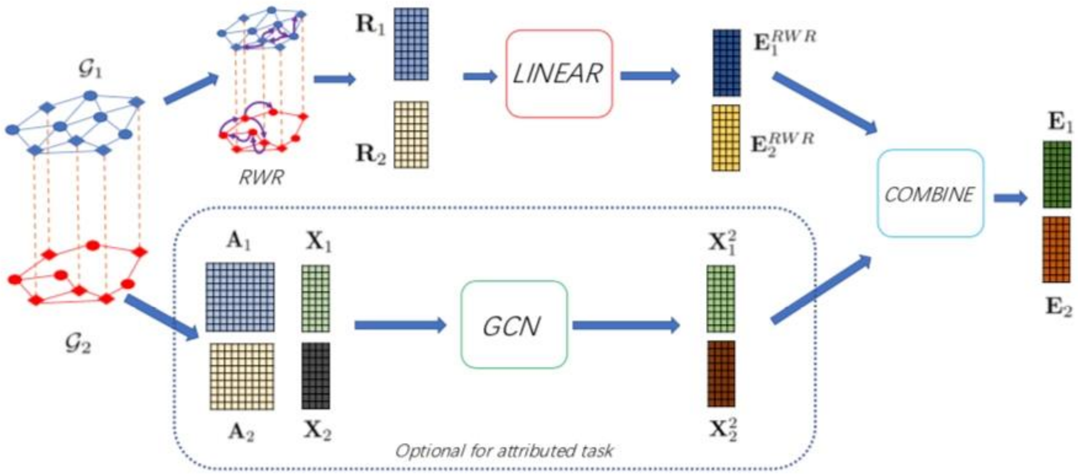
$$\mathbf{E}_1 = \mathbf{E}_1^{RWR} = LINEAR(\hat{\mathbf{W}}_{\mathbf{R}_1})$$



Part #2: BRIGHT-A (Attributed Network)

- Compute RWR embedding same as BRIGHT-U
- Use a shared two-layer GCN to capture attribute
- Combine RWR embedding and GCN embedding

$$\mathbf{E}_1 = COMBINE([\mathbf{E}_1^{RWR} \| \mathbf{X}_1^2])$$



Part #3: Model Training

- Ranking loss

Negative sample set

Margin parameter

$$\mathcal{J}_i = \frac{1}{|\mathcal{L}|} \sum_{l \in \mathcal{L}} \frac{1}{|U_{l_i}|} \sum_{u \in U_{l_i}} \max\{0, \gamma + d(l_1, l_2) - d(l_i, u)\}$$

Negative sample set

$$\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$$

- Advanced negative sampling

  —Sort-then-select

- Construct the alignment matrix

$$\mathrm{S}(x, a) = e^{-d(x,a)}$$

Outline

- Introduction
- Theoretical Analysis
- Proposed Model ✓
- Experimental Results
- Conclusions

Experimental Setup

- Datasets

Categories	Networks	# of Nodes	# of Edges	# of Attributes
Plain Networks	*Foursquare*	5,313	54,233	–
	Twitter	5,120	130,575	–
	ACM	9,916	44,808	–
	DBLP	9,872	39,561	–
Attributed Networks	*ACM(A)*	9,916	44,808	17
	DBLP(A)	9,872	39,561	17
	Cora-1	2,708	5,806	1,433
	Cora-2	2,708	4,547	1,433

- Metrics: Hit@K, MRR
- Baseline methods:
  - Plain network: CrossMNA, IONE, FINAL-P
  - Attributed network: REGAL, FINAL-N, NetTrans

[1] Xiaokai Chu, Xinxin Fan, Di Yao, Zhihua Zhu, Jianhui Huang, and Jingping Bi. 2019. Cross-Network Embedding for Multi-Network Alignment. In The World Wide Web Conference (WWW '19). ACM, New York, NY, USA, 273–284. https://doi.org/10.1145/3308558.3313499
[2] Li Liu, William K Cheung, Xin Li, and Lejian Liao. [n.d.]. Aligning Users across Social Networks Using Network Embedding.
[3] Si Zhang and Hanghang Tong. 2016. Final: Fast attributed network alignment. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 1345–1354.
[4] Si Zhang, Hanghang Tong, Yinglong Xia, Liang Xiong, and Jiejun Xu. 2020. NetTrans: Neural Cross-Network Transformation. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '20). Association for Computing Machinery, New York, NY, USA, 986–996. https://doi.org/10.1145/3394486.3403141
[5] Mark Heimann, Haoming Shen, Tara Safavi, and Danai Koutra. 2018. REGAL: Representation Learning-based Graph Alignment. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM '18). ACM, New York, NY, USA, 117–126. https://doi.org/10.1145/3269206.3271788

Experimental Results-Alignment

Plain Task	*DBLP vs. ACM*				*Foursquare vs. Twitter*			
Metrics	Hit@1	Hit@10	Hit@30	MRR	Hit@1	Hit@10	Hit@30	MRR
CrossMNA	7.90%	62.53%	79.48%	23.42%	0.00%	3.26%	12.03%	1.48%
IONE	30.91%	74.25%	84.11%	46.26%	4.50%	16.69%	27.80%	8.56%
FINAL-P	19.49%	68.75%	81.23%	35.00%	4.97%	22.22%	32.25%	10.31%
BRIGHT-U	40.45%	81.26%	84.13%	53.85%	6.37%	25.24%	33.54%	13.04%
Attributed Task	*DBLP(A) vs. ACM(A)*				*Cora-1 vs. Cora-2*			
Metrics	Hit@1	Hit@10	Hit@30	MRR	Hit@1	Hit@10	Hit@30	MRR
REGAL	36.26%	60.36%	69.51%	44.92%	45.66%	60.90%	69.21%	51.11%
FINAL-N	38.18%	79.74%	89.07%	52.15%	86.29%	91.32%	91.37%	88.70%
NetTrans	11.84%	84.11%	94.53%	30.11%	27.56%	90.95%	97.51%	49.67%
BRIGHT-A	45.26%	86.76%	92.17%	59.87%	83.85%	99.08%	99.68%	90.41%

Observation: (1) An advantage of 10% in Hit@1 on *DBLP vs. ACM;*
(2) 99% in Hit@10 on Cora.

Experimental Results-Ablation Study

plain Task	*DBLP vs. ACM*		*Foursquare vs. Twitter*	
Metrics	Hit@10	MRR	Hit@10	MRR
BRIGHT-U(SPD)	75.75%	48.78%	6.06%	3.07%
BRIGHT-U	81.26%	53.85%	25.24%	13.04%

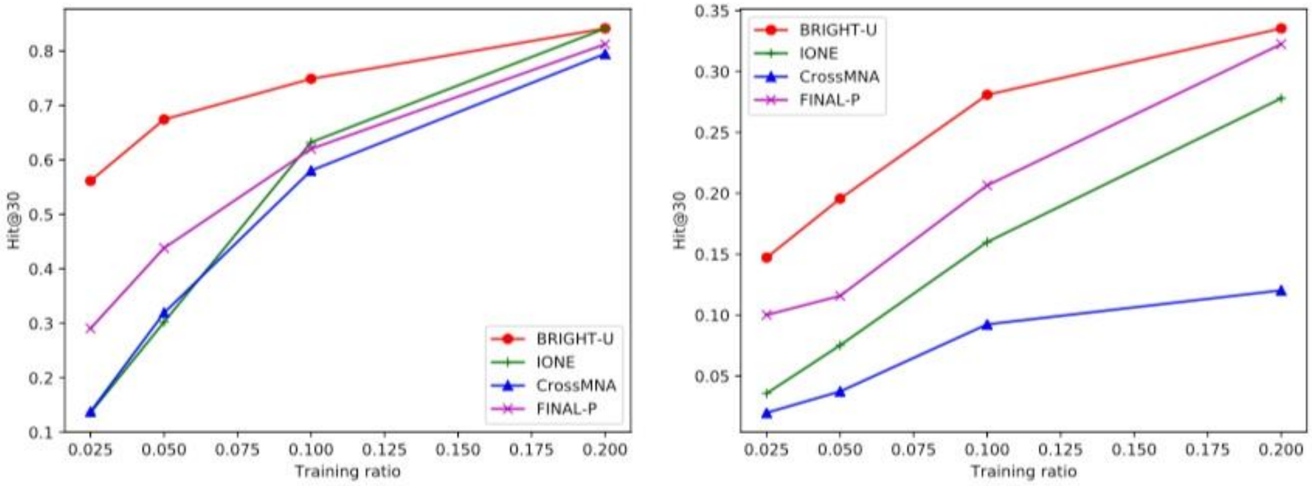
Ablation Study for BRIGHT-U

Attributed Task	*DBLP(A) vs. ACM(A)*		*Cora-1 vs. Cora-2*	
Metrics	Hit@10	MRR	Hit@10	MRR
BRIGHT-A(-RWR)	79.43%	51.61%	99.08%	90.12%
BRIGHT-A(-RWR:3500)	84.31%	58.01%	99.08%	90.12%
BRIGHT-A	86.76%	59.87%	99.08%	90.41%

Ablation Study for BRIGHT-A

Observation: (1) RWR module performs better than SPD;
(2) Attribute plays an important role in BRIGHT-A.

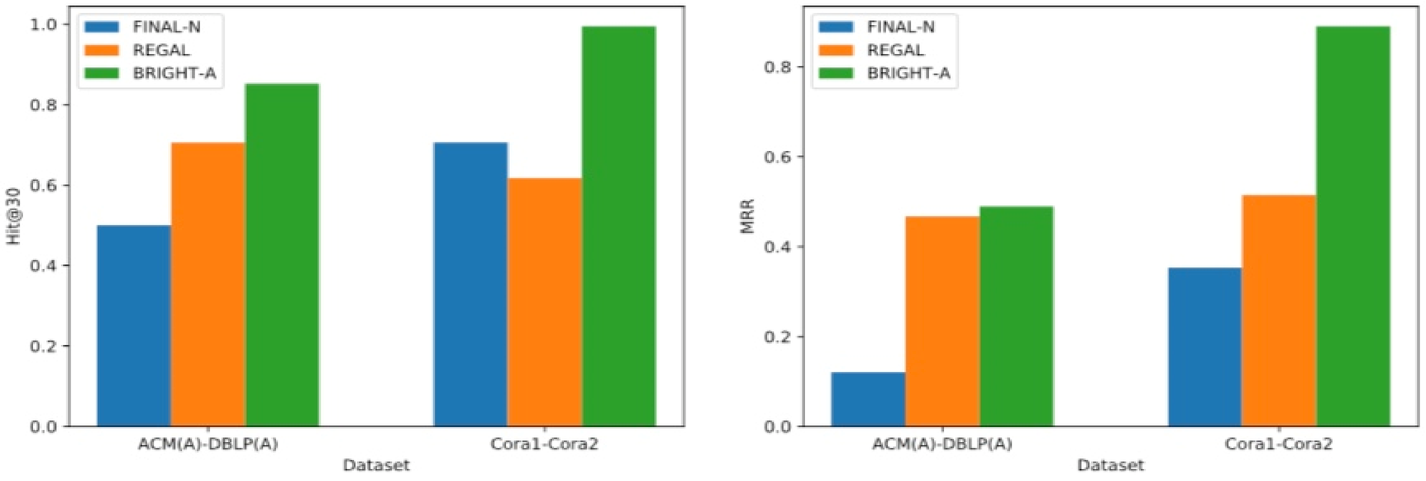
Experimental Results-Small Training Ratio



(a) Hit@30 with small training ratio on ***S1***(b) Hit@30 with small training ratio on ***S2***.

Observation: (1) Perform well under small training ratio;
(2) Avoid the space disparity issue better.

Experimental Results-Unsupervised Setting



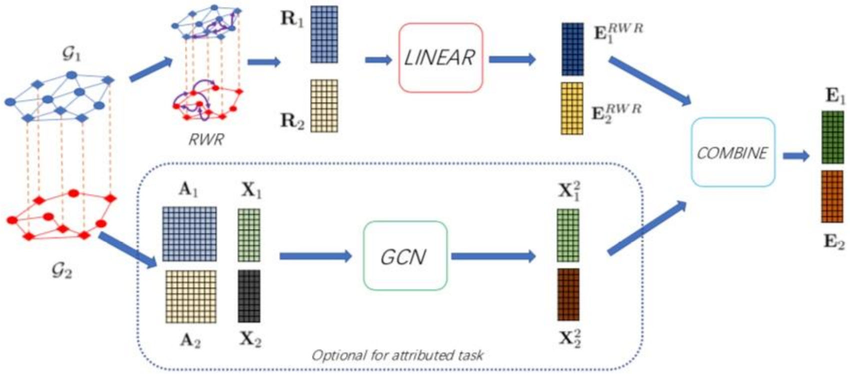
(a) Hit@30 (b) MRR

Observation: Good generalization to unsupervised setting.

Outline

- Introduction
- Theoretical Analysis
- Proposed Model
- Experimental Results ✓
- Conclusions

Conclusions



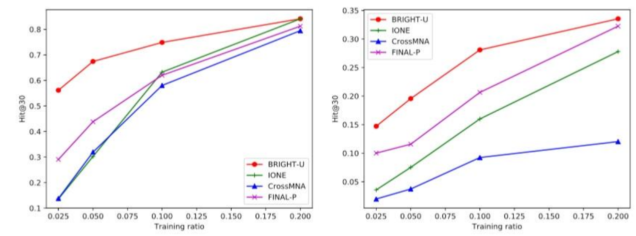
Problem: Network Alignment

Solution:

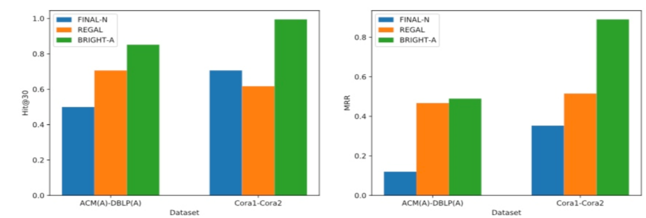
- BRIGHT-U
  - RWR for flexible propagation
  - Anchor link as basis
  - Weight sharing
- BRIGHT-A
  - Shared GCN to capture attribute

Results:

- Outperform all baselines
- Perform well with small training ratio
- Generalization to unsupervised setting



(a) Hit@30 with small training ratio on *S1*(b) Hit@30 with small training ratio on *S2*.



(a) Hit@30 (b) MRR