Multiplex Heterogeneous Graph Convolutional Network

Pengyang Yu
Ocean University of China
Qingdao, China
ypy@stu.ouc.edu.cn

Chaofan Fu
Ocean University of China
Qingdao, China
fuchaofan@stu.ouc.edu.cn

Yanwei Yu*
Ocean University of China
Qingdao, China
yuyanwei@ouc.edu.cn

Chao Huang*
The University of Hong Kong
Hong Kong, China
chaohuang75@gmail.com

Zhongying Zhao
Shandong University of Science and Technology
Qingdao, China
zyzhao@sdust.edu.cn

Junyu Dong
Ocean University of China
Qingdao, China
dongjunyu@ouc.edu.cn

—KDD 2022

Code: https://github.com/NSSSJSS/MHGCN.

Reported by Yabo Yin

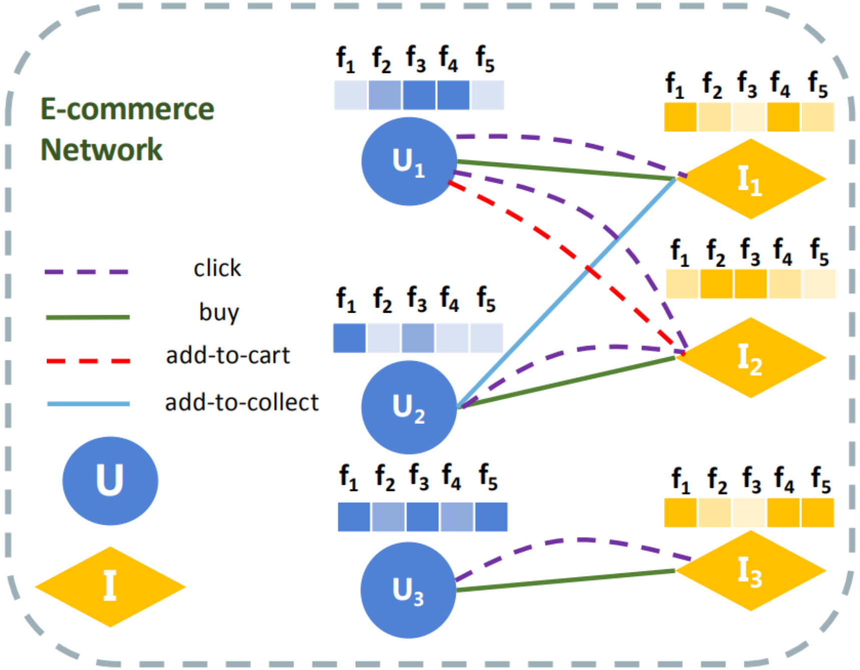
1.Introduction

2.Method

3.Experiments

Introduction

1. Most existing works ignore the relation heterogeneity with multiplex network between multi-typed nodes and the different importance of relations in meta-paths for node embedding.
2. Previous works cannot effectively preserve the heterogeneous and multiplex graph characteristics for the network representation task.



Method

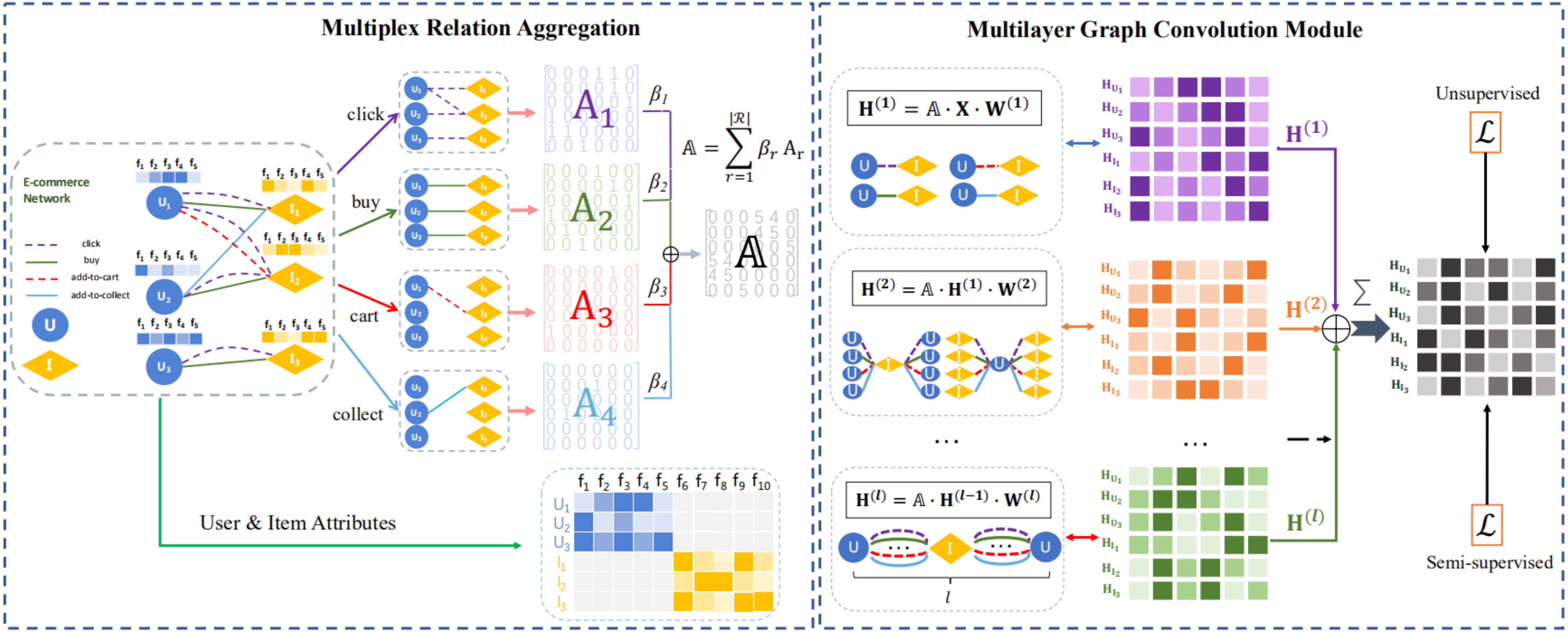


Figure 1: The overview of the proposed MHGCN.

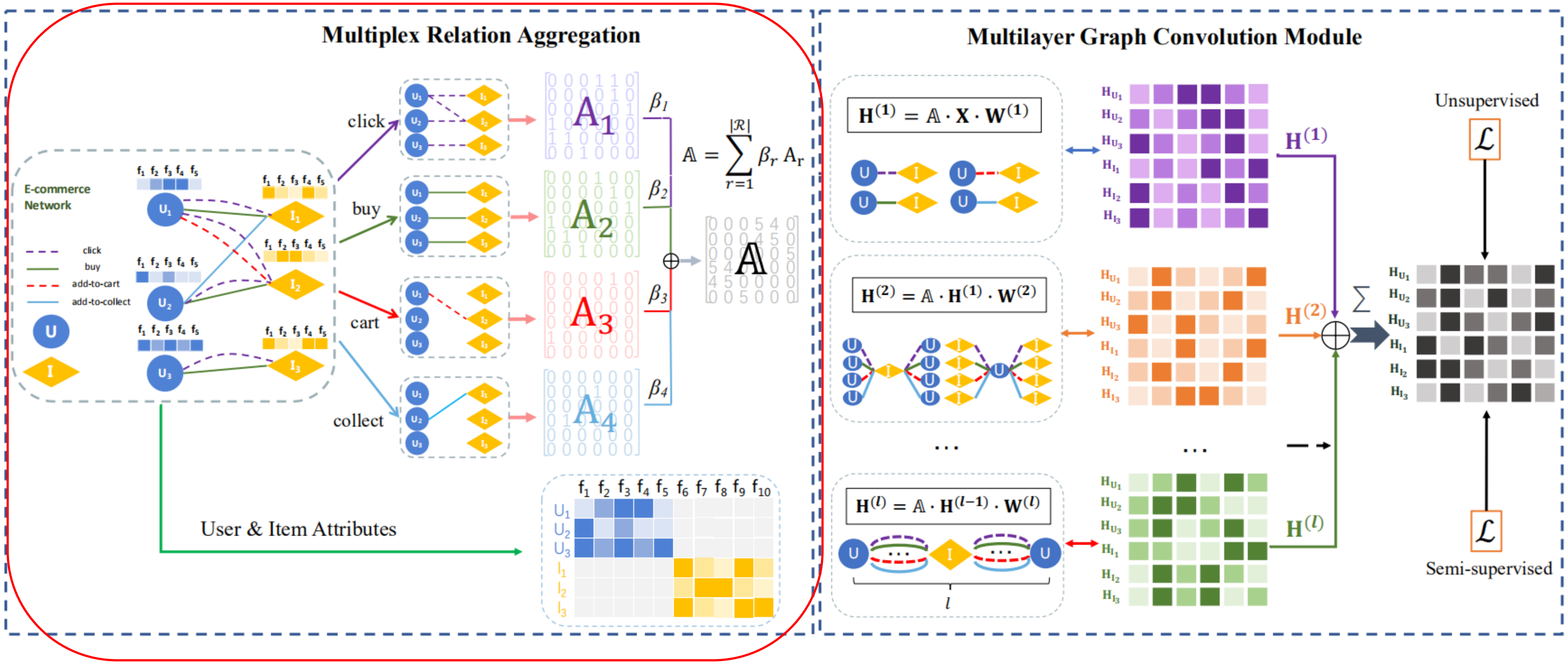


Figure 1: The overview of the proposed MHGCN.

$\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ $\mathbf{X} \in \mathbb{R}^{n \times m}$.

same pair of nodes can be connected through multi-typed edges.

sub-graph as $\{\mathcal{G}_r | r = 1, 2, \dots, |\mathcal{R}|\}$

corresponding adjacent matrix $\{\mathbf{A}_r | r = 1, 2, \dots, |\mathcal{R}|\}$.

$\mathbb{A} = \sum_{r=1}^{|\mathcal{R}|} \beta_r \mathbf{A}_r$ $\{\beta_r | r = 1, 2, \dots, |\mathcal{R}|\}$

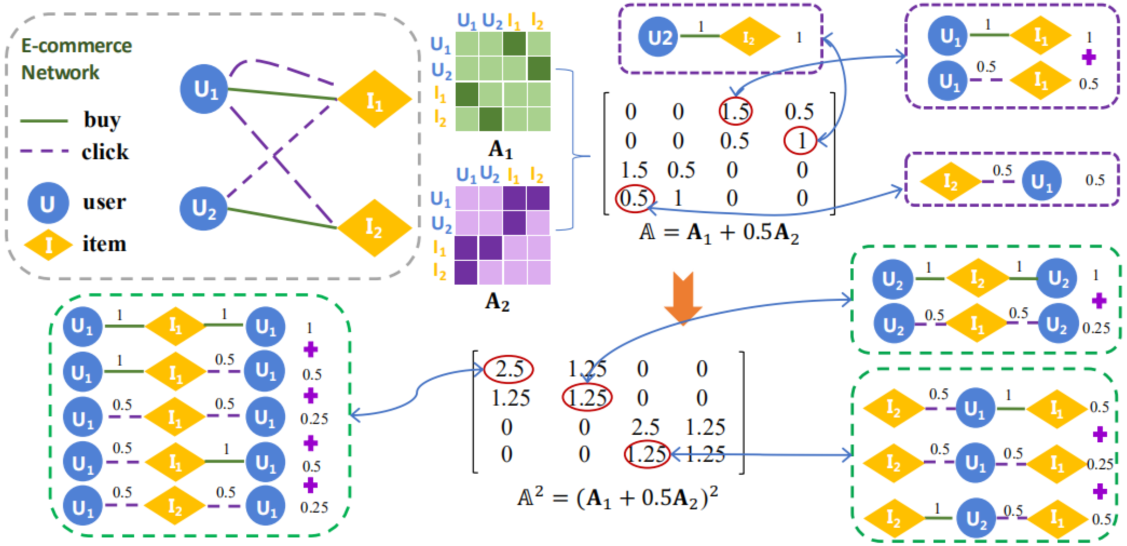


Figure 2: Illustration of meta-paths with importance for a toy example

$$H^{(1)} = \mathbb{A} \cdot X \cdot W^{(1)}, \qquad (1)$$

$$H^{(1)} \in \mathbb{R}^{n\times d} \quad X \in \mathbb{R}^{n\times m} \quad W^{(1)} \in \mathbb{R}^{m\times d}$$

$$\begin{aligned} H^{(2)} &= \mathbb{A} \cdot H^{(1)} \cdot W^{(2)} \\ &= \mathbb{A} \cdot (\mathbb{A} \cdot X \cdot W^{(1)}) \cdot W^{(2)} \qquad (2) \\ &= \mathbb{A}^2 \cdot X \cdot W^{(1)} \cdot W^{(2)}, \end{aligned}$$

$$W^{(2)} \in \mathbb{R}^{d\times d}$$

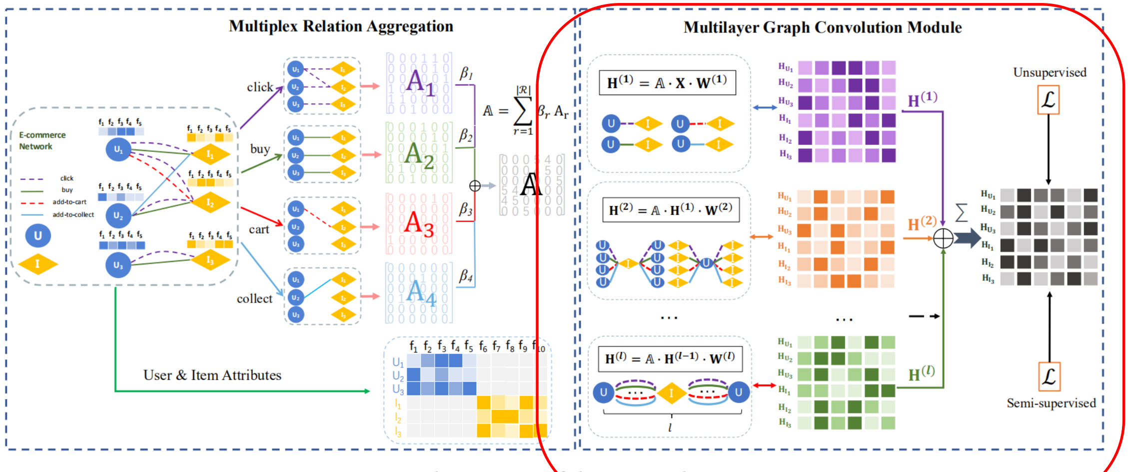


Figure 1: The overview of the proposed MHGCN.

$$H = \frac{1}{2}(H^{(1)} + H^{(2)}). \qquad (3)$$

$$\begin{aligned} H^{(l)} &= \mathbb{A} \cdot H^{(l-1)} \cdot W^{(l)} \\ &= \mathbb{A} \cdot (\mathbb{A} \cdot H^{(l-2)} \cdot W^{(l-1)}) \cdot W^{(l)} \\ &= \underbrace{\mathbb{A} \cdots (\mathbb{A}}_{l} \cdot X \cdot \underbrace{W^{(1)}) \cdots W^{(l)}}_{l} \qquad (4) \\ &= \mathbb{A}^{l} \cdot X \cdot \underbrace{W^{(1)} \cdots W^{(l)}}_{l} \end{aligned}$$

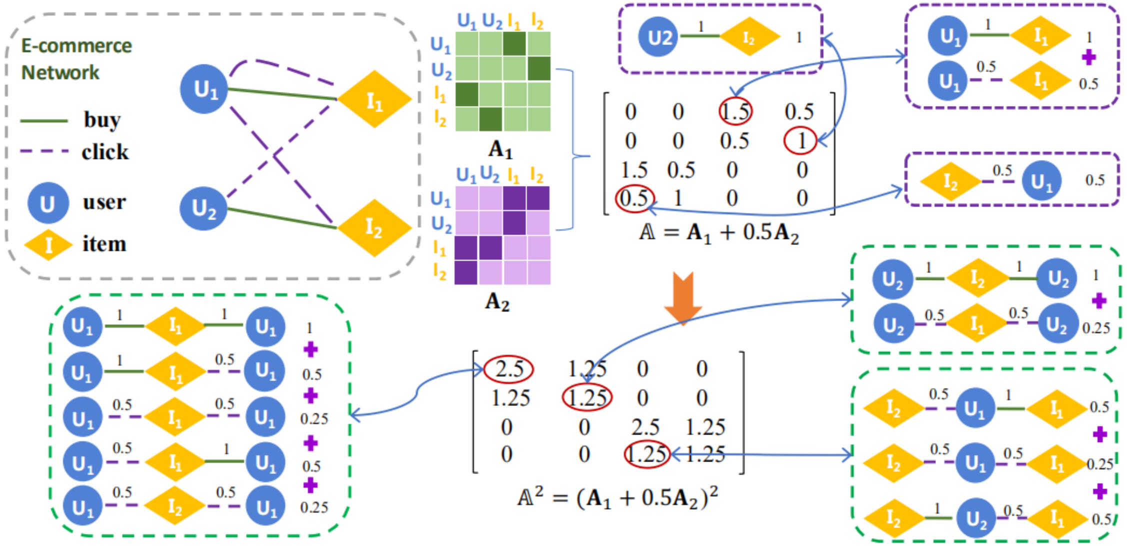
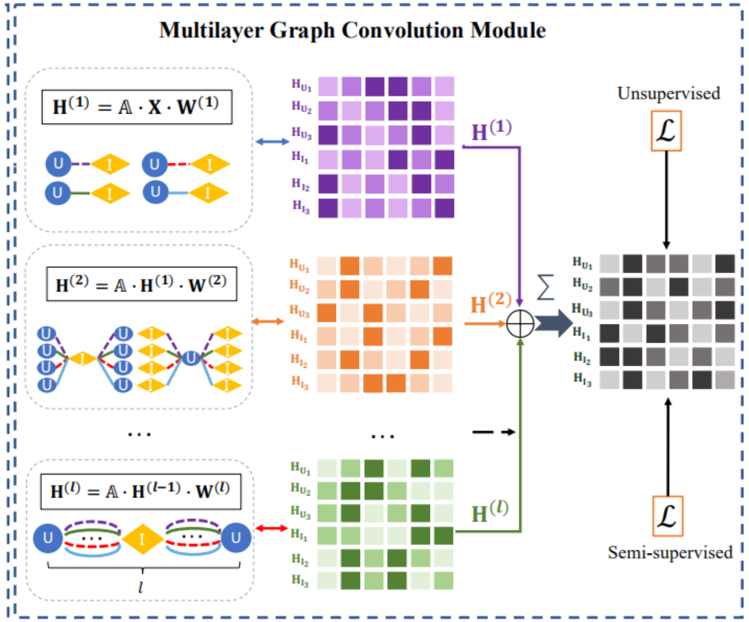


Figure 2: Illustration of meta-paths with importance for a toy example



$$\mathbf{H} = \frac{1}{l}\sum_{i=1}^{l} \mathbf{H}^{(i)}$$
$$= \frac{1}{l}\sum_{i=1}^{l} \mathbb{A}\cdot \mathbf{H}^{(i-1)}\cdot \mathbf{W}^{(i)}, \qquad (5)$$

where $\mathbf{H}^{(0)}$ is the node attribute matrix $\mathbf{X}$.

$$\mathcal{L} = -\sum_{(u,v)\in\Omega} \log\sigma(<\mathbf{H}_u^{\mathsf{T}}, \mathbf{H}_v>) - \sum_{(u',v')\in\Omega^-} \log\sigma(-<\mathbf{H}_{u'}^{\mathsf{T}}, \mathbf{H}_{v'}>), \qquad (6)$$

$$\mathcal{L} = -\sum_{i\in\mathcal{V}_{ids}} \mathbf{Y}_i \ln(\mathbf{C}\cdot\mathbf{H}_i), \qquad (7)$$

Experiments

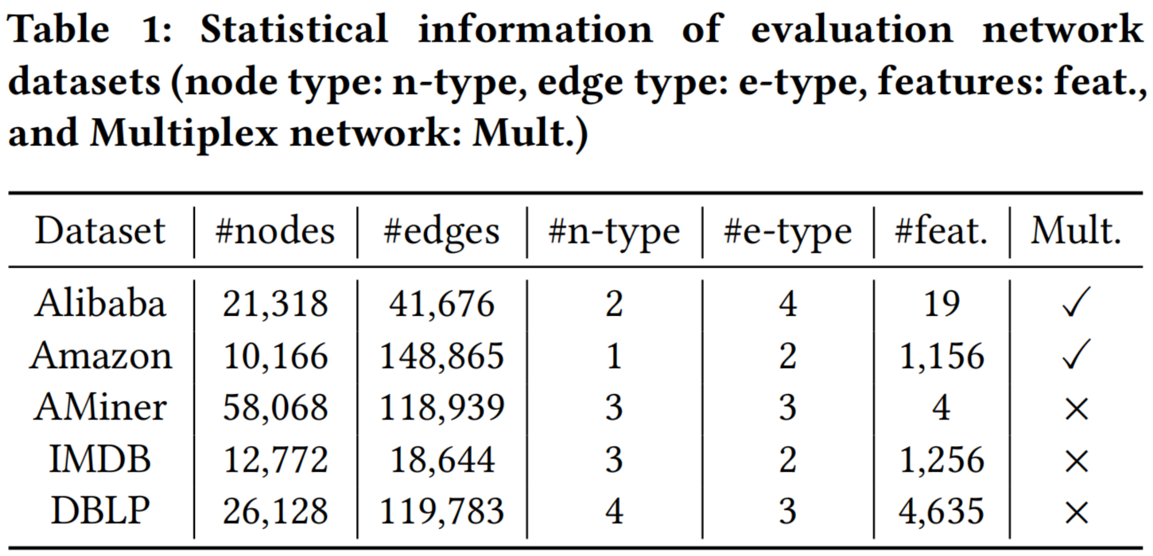
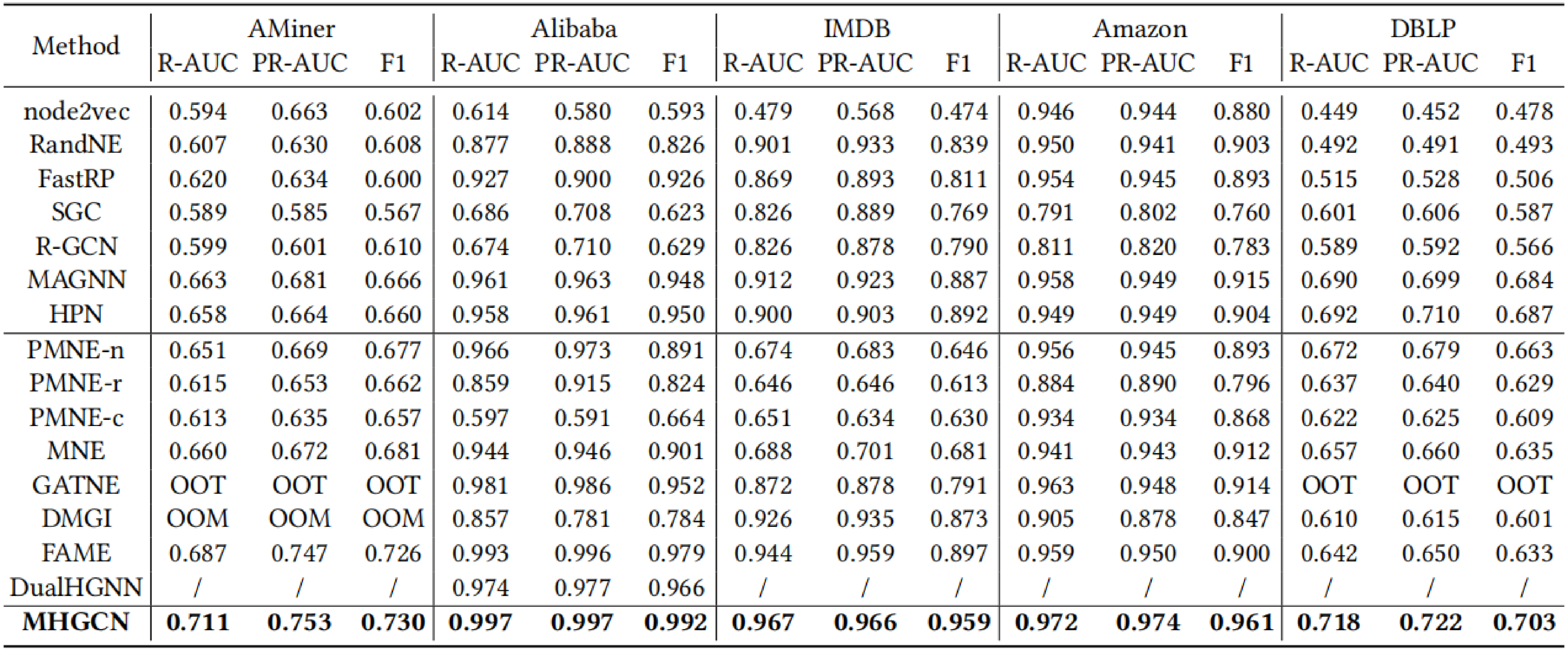


Table 1: Statistical information of evaluation network datasets (node type: n-type, edge type: e-type, features: feat., and Multiplex network: Mult.)

Dataset	#nodes	#edges	#n-type	#e-type	#feat.	Mult.
Alibaba	21,318	41,676	2	4	19	✓
Amazon	10,166	148,865	1	2	1,156	✓
AMiner	58,068	118,939	3	3	4	×
IMDB	12,772	18,644	3	2	1,256	×
DBLP	26,128	119,783	4	3	4,635	×

Experiments

Table 2: Model performance comparison for the task of link prediction on different datasets.

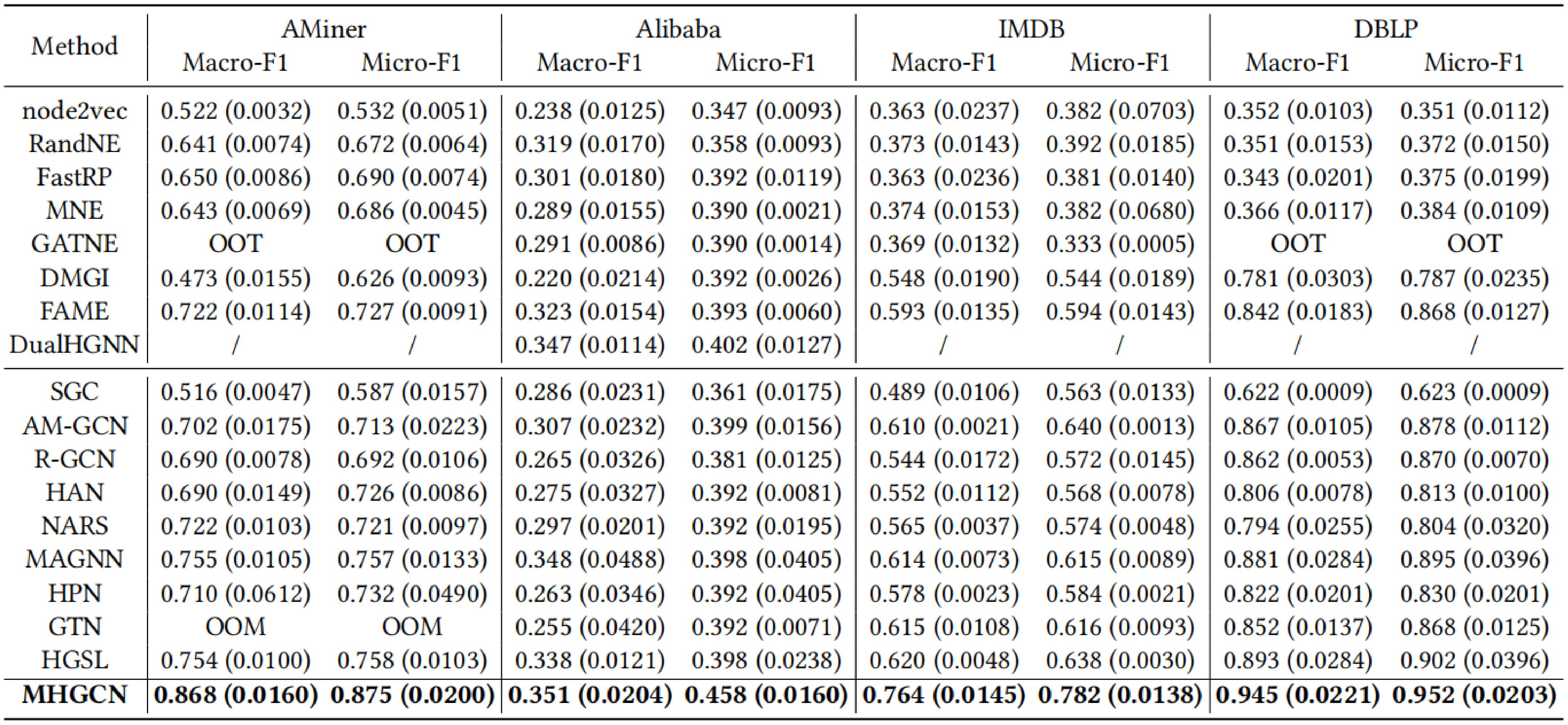


Method	AMiner			Alibaba			IMDB			Amazon			DBLP		
	R-AUC	PR-AUC	F1	R-AUC	PR-AUC	F1	R-AUC	PR-AUC	F1	R-AUC	PR-AUC	F1	R-AUC	PR-AUC	F1
node2vec	0.594	0.663	0.602	0.614	0.580	0.593	0.479	0.568	0.474	0.946	0.944	0.880	0.449	0.452	0.478
RandNE	0.607	0.630	0.608	0.877	0.888	0.826	0.901	0.933	0.839	0.950	0.941	0.903	0.492	0.491	0.493
FastRP	0.620	0.634	0.600	0.927	0.900	0.926	0.869	0.893	0.811	0.954	0.945	0.893	0.515	0.528	0.506
SGC	0.589	0.585	0.567	0.686	0.708	0.623	0.826	0.889	0.769	0.791	0.802	0.760	0.601	0.606	0.587
R-GCN	0.599	0.601	0.610	0.674	0.710	0.629	0.826	0.878	0.790	0.811	0.820	0.783	0.589	0.592	0.566
MAGNN	0.663	0.681	0.666	0.961	0.963	0.948	0.912	0.923	0.887	0.958	0.949	0.915	0.690	0.699	0.684
HPN	0.658	0.664	0.660	0.958	0.961	0.950	0.900	0.903	0.892	0.949	0.949	0.904	0.692	0.710	0.687
PMNE-n	0.651	0.669	0.677	0.966	0.973	0.891	0.674	0.683	0.646	0.956	0.945	0.893	0.672	0.679	0.663
PMNE-r	0.615	0.653	0.662	0.859	0.915	0.824	0.646	0.646	0.613	0.884	0.890	0.796	0.637	0.640	0.629
PMNE-c	0.613	0.635	0.657	0.597	0.591	0.664	0.651	0.634	0.630	0.934	0.934	0.868	0.622	0.625	0.609
MNE	0.660	0.672	0.681	0.944	0.946	0.901	0.688	0.701	0.681	0.941	0.943	0.912	0.657	0.660	0.635
GATNE	OOT	OOT	OOT	0.981	0.986	0.952	0.872	0.878	0.791	0.963	0.948	0.914	OOT	OOT	OOT
DMGI	OOM	OOM	OOM	0.857	0.781	0.784	0.926	0.935	0.873	0.905	0.878	0.847	0.610	0.615	0.601
FAME	0.687	0.747	0.726	0.993	0.996	0.979	0.944	0.959	0.897	0.959	0.950	0.900	0.642	0.650	0.633
DualHGNN	/	/	/	0.974	0.977	0.966	/	/	/	/	/	/	/	/	/
MHGCN	**0.711**	**0.753**	**0.730**	**0.997**	**0.997**	**0.992**	**0.967**	**0.966**	**0.959**	**0.972**	**0.974**	**0.961**	**0.718**	**0.722**	**0.703**

OOT: Out Of Time (36 hours). OOM: Out Of Memory; DMGI runs out of memory on the entire AMiner data. R-AUC: ROC-AUC.

Experiments

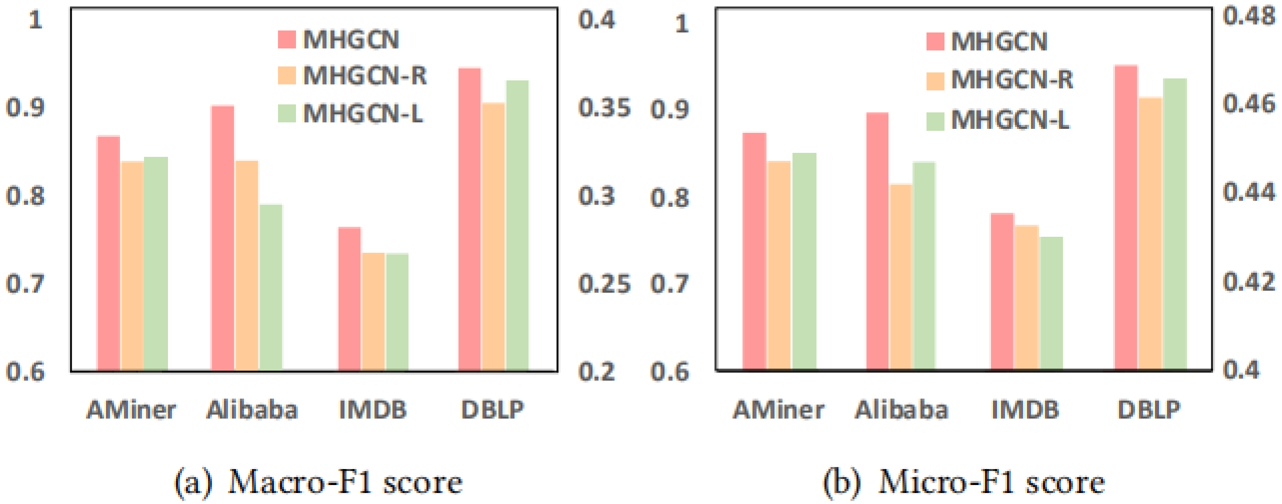
Table 3: Node classification performance comparison of different methods on four datasets



Method	AMiner		Alibaba		IMDB		DBLP	
	Macro-F1	Micro-F1	Macro-F1	Micro-F1	Macro-F1	Micro-F1	Macro-F1	Micro-F1
node2vec	0.522 (0.0032)	0.532 (0.0051)	0.238 (0.0125)	0.347 (0.0093)	0.363 (0.0237)	0.382 (0.0703)	0.352 (0.0103)	0.351 (0.0112)
RandNE	0.641 (0.0074)	0.672 (0.0064)	0.319 (0.0170)	0.358 (0.0093)	0.373 (0.0143)	0.392 (0.0185)	0.351 (0.0153)	0.372 (0.0150)
FastRP	0.650 (0.0086)	0.690 (0.0074)	0.301 (0.0180)	0.392 (0.0119)	0.363 (0.0236)	0.381 (0.0140)	0.343 (0.0201)	0.375 (0.0199)
MNE	0.643 (0.0069)	0.686 (0.0045)	0.289 (0.0155)	0.390 (0.0021)	0.374 (0.0153)	0.382 (0.0680)	0.366 (0.0117)	0.384 (0.0109)
GATNE	OOT	OOT	0.291 (0.0086)	0.390 (0.0014)	0.369 (0.0132)	0.333 (0.0005)	OOT	OOT
DMGI	0.473 (0.0155)	0.626 (0.0093)	0.220 (0.0214)	0.392 (0.0026)	0.548 (0.0190)	0.544 (0.0189)	0.781 (0.0303)	0.787 (0.0235)
FAME	0.722 (0.0114)	0.727 (0.0091)	0.323 (0.0154)	0.393 (0.0060)	0.593 (0.0135)	0.594 (0.0143)	0.842 (0.0183)	0.868 (0.0127)
DualHGNN	/	/	0.347 (0.0114)	0.402 (0.0127)	/	/	/	/
SGC	0.516 (0.0047)	0.587 (0.0157)	0.286 (0.0231)	0.361 (0.0175)	0.489 (0.0106)	0.563 (0.0133)	0.622 (0.0009)	0.623 (0.0009)
AM-GCN	0.702 (0.0175)	0.713 (0.0223)	0.307 (0.0232)	0.399 (0.0156)	0.610 (0.0021)	0.640 (0.0013)	0.867 (0.0105)	0.878 (0.0112)
R-GCN	0.690 (0.0078)	0.692 (0.0106)	0.265 (0.0326)	0.381 (0.0125)	0.544 (0.0172)	0.572 (0.0145)	0.862 (0.0053)	0.870 (0.0070)
HAN	0.690 (0.0149)	0.726 (0.0086)	0.275 (0.0327)	0.392 (0.0081)	0.552 (0.0112)	0.568 (0.0078)	0.806 (0.0078)	0.813 (0.0100)
NARS	0.722 (0.0103)	0.721 (0.0097)	0.297 (0.0201)	0.392 (0.0195)	0.565 (0.0037)	0.574 (0.0048)	0.794 (0.0255)	0.804 (0.0320)
MAGNN	0.755 (0.0105)	0.757 (0.0133)	0.348 (0.0488)	0.398 (0.0405)	0.614 (0.0073)	0.615 (0.0089)	0.881 (0.0284)	0.895 (0.0396)
HPN	0.710 (0.0612)	0.732 (0.0490)	0.263 (0.0346)	0.392 (0.0405)	0.578 (0.0023)	0.584 (0.0021)	0.822 (0.0201)	0.830 (0.0201)
GTN	OOM	OOM	0.255 (0.0420)	0.392 (0.0071)	0.615 (0.0108)	0.616 (0.0093)	0.852 (0.0137)	0.868 (0.0125)
HGSL	0.754 (0.0100)	0.758 (0.0103)	0.338 (0.0121)	0.398 (0.0238)	0.620 (0.0048)	0.638 (0.0030)	0.893 (0.0284)	0.902 (0.0396)
MHGCN	**0.868 (0.0160)**	**0.875 (0.0200)**	**0.351 (0.0204)**	**0.458 (0.0160)**	**0.764 (0.0145)**	**0.782 (0.0138)**	**0.945 (0.0221)**	**0.952 (0.0203)**

OOT: Out Of Time (36 hours), OOM: Out Of Memory. The standard deviations are reported in the parentheses.

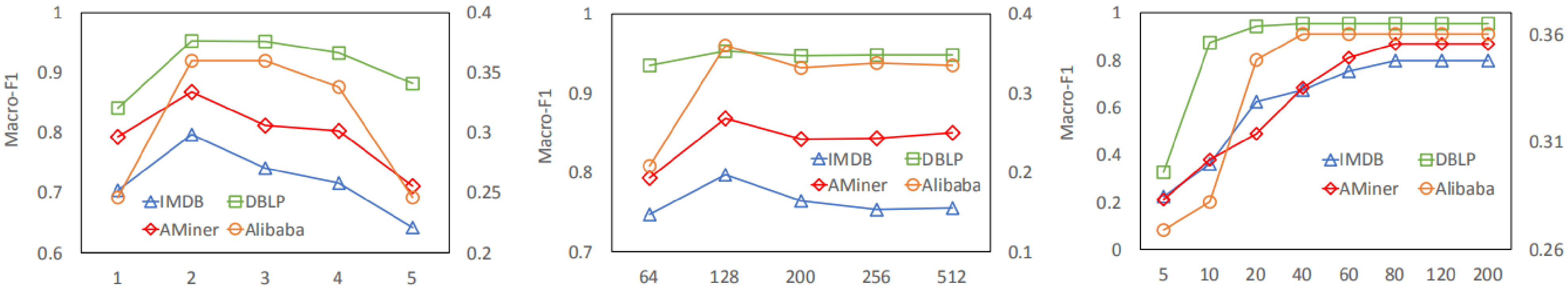
Experiments



(a) Macro-F1 score (b) Micro-F1 score

Figure 3: Experimental results of ablation study

Experiments



(a) Macro-F1 score *w.r.t.* #layers (b) Macro-F1 score *w.r.t.* dimension *d* (c) Macro-F1 score *w.r.t.* #rounds

Figure 4: Hyperparameter impact study of the proposed method *w.r.t.* #layers, dimension *d*, and #rounds.

Experiments

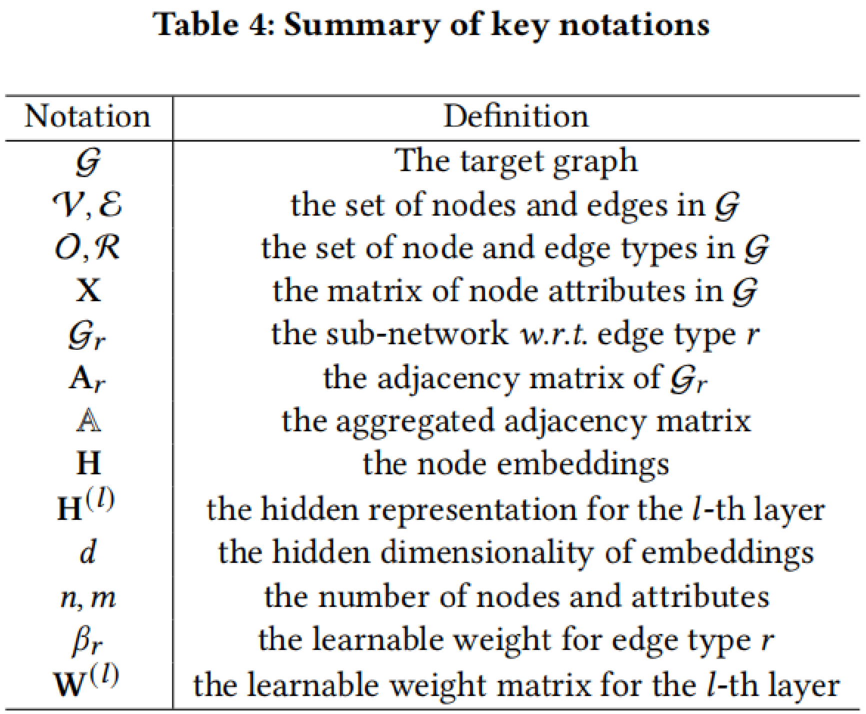
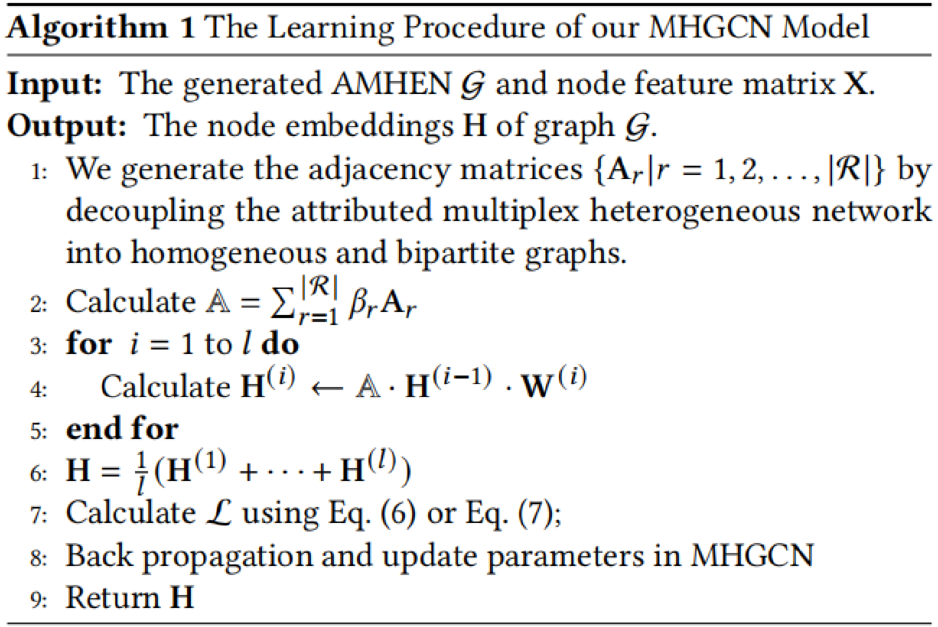


Table 4: Summary of key notations

Notation	Definition
$\mathcal{G}$	The target graph
$\mathcal{V}, \mathcal{E}$	the set of nodes and edges in $\mathcal{G}$
$\mathcal{O}, \mathcal{R}$	the set of node and edge types in $\mathcal{G}$
$\mathbf{X}$	the matrix of node attributes in $\mathcal{G}$
$\mathcal{G}_r$	the sub-network *w.r.t.* edge type r
$\mathbf{A}_r$	the adjacency matrix of $\mathcal{G}_r$
$\mathbb{A}$	the aggregated adjacency matrix
$\mathbf{H}$	the node embeddings
$\mathbf{H}^{(l)}$	the hidden representation for the l-th layer
d	the hidden dimensionality of embeddings
n, m	the number of nodes and attributes
β_r	the learnable weight for edge type r
$\mathbf{W}^{(l)}$	the learnable weight matrix for the l-th layer



Algorithm 1 The Learning Procedure of our MHGCN Model

```
Input: The generated AMHEN $\mathcal{G}$ and node feature matrix $\mathbf{X}$.
Output: The node embeddings $\mathbf{H}$ of graph $\mathcal{G}$.
We generate the adjacency matrices $\{\mathbf{A}_r | r = 1, 2, \ldots, |\mathcal{R}|\}$ by decoupling the attributed multiplex heterogeneous network into homogeneous and bipartite graphs.
Calculate $\mathbb{A} = \sum_{r=1}^{|\mathcal{R}|} \beta_r \mathbf{A}_r$
for $i = 1$ to $l$ do
    Calculate $\mathbf{H}^{(i)} \leftarrow \mathbb{A} \cdot \mathbf{H}^{(i-1)} \cdot \mathbf{W}^{(i)}$
end for
$\mathbf{H} = \frac{1}{l}(\mathbf{H}^{(1)} + \cdots + \mathbf{H}^{(l)})$
Calculate $\mathcal{L}$ using Eq. (6) or Eq. (7);
Back propagation and update parameters in MHGCN
Return $\mathbf{H}$
```

Experiments

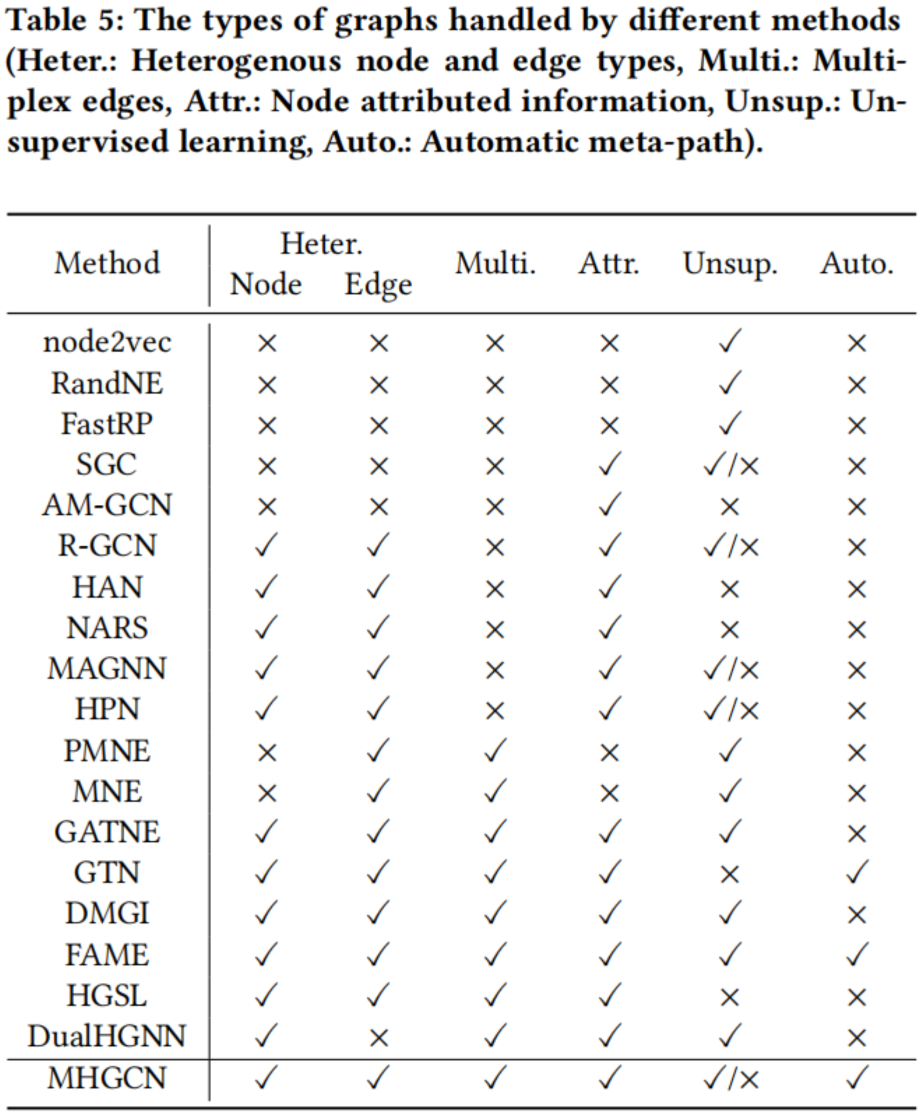


Table 5: The types of graphs handled by different methods (Heter.: Heterogenous node and edge types, Multi.: Multiplex edges, Attr.: Node attributed information, Unsup.: Unsupervised learning, Auto.: Automatic meta-path).

Method	Heter.		Multi.	Attr.	Unsup.	Auto.
	Node	Edge				
node2vec	×	×	×	×	✓	×
RandNE	×	×	×	×	✓	×
FastRP	×	×	×	×	✓	×
SGC	×	×	×	✓	✓/×	×
AM-GCN	×	×	×	✓	×	×
R-GCN	✓	✓	×	✓	✓/×	×
HAN	✓	✓	×	✓	×	×
NARS	✓	✓	×	✓	×	×
MAGNN	✓	✓	×	✓	✓/×	×
HPN	✓	✓	×	✓	✓/×	×
PMNE	×	✓	✓	×	✓	×
MNE	×	✓	✓	×	✓	×
GATNE	✓	✓	✓	✓	✓	×
GTN	✓	✓	✓	✓	×	✓
DMGI	✓	✓	✓	✓	✓	×
FAME	✓	✓	✓	✓	✓	✓
HGSL	✓	✓	✓	✓	×	×
DualHGNN	✓	×	✓	✓	✓	×
MHGCN	✓	✓	✓	✓	✓/×	✓

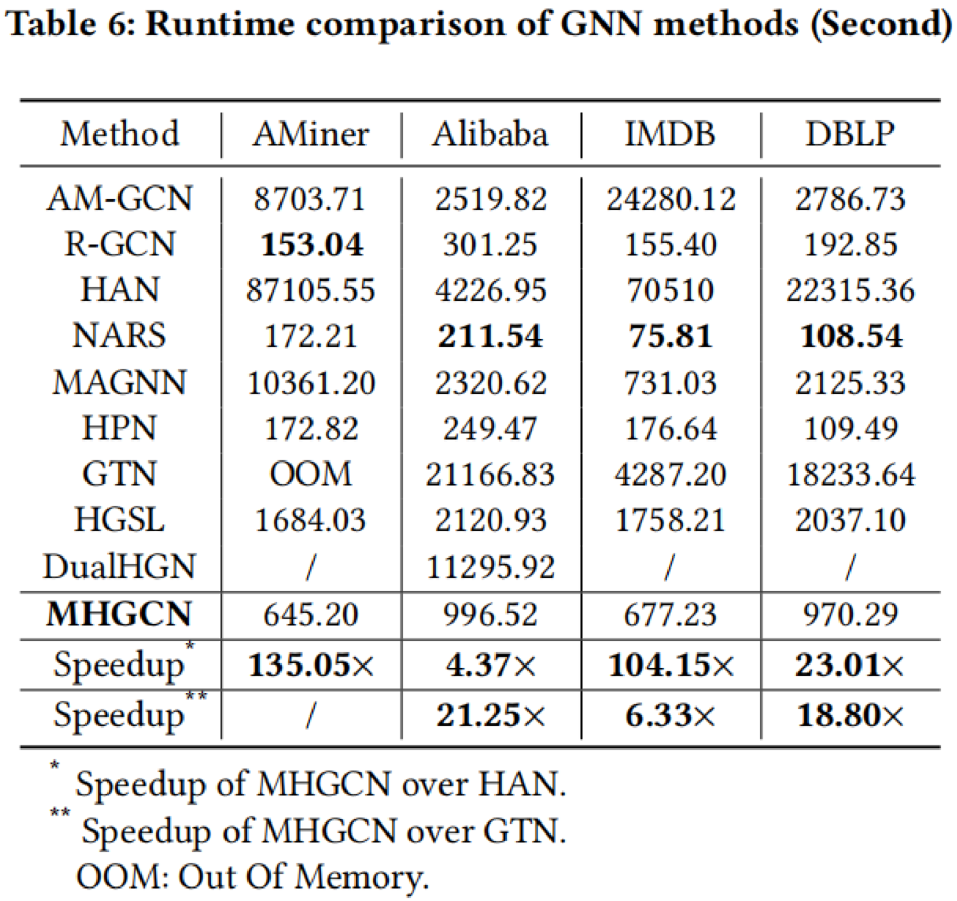


Table 6: Runtime comparison of GNN methods (Second)

Method	AMiner	Alibaba	IMDB	DBLP
AM-GCN	8703.71	2519.82	24280.12	2786.73
R-GCN	**153.04**	301.25	155.40	192.85
HAN	87105.55	4226.95	70510	22315.36
NARS	172.21	**211.54**	**75.81**	**108.54**
MAGNN	10361.20	2320.62	731.03	2125.33
HPN	172.82	249.47	176.64	109.49
GTN	OOM	21166.83	4287.20	18233.64
HGSL	1684.03	2120.93	1758.21	2037.10
DualHGN	/	11295.92	/	/
MHGCN	645.20	996.52	677.23	970.29
Speedup*	**135.05×**	**4.37×**	**104.15×**	**23.01×**
Speedup**	/	**21.25×**	**6.33×**	**18.80×**

* Speedup of MHGCN over HAN.
** Speedup of MHGCN over GTN.
OOM: Out Of Memory.

Thanks !