CIKM 2019 | BEIJING CHINA
November 03-07. 2019

Collective Link Prediction Oriented Network Embedding with Hierarchical Graph Attention

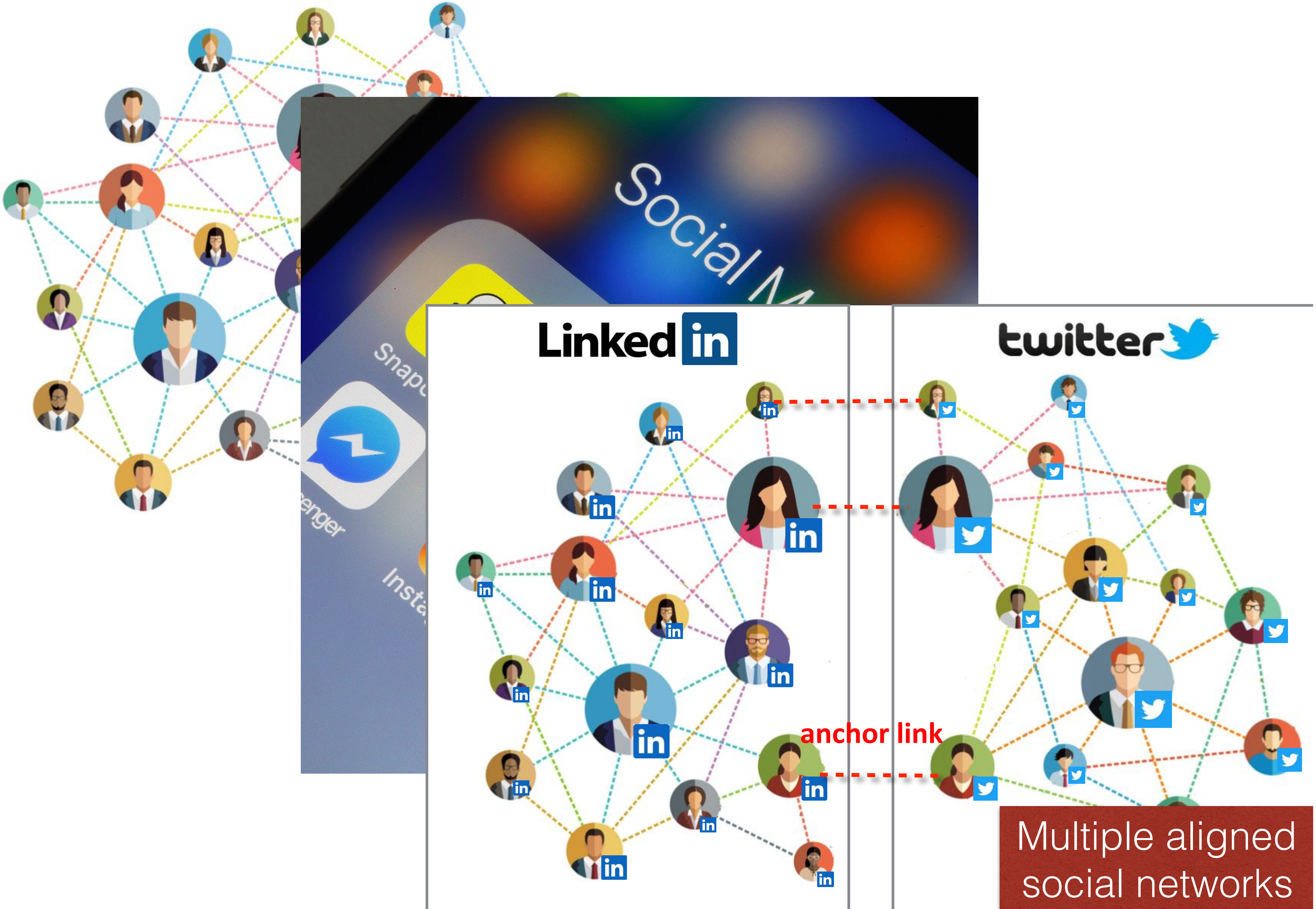
Yizhu Jiao[1], Yun Xiong[1], Jiawei Zhang[2] and Yangyong Zhu[1]

[1] Shanghai Key Laboratory of Data Science, Fudan University, China
[2] IFM Lab, Florida State University, USA

Datalogy 上海市数据科学重点实验室 Shanghai Key Laboratory of Data Science

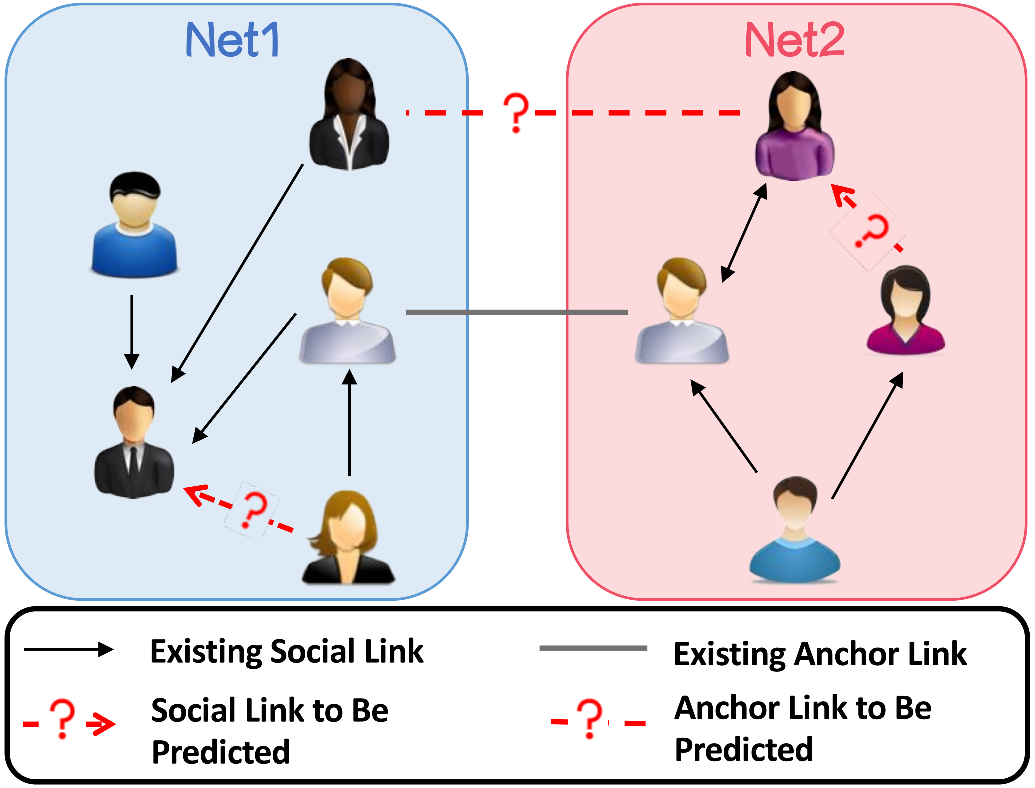
IFM Lab Information Fusion and Mining Laboratory

Multiple aligned social networks



Multiple aligned social networks

Problem Studied: **Collective Link Prediction**



Problem Definition:

- It covers several different link formation prediction tasks simultaneously including both the *intra-network social link prediction* and the *inter-network anchor link prediction*.

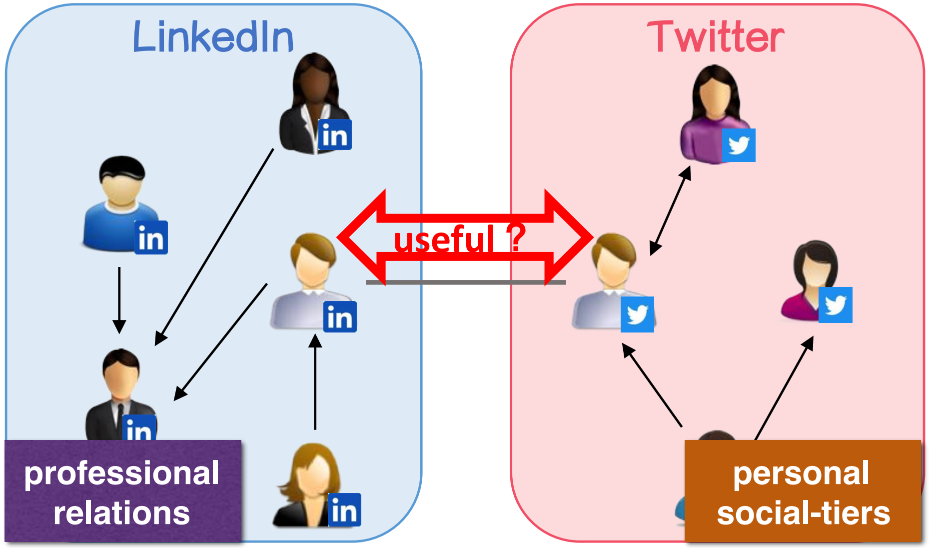
Why Study Together:

- take advantage of their strong correlations to enhance the prediction performance

Main Challenges

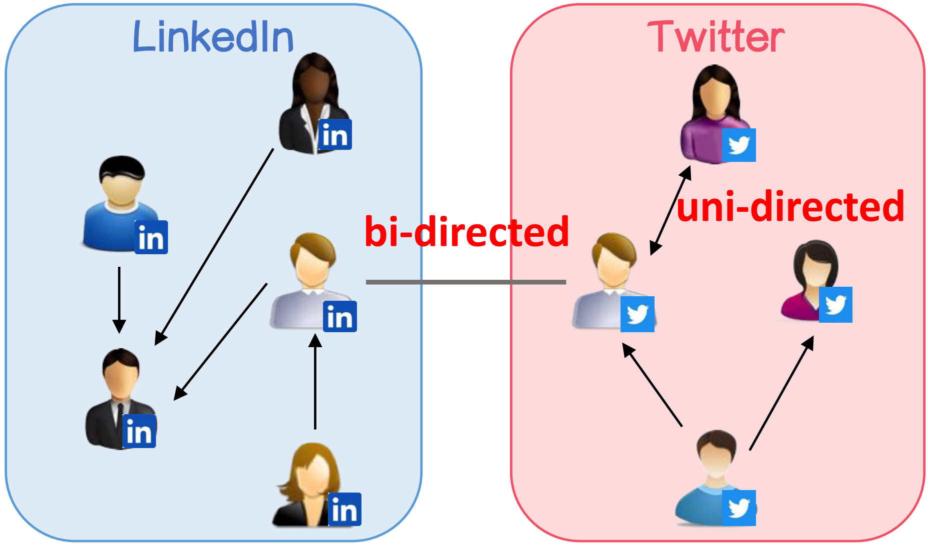
Network characteristic differences

- Each network reveals different aspects of the users.
- Information transfer could also **deteriorate** the performance of intra-network link prediction

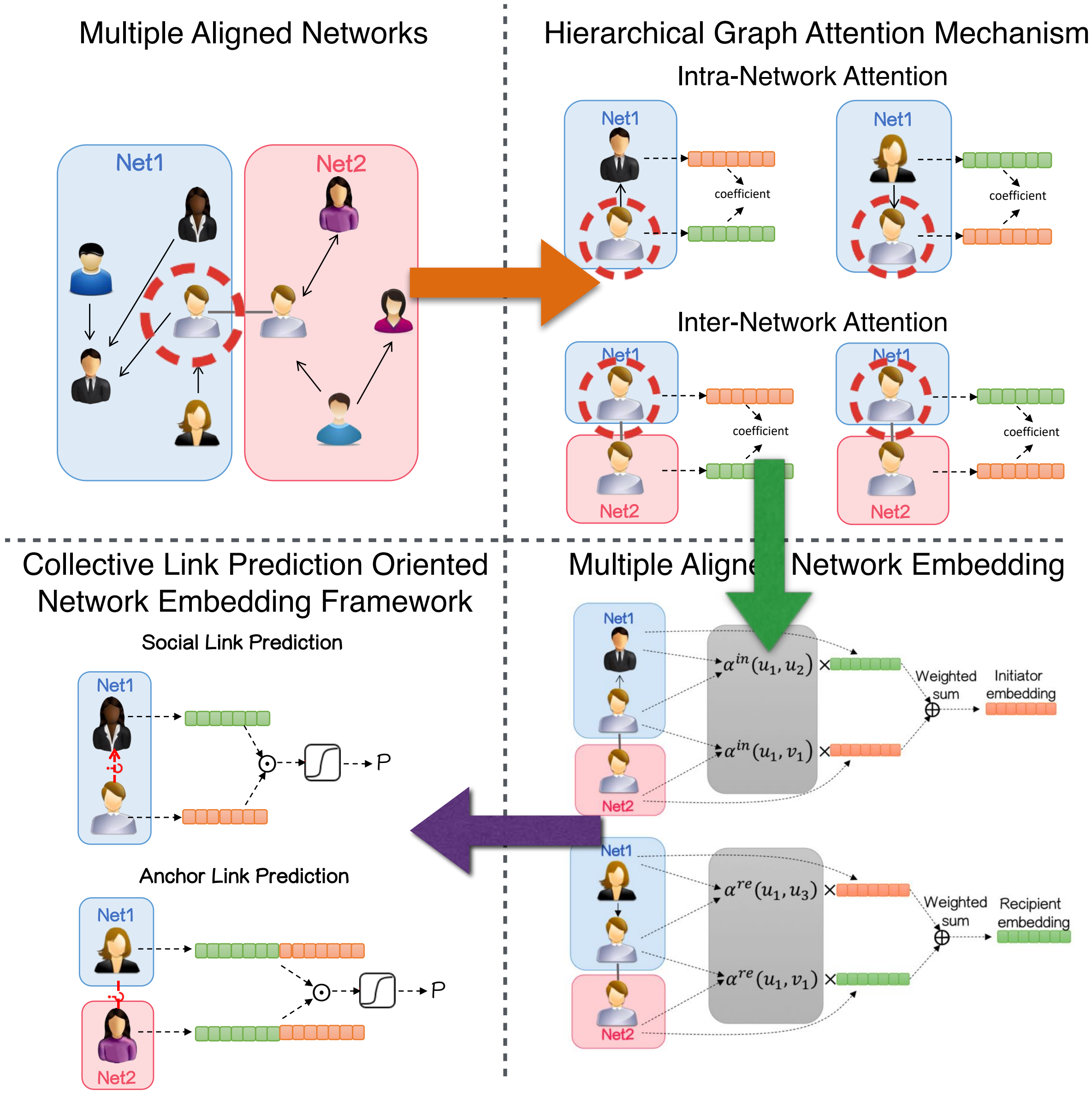


Link directivity differences

- The intra-network social links are usually **uni-directed** while the inter-network anchor links are **bi-directed**.
- Such different directivity properties should be carefully considered.



Proposed Framework: **HGANE**
(Hierarchical Graph Attention based Network Embedding)



step 1: Hierarchical graph attention mechanism can pay more or less attention to different neighbor nodes and networks to resolve the *network characteristic differences* problems.

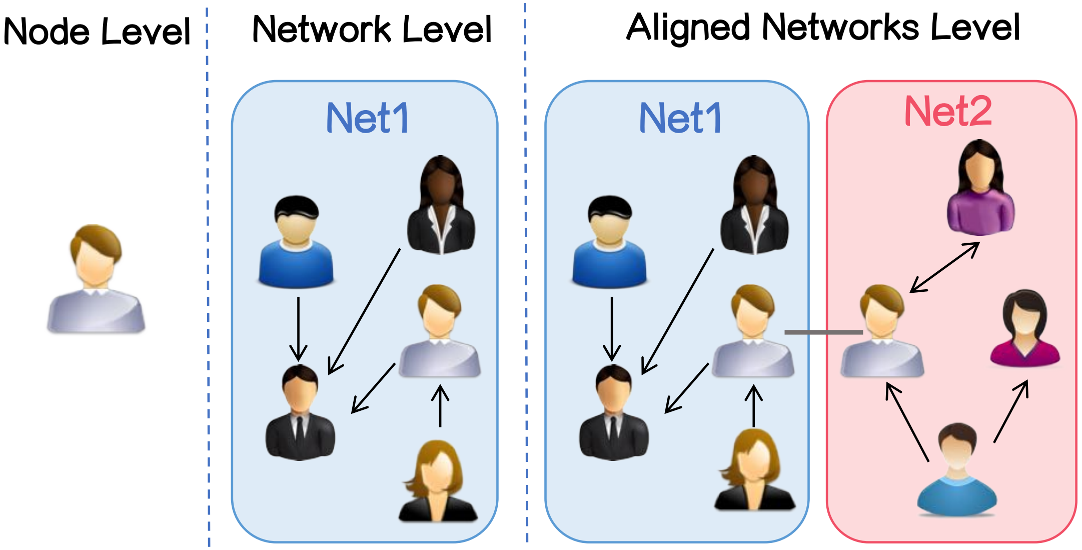
step 2: Multiple aligned network embedding can be learnt by aggregating information from both the *intra-network neighbors* and *inter-network partners*.

step 3: HGANE incorporates the collective link prediction task into a unified framework and balances between multiple prediction tasks.

Hierarchical Graph Attention Mechanism

Multiple aligned networks have a hierarchical structure…

- For the target node, the relevance of different neighbors is different.
- For the target network, other networks are differentially informative since they have different characteristics.



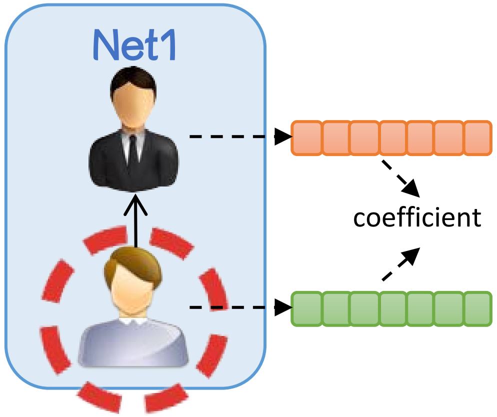
We proposed hierarchical graph attention mechanism including

- *Intra-network social attention* (at the node level)
- *Inter-network anchor attention* (at the network level)

to resolve the problems of *network characteristic differences* and *link directivity challenges* in the multiple aligned networks.

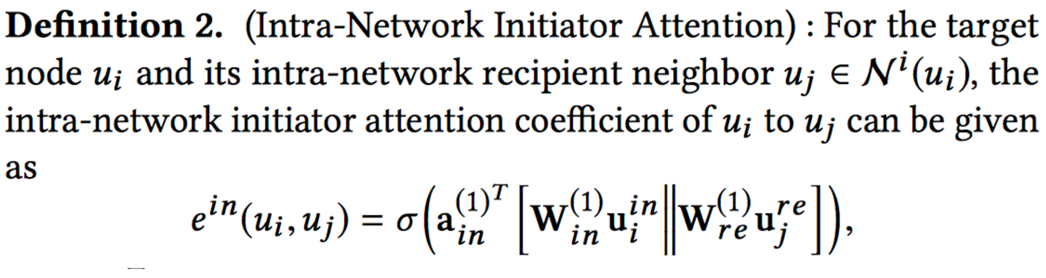
Intra-Network Social Attention

Each node is represented with two vector representations, the initiator feature and the recipient feature.



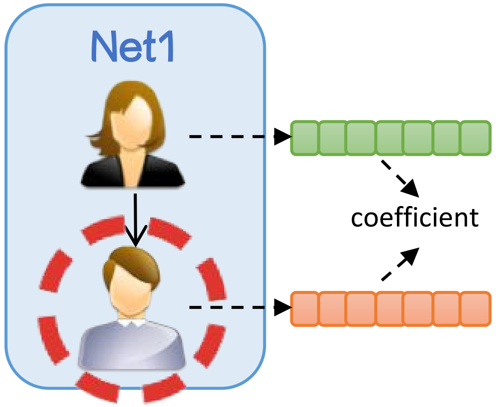
Intra-Network Initiator Attention

- To judge the importance of the **intra-network recipient neighbor** to the **target initiator.**



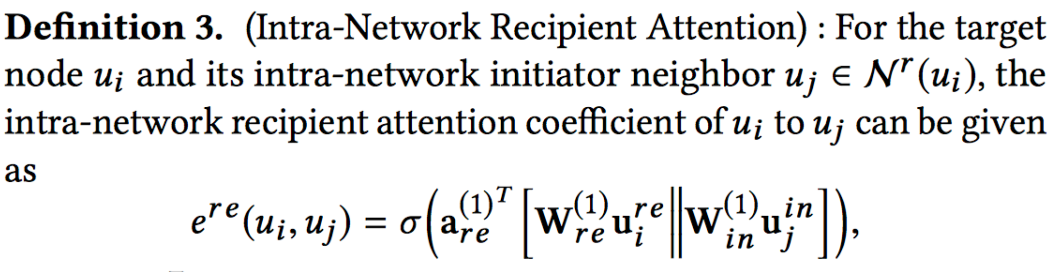
Definition 2. (Intra-Network Initiator Attention) : For the target node u_i and its intra-network recipient neighbor $u_j \in \mathcal{N}^i(u_i)$, the intra-network initiator attention coefficient of u_i to u_j can be given as

$$e^{in}(u_i,u_j) = \sigma\left(\mathbf{a}_{in}^{(1)^T}\left[\mathbf{W}_{in}^{(1)}\mathbf{u}_i^{in} \| \mathbf{W}_{re}^{(1)}\mathbf{u}_j^{re}\right]\right),$$



Intra-Network Recipient Attention

- To judge the importance of the **intra-network initiator neighbor** to the **target recipient.**

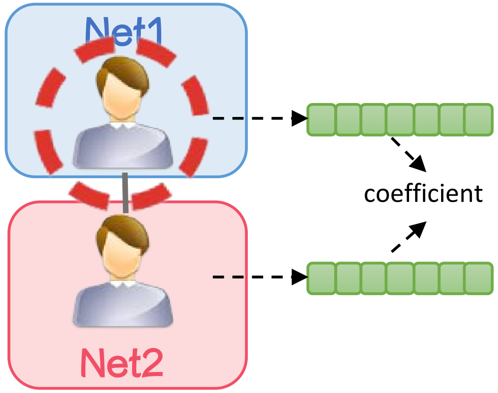


Definition 3. (Intra-Network Recipient Attention) : For the target node u_i and its intra-network initiator neighbor $u_j \in \mathcal{N}^r(u_i)$, the intra-network recipient attention coefficient of u_i to u_j can be given as

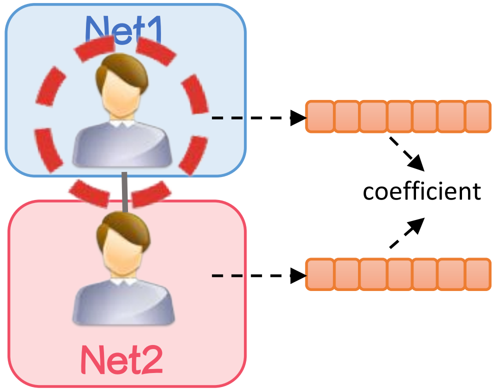
$$e^{re}(u_i,u_j) = \sigma\left(\mathbf{a}_{re}^{(1)^T}\left[\mathbf{W}_{re}^{(1)}\mathbf{u}_i^{re} \| \mathbf{W}_{in}^{(1)}\mathbf{u}_j^{in}\right]\right),$$

Inter-Network Anchor Attention

To handle the problem of network characteristic differences, the inter-network anchor attention is designed for the network level.



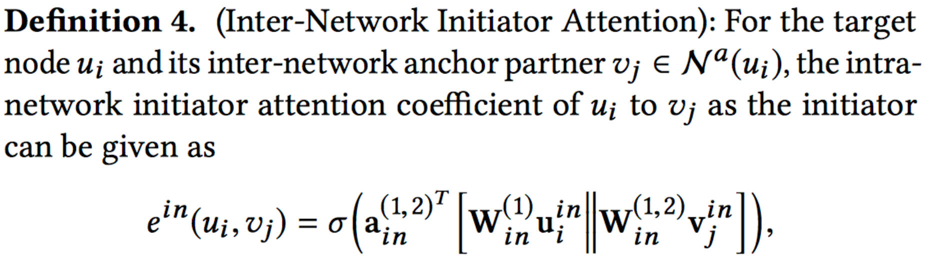
Inter-Network Initiator Attention



Inter-Network Recipient Attention

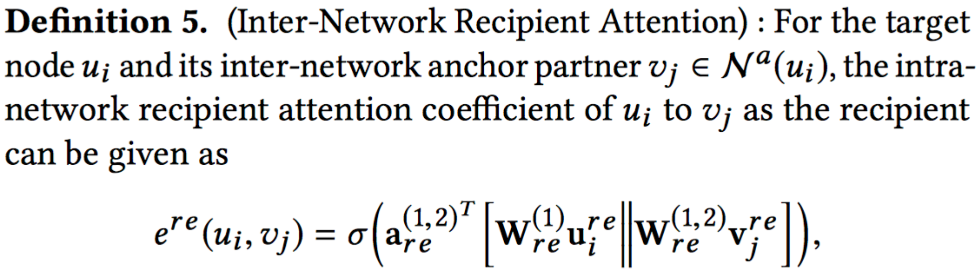
- To indicate the importance of information from other networks to the target node as the initiator.

- To indicate the importance of information from other networks to the target node as the recipient.



Definition 4. (Inter-Network Initiator Attention): For the target node u_i and its inter-network anchor partner $v_j \in \mathcal{N}^a(u_i)$, the intra-network initiator attention coefficient of u_i to v_j as the initiator can be given as

$$e^{in}(u_i, v_j) = \sigma\left(\mathbf{a}_{in}^{(1,2)^T}\left[\mathbf{W}_{in}^{(1)}\mathbf{u}_i^{in}\middle\|\mathbf{W}_{in}^{(1,2)}\mathbf{v}_j^{in}\right]\right),$$

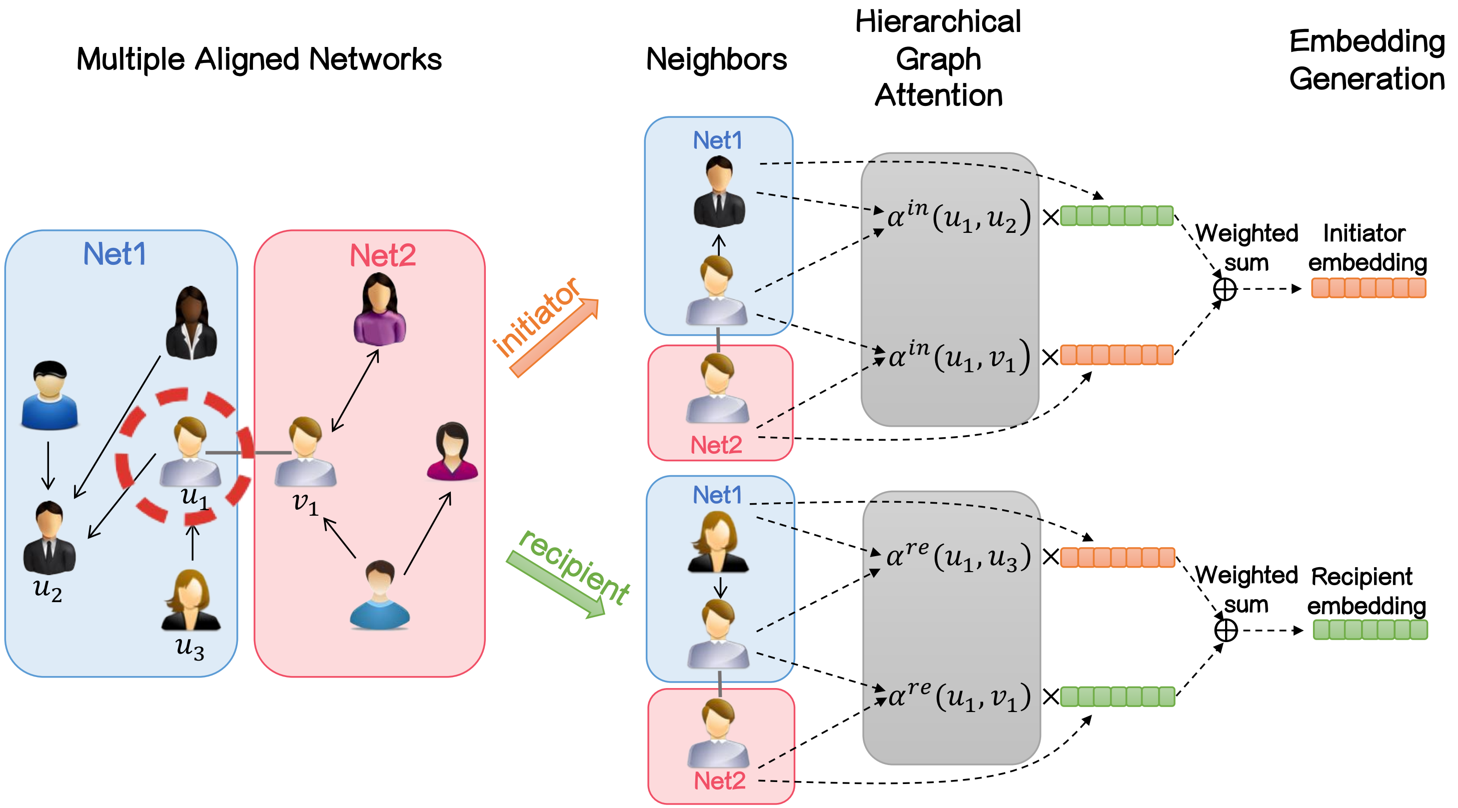


Definition 5. (Inter-Network Recipient Attention) : For the target node u_i and its inter-network anchor partner $v_j \in \mathcal{N}^a(u_i)$, the intra-network recipient attention coefficient of u_i to v_j as the recipient can be given as

$$e^{re}(u_i, v_j) = \sigma\left(\mathbf{a}_{re}^{(1,2)^T}\left[\mathbf{W}_{re}^{(1)}\mathbf{u}_i^{re}\middle\|\mathbf{W}_{re}^{(1,2)}\mathbf{v}_j^{re}\right]\right),$$

Multiple Aligned Network Embedding

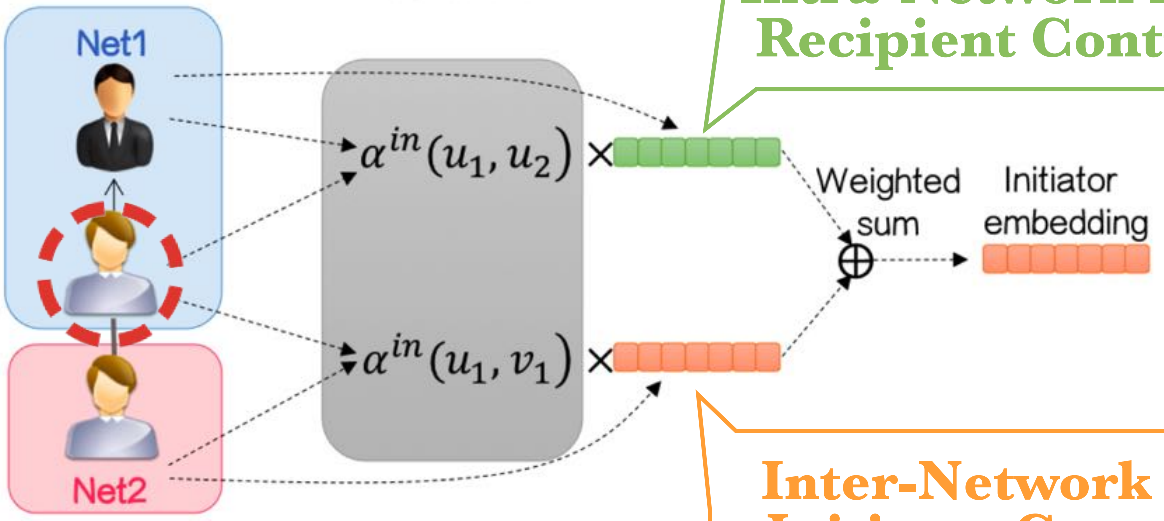
- HGANE extends the traditional graph neural network to the *multiple aligned social networks scenario*.
- HGANE learns the node representations by aggregating information from both the intra-network neighbors and inter-network partners.



- **capture the localized structural features** by utilizing information propagated from the intra-network social neighbors.
- **preserve more comprehensive features** by leveraging cross-network information transferred by anchor links.

Multiple Aligned Network Embedding

- Formally, each node is represented by two embeddings according to **its two roles** in social networks considering the problems of link directivity challenges.
- The initiator embedding indicates its features as the initiator in the social network, which depends on its intra-network recipient neighbors and inter-network anchor partners.

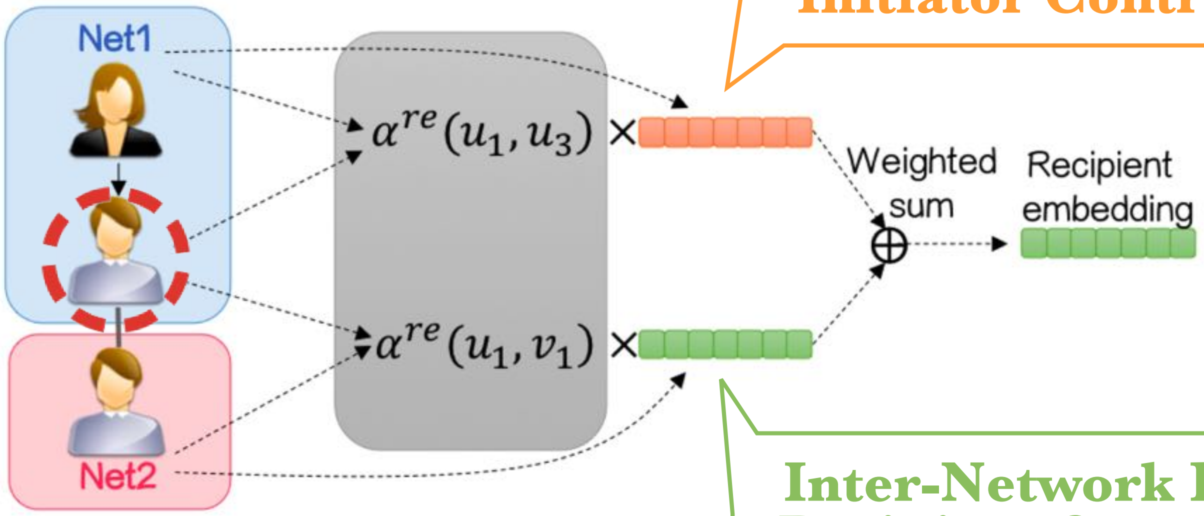


Intra-Network Neighbor Recipient Contribution

Inter-Network Partner Initiator Contribution

Multiple Aligned Network Embedding

- Formally, each node is represented by two embeddings according to **its two role** in social networks considering the problems of link directivity challenges.
- The recipient embedding indicates its features as the recipient in the social network, which depends on its intra-network initiator neighbors and inter-network anchor partners.



Intra-Network Neighbor Initiator Contribution

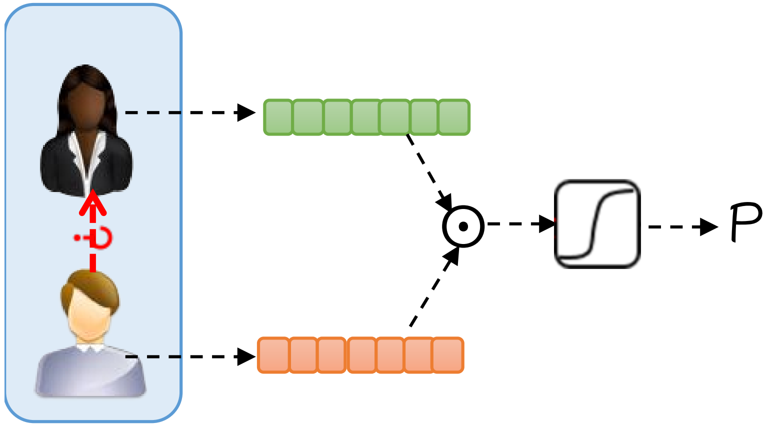
Inter-Network Partner Recipient Contribution

Collective Link Prediction Oriented Network Embedding Optimization Framework

- *The probability of the intra-network social link formation* from the initiator u_i pointing to the recipient u_j as

$$p(u_i, u_j) = \sigma(\mathbf{u}_i^{in^T} \cdot \mathbf{u}_j^{re})$$

Social Link Prediction



- The *objective* of intra-network social link formation

$$\mathcal{L}_{soc}(u_i, u_j) = \log p(u_i, u_j) + \sum_{\{(u_m, u_n)\}} \log(1 - p(u_m, u_n))$$

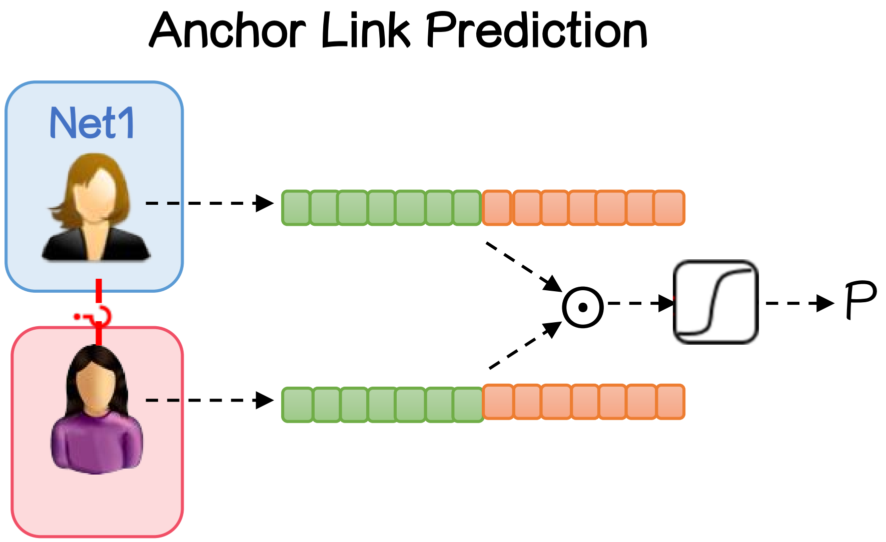
- The *final objective* for $G^{(1)}$

$$\mathcal{L}^{(1)} = \sum_{(u_i, u_j) \in E^{(1)}} \mathcal{L}_{soc}(u_i, u_j)$$

Collective Link Prediction Oriented Network Embedding Optimization Framework

- *The probability of the inter-network anchor link formation* between the target node u_i and the anchor node v_j

$$p(u_i, v_j) = \sigma\left(\left(\mathbf{u}_i^{in} \| \mathbf{u}_i^{re}\right)^T \cdot \left(\mathbf{v}_j^{in} \| \mathbf{v}_j^{re}\right)\right)$$



- The *objective* of inter-network anchor link formation

$$\mathcal{L}_{ach}(u_i, v_j) = \log p(u_i, v_j) + \sum_{\{(u_m, v_n)\}} \log\left(1 - p(u_m, v_n)\right)$$

- The *final objective* of network alignment

$$\mathcal{L}^{(1,2)} = \sum_{(u_i, v_j) \in A^{(1,2)}} \left(\mathcal{L}_{ach}(u_i, v_j)\right)$$

Collective Link Prediction Oriented Network Embedding Optimization Framework

- To incorporate the collective link prediction task into a unified framework, we learn the node representations with rich information by jointly training the objective function

$$\mathcal{L}\left(G^{(1)}, G^{(2)}\right) = \mathcal{L}^{(1)} + \mathcal{L}^{(2)} + \underline{\alpha} \cdot \mathcal{L}^{(1,2)} + \underline{\beta} \cdot \mathcal{L}_{reg}$$

    - the objective for the two social networks
    - the objective of network alignment
    - the regularization term

- Hyperparameter

Experiments

- Datasets
  - Twitter-Foursquare
  - Facebook-Twitter

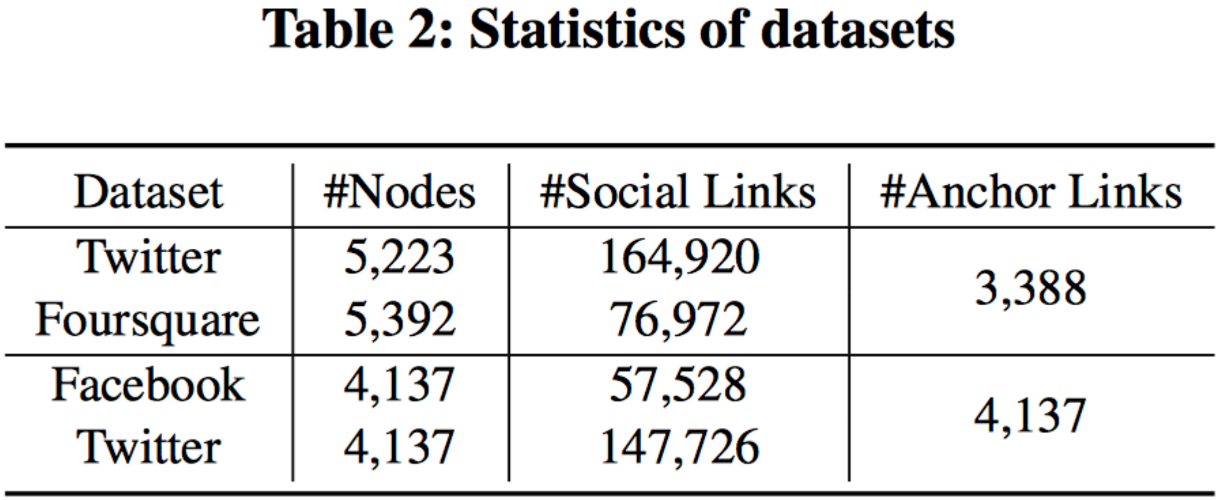


Table 2: Statistics of datasets

Dataset	#Nodes	#Social Links	#Anchor Links
Twitter	5,223	164,920	3,388
Foursquare	5,392	76,972	
Facebook	4,137	57,528	4,137
Twitter	4,137	147,726	

- Evaluation Metrics
  - AUC

Experiment Settings

- Comparison Methods
  - Deepwalk, Node2vec, GAT learn node embeddings in the single network
  - IONE, DIME, MNN study the single kind of link prediction task.
  - CLF proposes and studies the collective link prediction task.

	Multiple Networks	**Links Directi. Differe.**	**Network Charact. Differe.**	**Predict Social Link**	**Predict Anchor Link**
DeepWalk				√	
Node2Vec				√	
GAT				√	
IONE	√	√			√
DIME	√	√		√	
MNN	√	√		√	
CLF	√			√	√
HGANE	√	√	√	√	√

Experiment Results

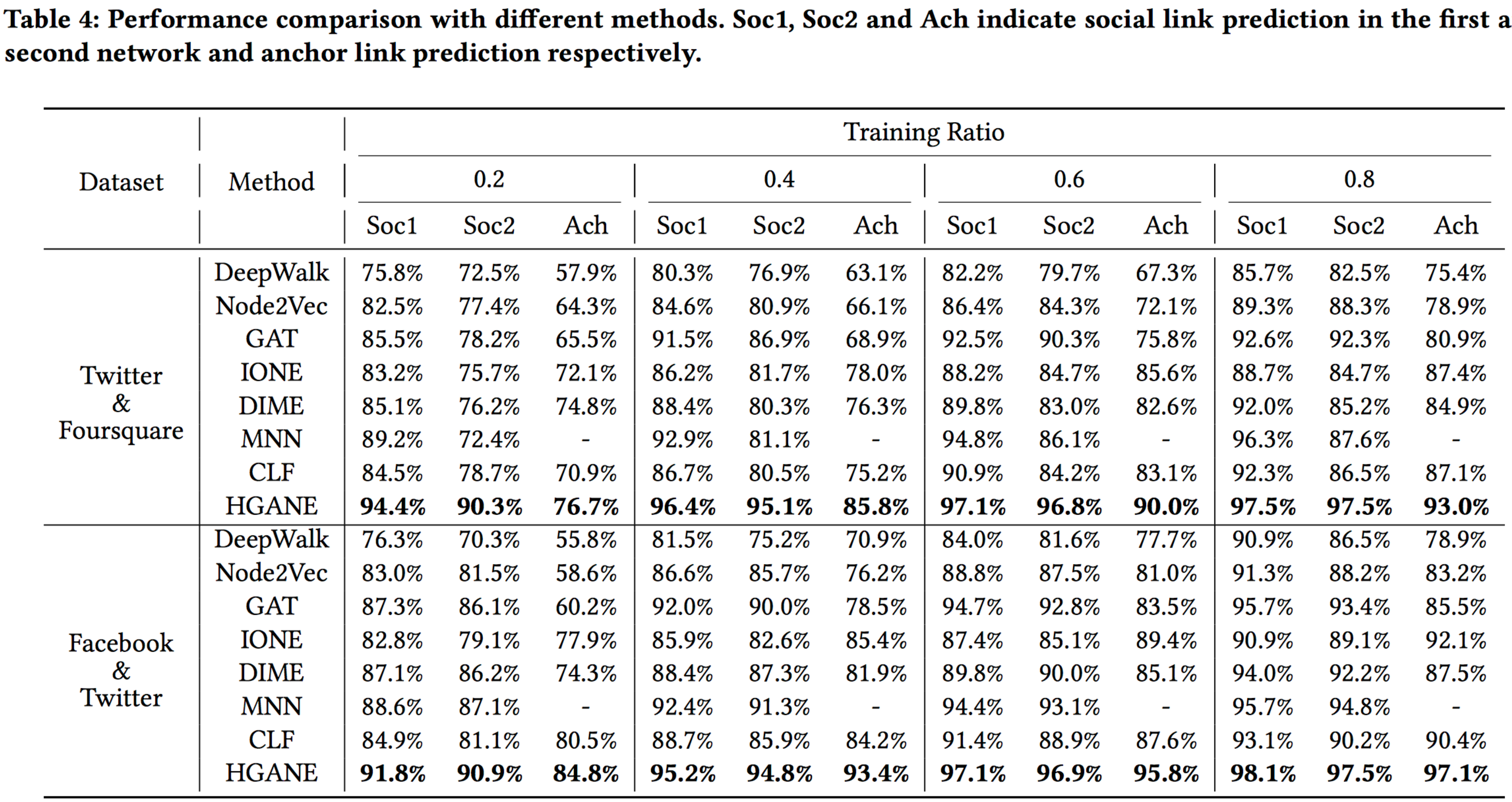
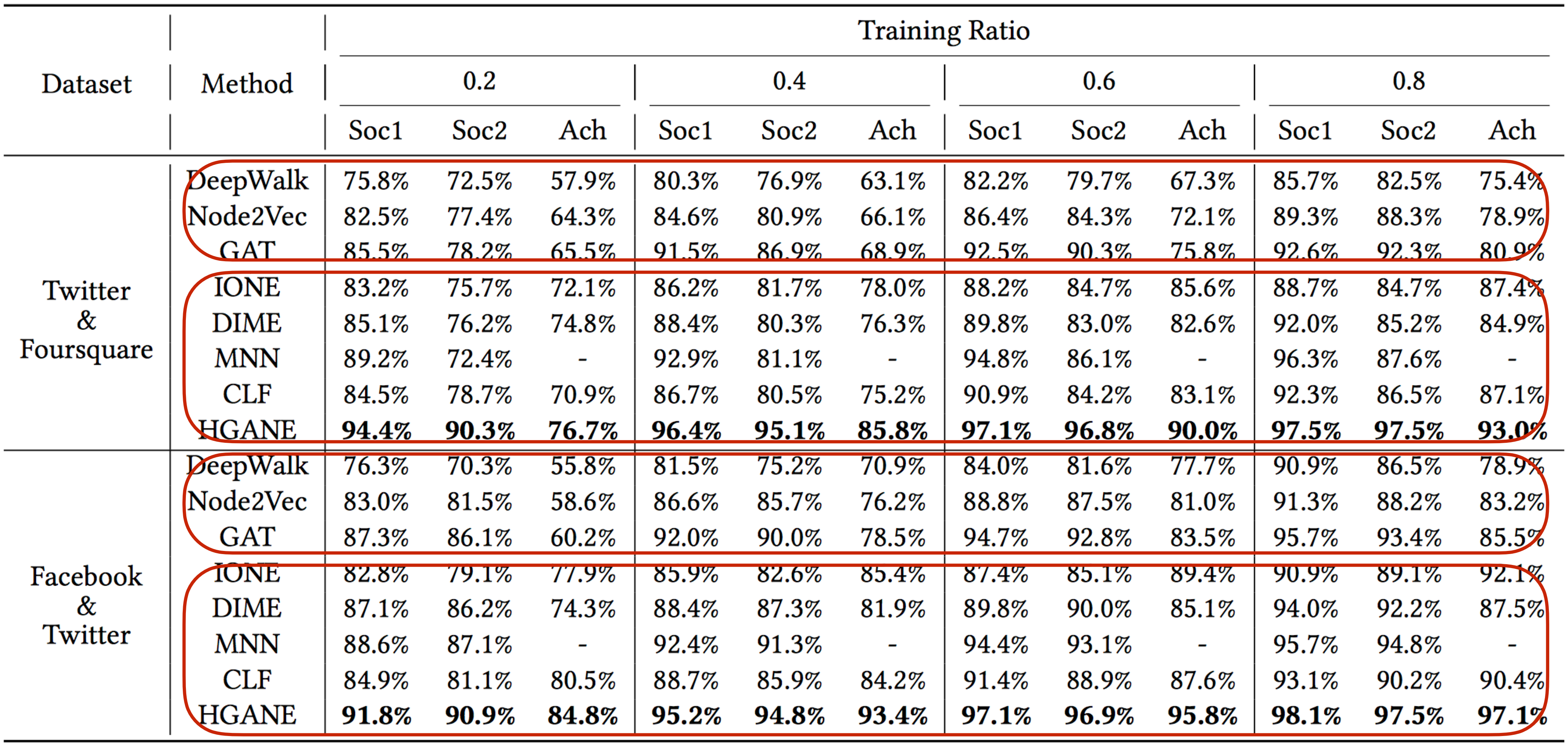


Table 4: Performance comparison with different methods. Soc1, Soc2 and Ach indicate social link prediction in the first and second network and anchor link prediction respectively.

Dataset	Method	Training Ratio											
		0.2			0.4			0.6			0.8		
		Soc1	Soc2	Ach	Soc1	Soc2	Ach	Soc1	Soc2	Ach	Soc1	Soc2	Ach
Twitter & Foursquare	DeepWalk	75.8%	72.5%	57.9%	80.3%	76.9%	63.1%	82.2%	79.7%	67.3%	85.7%	82.5%	75.4%
	Node2Vec	82.5%	77.4%	64.3%	84.6%	80.9%	66.1%	86.4%	84.3%	72.1%	89.3%	88.3%	78.9%
	GAT	85.5%	78.2%	65.5%	91.5%	86.9%	68.9%	92.5%	90.3%	75.8%	92.6%	92.3%	80.9%
	IONE	83.2%	75.7%	72.1%	86.2%	81.7%	78.0%	88.2%	84.7%	85.6%	88.7%	84.7%	87.4%
	DIME	85.1%	76.2%	74.8%	88.4%	80.3%	76.3%	89.8%	83.0%	82.6%	92.0%	85.2%	84.9%
	MNN	89.2%	72.4%	-	92.9%	81.1%	-	94.8%	86.1%	-	96.3%	87.6%	-
	CLF	84.5%	78.7%	70.9%	86.7%	80.5%	75.2%	90.9%	84.2%	83.1%	92.3%	86.5%	87.1%
	HGANE	**94.4%**	**90.3%**	**76.7%**	**96.4%**	**95.1%**	**85.8%**	**97.1%**	**96.8%**	**90.0%**	**97.5%**	**97.5%**	**93.0%**
Facebook & Twitter	DeepWalk	76.3%	70.3%	55.8%	81.5%	75.2%	70.9%	84.0%	81.6%	77.7%	90.9%	86.5%	78.9%
	Node2Vec	83.0%	81.5%	58.6%	86.6%	85.7%	76.2%	88.8%	87.5%	81.0%	91.3%	88.2%	83.2%
	GAT	87.3%	86.1%	60.2%	92.0%	90.0%	78.5%	94.7%	92.8%	83.5%	95.7%	93.4%	85.5%
	IONE	82.8%	79.1%	77.9%	85.9%	82.6%	85.4%	87.4%	85.1%	89.4%	90.9%	89.1%	92.1%
	DIME	87.1%	86.2%	74.3%	88.4%	87.3%	81.9%	89.8%	90.0%	85.1%	94.0%	92.2%	87.5%
	MNN	88.6%	87.1%	-	92.4%	91.3%	-	94.4%	93.1%	-	95.7%	94.8%	-
	CLF	84.9%	81.1%	80.5%	88.7%	85.9%	84.2%	91.4%	88.9%	87.6%	93.1%	90.2%	90.4%
	HGANE	**91.8%**	**90.9%**	**84.8%**	**95.2%**	**94.8%**	**93.4%**	**97.1%**	**96.9%**	**95.8%**	**98.1%**	**97.5%**	**97.1%**

Experiment Results

Table 4: Performance comparison with different methods. Soc1, Soc2 and Ach indicate social link prediction in the first and second network and anchor link prediction respectively.



Dataset	Method	Training Ratio											
		0.2			0.4			0.6			0.8		
		Soc1	Soc2	Ach	Soc1	Soc2	Ach	Soc1	Soc2	Ach	Soc1	Soc2	Ach
Twitter & Foursquare	DeepWalk	75.8%	72.5%	57.9%	80.3%	76.9%	63.1%	82.2%	79.7%	67.3%	85.7%	82.5%	75.4%
	Node2Vec	82.5%	77.4%	64.3%	84.6%	80.9%	66.1%	86.4%	84.3%	72.1%	89.3%	88.3%	78.9%
	GAT	85.5%	78.2%	65.5%	91.5%	86.9%	68.9%	92.5%	90.3%	75.8%	92.6%	92.3%	80.9%
	IONE	83.2%	75.7%	72.1%	86.2%	81.7%	78.0%	88.2%	84.7%	85.6%	88.7%	84.7%	87.4%
	DIME	85.1%	76.2%	74.8%	88.4%	80.3%	76.3%	89.8%	83.0%	82.6%	92.0%	85.2%	84.9%
	MNN	89.2%	72.4%	-	92.9%	81.1%	-	94.8%	86.1%	-	96.3%	87.6%	-
	CLF	84.5%	78.7%	70.9%	86.7%	80.5%	75.2%	90.9%	84.2%	83.1%	92.3%	86.5%	87.1%
	HGANE	**94.4%**	**90.3%**	**76.7%**	**96.4%**	**95.1%**	**85.8%**	**97.1%**	**96.8%**	**90.0%**	**97.5%**	**97.5%**	**93.0%**
Facebook & Twitter	DeepWalk	76.3%	70.3%	55.8%	81.5%	75.2%	70.9%	84.0%	81.6%	77.7%	90.9%	86.5%	78.9%
	Node2Vec	83.0%	81.5%	58.6%	86.6%	85.7%	76.2%	88.8%	87.5%	81.0%	91.3%	88.2%	83.2%
	GAT	87.3%	86.1%	60.2%	92.0%	90.0%	78.5%	94.7%	92.8%	83.5%	95.7%	93.4%	85.5%
	IONE	82.8%	79.1%	77.9%	85.9%	82.6%	85.4%	87.4%	85.1%	89.4%	90.9%	89.1%	92.1%
	DIME	87.1%	86.2%	74.3%	88.4%	87.3%	81.9%	89.8%	90.0%	85.1%	94.0%	92.2%	87.5%
	MNN	88.6%	87.1%	-	92.4%	91.3%	-	94.4%	93.1%	-	95.7%	94.8%	-
	CLF	84.9%	81.1%	80.5%	88.7%	85.9%	84.2%	91.4%	88.9%	87.6%	93.1%	90.2%	90.4%
	HGANE	**91.8%**	**90.9%**	**84.8%**	**95.2%**	**94.8%**	**93.4%**	**97.1%**	**96.9%**	**95.8%**	**98.1%**	**97.5%**	**97.1%**

- The methods studying multiple aligned networks alleviate the data insufficiency problem.

Experiment Results

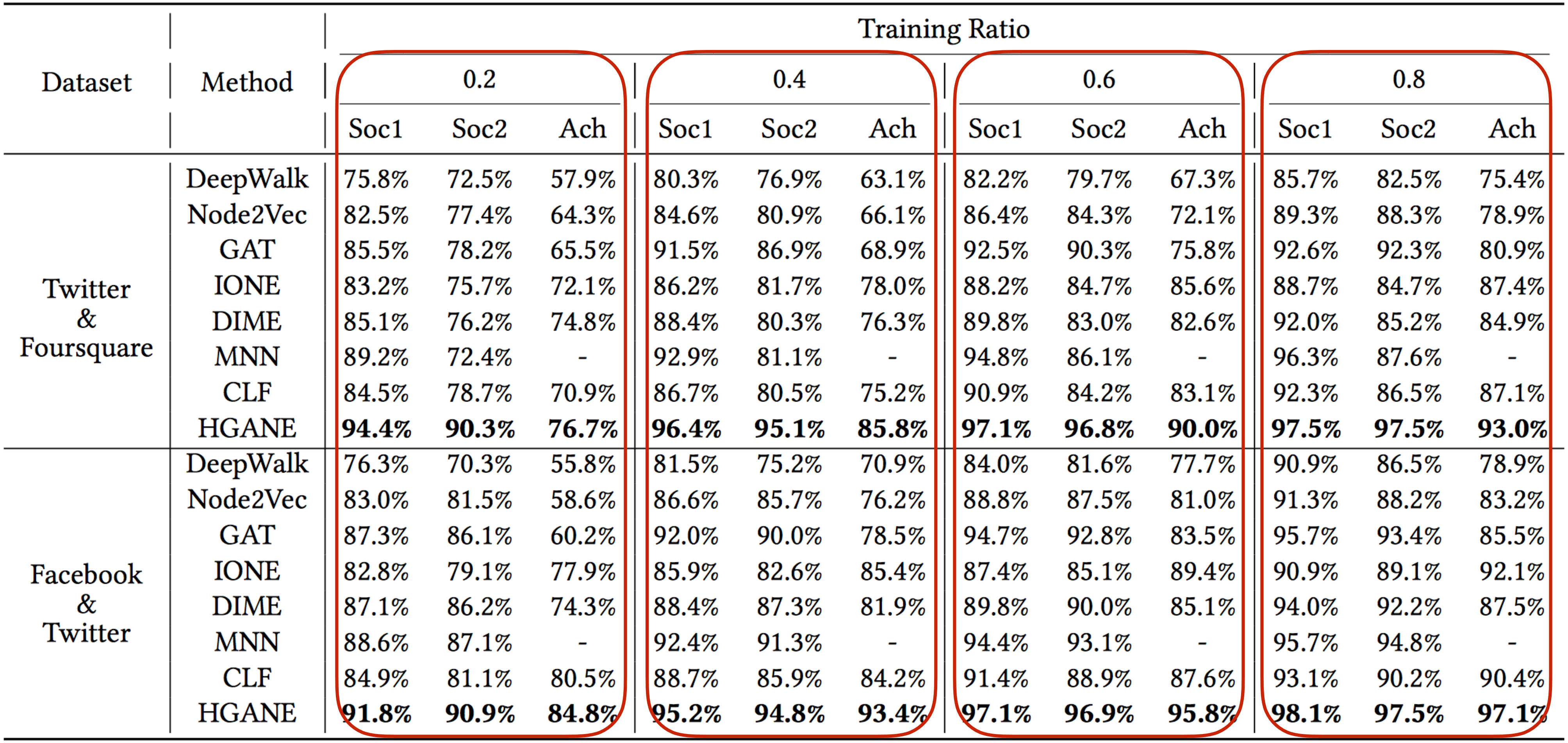


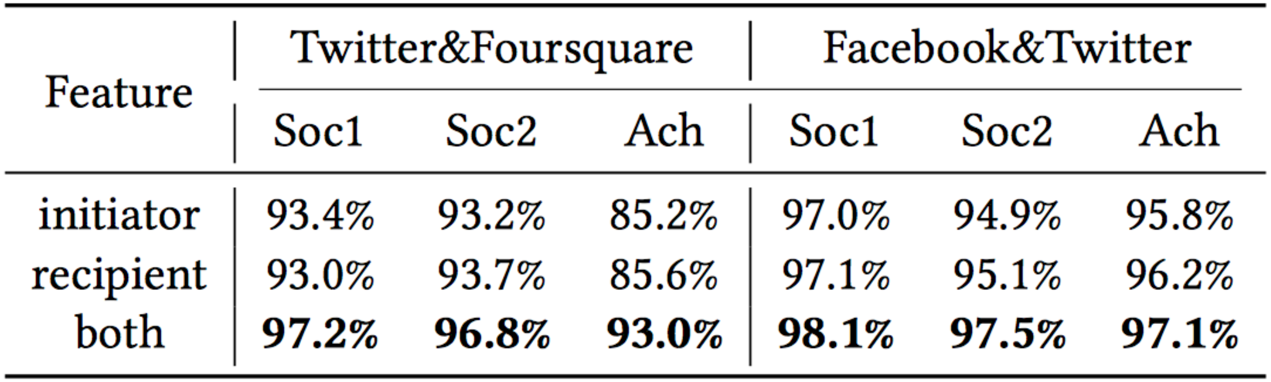
Table 4: Performance comparison with different methods. Soc1, Soc2 and Ach indicate social link prediction in the first and second network and anchor link prediction respectively.

Dataset	Method	Training Ratio											
		0.2			0.4			0.6			0.8		
		Soc1	Soc2	Ach	Soc1	Soc2	Ach	Soc1	Soc2	Ach	Soc1	Soc2	Ach
Twitter & Foursquare	DeepWalk	75.8%	72.5%	57.9%	80.3%	76.9%	63.1%	82.2%	79.7%	67.3%	85.7%	82.5%	75.4%
	Node2Vec	82.5%	77.4%	64.3%	84.6%	80.9%	66.1%	86.4%	84.3%	72.1%	89.3%	88.3%	78.9%
	GAT	85.5%	78.2%	65.5%	91.5%	86.9%	68.9%	92.5%	90.3%	75.8%	92.6%	92.3%	80.9%
	IONE	83.2%	75.7%	72.1%	86.2%	81.7%	78.0%	88.2%	84.7%	85.6%	88.7%	84.7%	87.4%
	DIME	85.1%	76.2%	74.8%	88.4%	80.3%	76.3%	89.8%	83.0%	82.6%	92.0%	85.2%	84.9%
	MNN	89.2%	72.4%	-	92.9%	81.1%	-	94.8%	86.1%	-	96.3%	87.6%	-
	CLF	84.5%	78.7%	70.9%	86.7%	80.5%	75.2%	90.9%	84.2%	83.1%	92.3%	86.5%	87.1%
	HGANE	**94.4%**	**90.3%**	**76.7%**	**96.4%**	**95.1%**	**85.8%**	**97.1%**	**96.8%**	**90.0%**	**97.5%**	**97.5%**	**93.0%**
Facebook & Twitter	DeepWalk	76.3%	70.3%	55.8%	81.5%	75.2%	70.9%	84.0%	81.6%	77.7%	90.9%	86.5%	78.9%
	Node2Vec	83.0%	81.5%	58.6%	86.6%	85.7%	76.2%	88.8%	87.5%	81.0%	91.3%	88.2%	83.2%
	GAT	87.3%	86.1%	60.2%	92.0%	90.0%	78.5%	94.7%	92.8%	83.5%	95.7%	93.4%	85.5%
	IONE	82.8%	79.1%	77.9%	85.9%	82.6%	85.4%	87.4%	85.1%	89.4%	90.9%	89.1%	92.1%
	DIME	87.1%	86.2%	74.3%	88.4%	87.3%	81.9%	89.8%	90.0%	85.1%	94.0%	92.2%	87.5%
	MNN	88.6%	87.1%	-	92.4%	91.3%	-	94.4%	93.1%	-	95.7%	94.8%	-
	CLF	84.9%	81.1%	80.5%	88.7%	85.9%	84.2%	91.4%	88.9%	87.6%	93.1%	90.2%	90.4%
	HGANE	**91.8%**	**90.9%**	**84.8%**	**95.2%**	**94.8%**	**93.4%**	**97.1%**	**96.9%**	**95.8%**	**98.1%**	**97.5%**	**97.1%**

- As the training rate λ drops, the performance degradation of HGANE is rather moderate since we handle the problem of network characteristic differences.

Experiment Results

- Validation of the design of represent each node with two embeddings to resolve the link directivity differences problem.

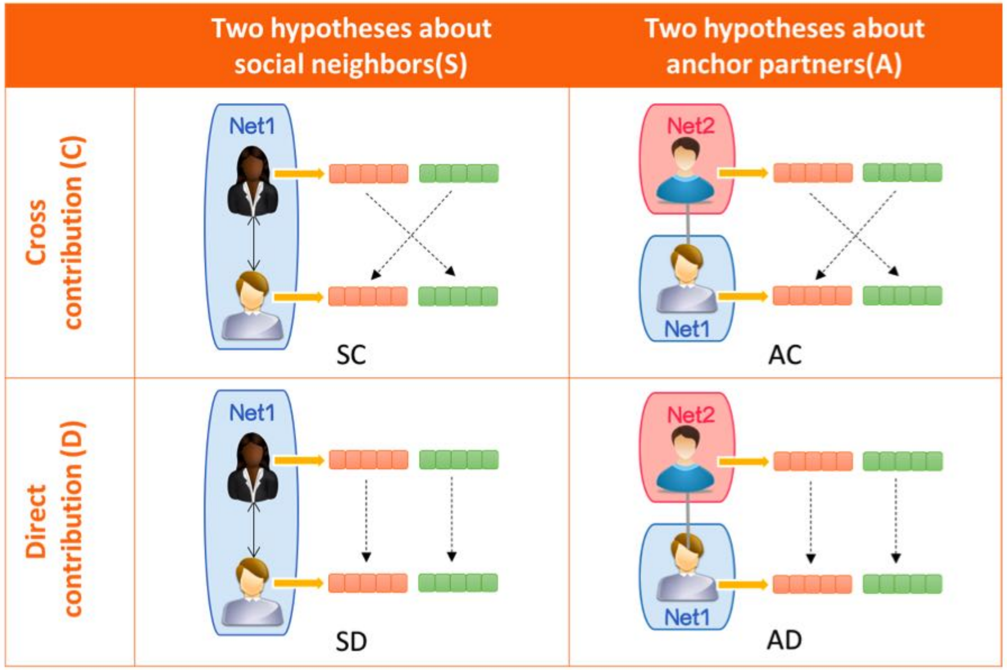


Feature	Twitter&Foursquare			Facebook&Twitter		
	Soc1	Soc2	Ach	Soc1	Soc2	Ach
initiator	93.4%	93.2%	85.2%	97.0%	94.9%	95.8%
recipient	93.0%	93.7%	85.6%	97.1%	95.1%	96.2%
both	**97.2%**	**96.8%**	**93.0%**	**98.1%**	**97.5%**	**97.1%**

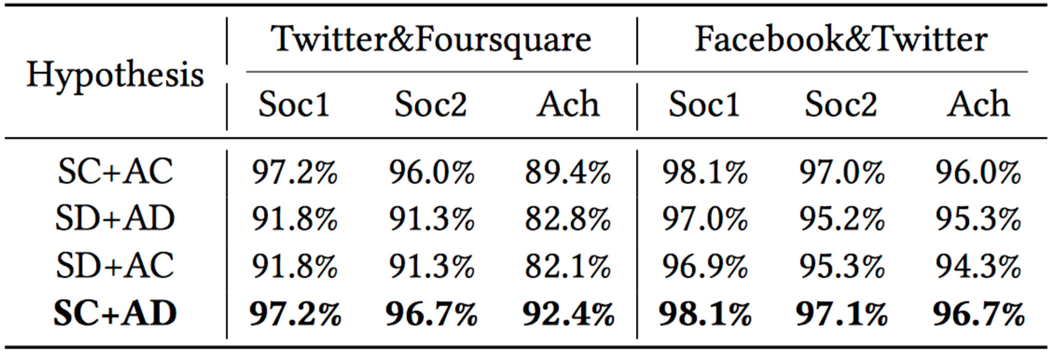
- Our full model outperforms two variants with either the initiator or recipient features.
- The performance of social link prediction within the more dense network can be improved more by distinguishing nodes' initiator and recipient roles.

Hypothesis Verification

- Four hypotheses about how the the target node aggregates information from social neighbors and anchor partners.

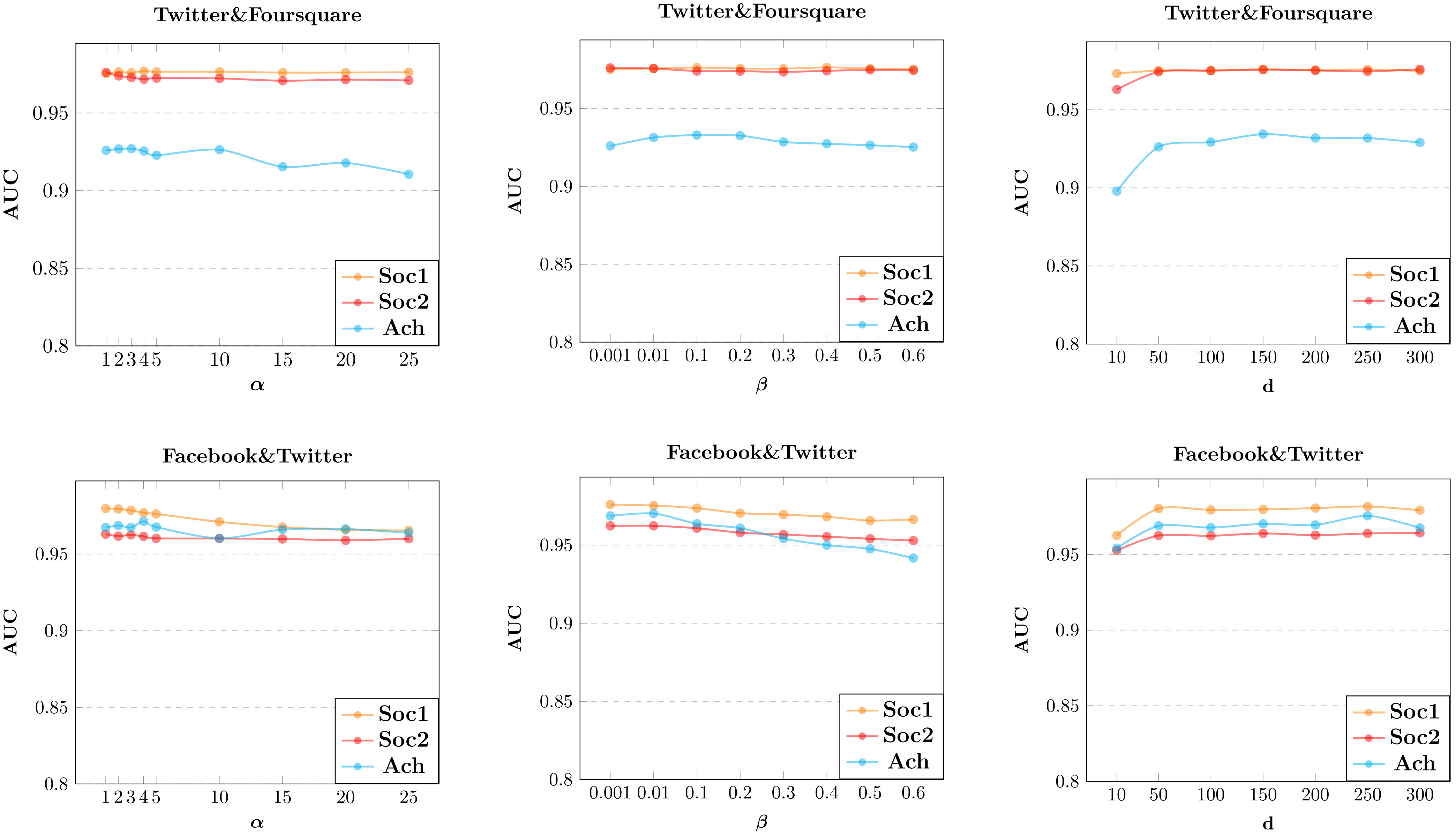


- The hypotheses SC+AD is adopted in our framework.
- If DC is replaced with CC, AUC of the anchor link prediction decreases a lot.
- AUC of the social link prediction are affected if we adopt DC for social links.



Hypothesis	Twitter&Foursquare			Facebook&Twitter		
	Soc1	Soc2	Ach	Soc1	Soc2	Ach
SC+AC	97.2%	96.0%	89.4%	98.1%	97.0%	96.0%
SD+AD	91.8%	91.3%	82.8%	97.0%	95.2%	95.3%
SD+AC	91.8%	91.3%	82.1%	96.9%	95.3%	94.3%
SC+AD	**97.2%**	**96.7%**	**92.4%**	**98.1%**	**97.1%**	**96.7%**

Hyperparameter analysis



Summary

- In this paper, we study the Collective Link Prediction problem over multiple aligned social networks
- Proposed a task oriented network embedding framework with hierarchical graph attention
  1. learn node representation by aggregating information from both the intra-network neighbors and inter-network partners
  2. hierarchical graph attention mechanism to resolve the network characteristic differences and link directivity challenges
  3. incorporate the collective link prediction task objectives into consideration and balances between different prediction tasks
- Conduct extensive experiments on two real-world datasets to test the effectiveness of HGANE

Collective Link Prediction Oriented Network Embedding with Hierarchical Graph Attention

Q&A

Yizhu Jiao[1], Yun Xiong[1], Jiawei Zhang[2] and Yangyong Zhu[1]
{yzjiao18, yunx, yyzhu}@fudan.edu.cn, jiawei@ifmlab.org

Code: https://github.com/yzjiao/HierarchicalGraphAttention