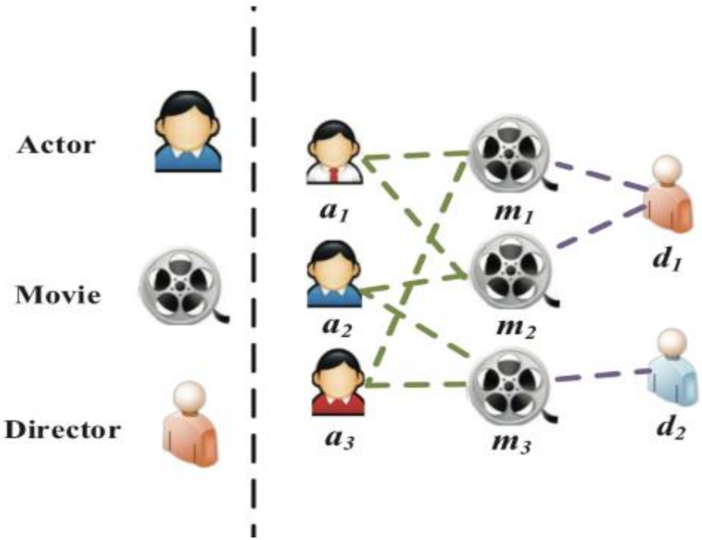
Heterogeneous Graph Neural Network: Models and Applications

叶泽凯

CONTENTS

CONTENTS

Introduction

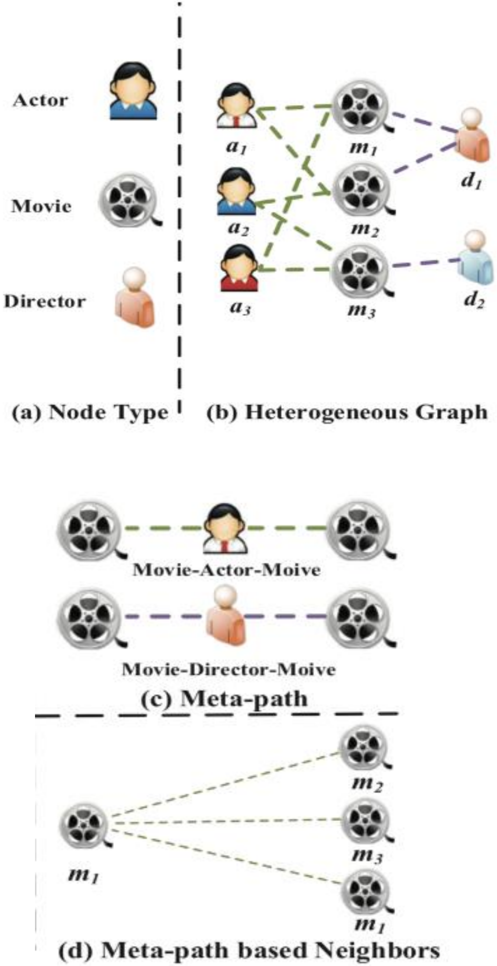


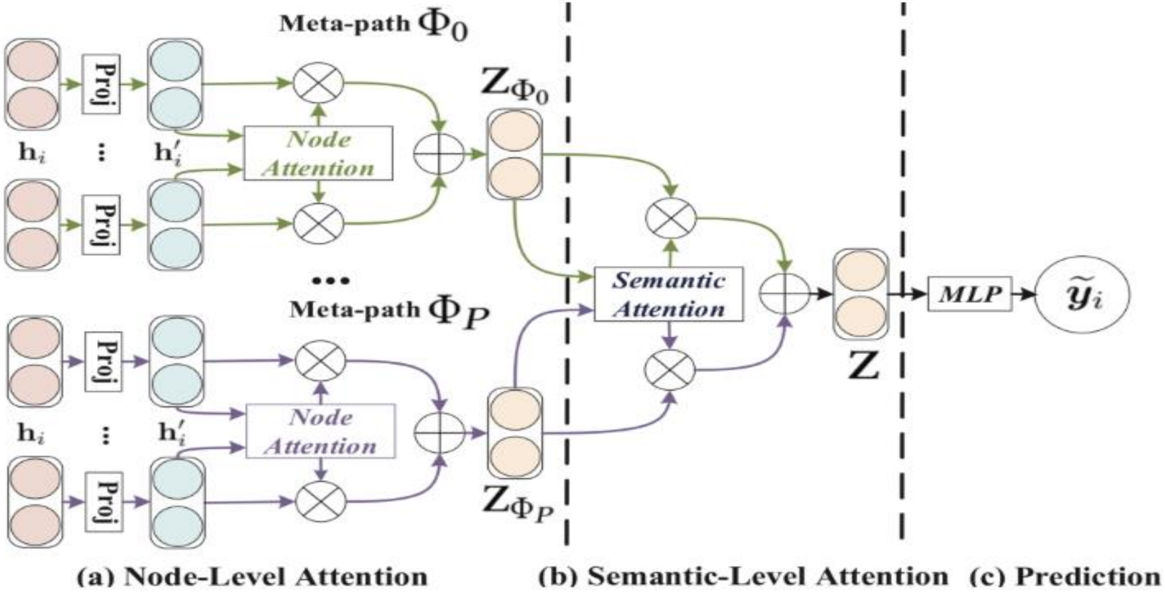
Existing graph neural networks focus on homogeneous graph:

- Cannot handle multiple types of nodes and edges.
- Cannot capture rich semantic information.

CONTENTS

- Introduction
- Models
- Applications
- Conclusion

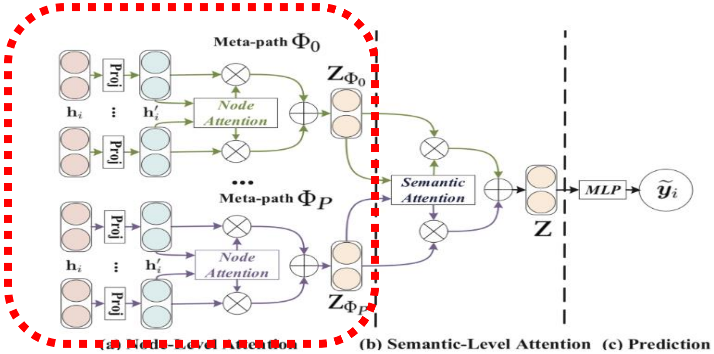




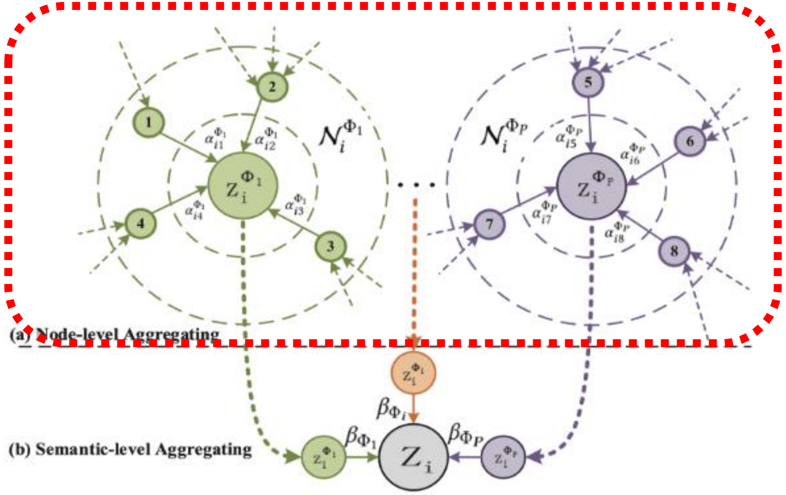
Graph Structure

Importance of Different Meta-paths

Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, Philip S. Yu
Heterogeneous Graph Attention Network
WWW,2019



Heterogeneous Graph Attention Network



type-specific transformation

$$\mathbf{h}'_i = \mathbf{M}_{\phi_i} \cdot \mathbf{h}_i,$$

Importance of Neighbors

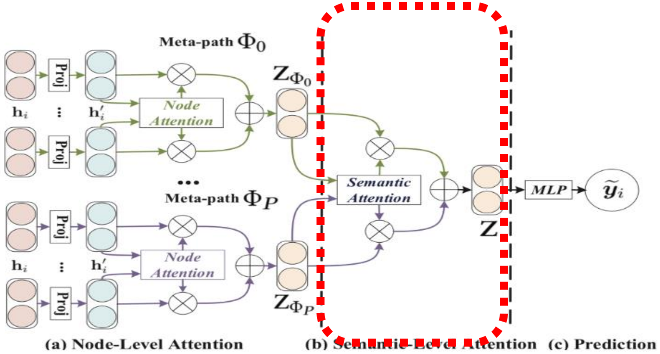
$$e_{ij}^{\Phi} = att_{node}(\mathbf{h}'_i, \mathbf{h}'_j; \Phi).$$

$$\alpha_{ij}^{\Phi} = softmax_j(e_{ij}^{\Phi}) = \frac{\exp\left(\sigma(\mathbf{a}_{\Phi}^{\mathrm{T}} \cdot [\mathbf{h}'_i \| \mathbf{h}'_j])\right)}{\sum_{k \in \mathcal{N}_i^{\Phi}} \exp\left(\sigma(\mathbf{a}_{\Phi}^{\mathrm{T}} \cdot [\mathbf{h}'_i \| \mathbf{h}'_k])\right)},$$

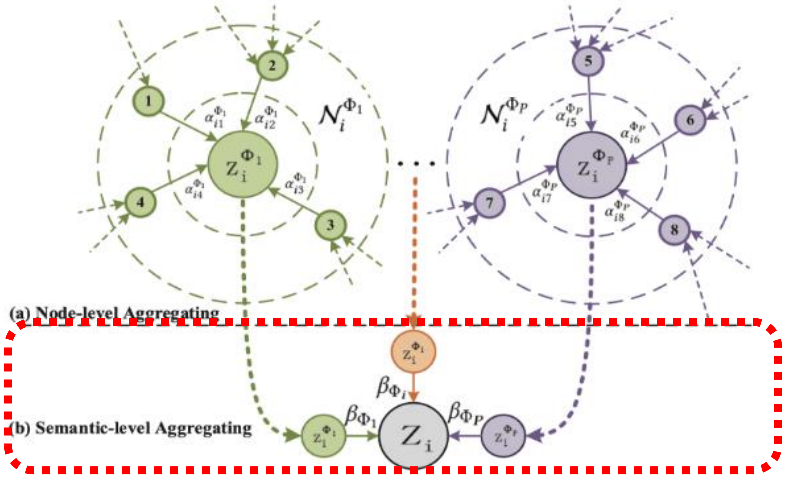
Node-Level Aggregating

$$\mathbf{z}_i^{\Phi} = \sigma\left(\sum_{j \in \mathcal{N}_i^{\Phi}} \alpha_{ij}^{\Phi} \cdot \mathbf{h}'_j\right).$$

Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, Philip S. Yu
Heterogeneous Graph Attention Network
WWW,2019



Heterogeneous Graph Attention Network



Semantic-Level Attention

$$(\beta_{\Phi_0}, \beta_{\Phi_1}, \ldots, \beta_{\Phi_P}) = att_{sem}(\mathbf{Z}_{\Phi_0}, \mathbf{Z}_{\Phi_1}, \ldots, \mathbf{Z}_{\Phi_P}).$$

Importance of Meta-path

$$w_{\Phi_i} = \frac{1}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} \mathbf{q}^{\mathrm{T}} \cdot \tanh(\mathbf{W} \cdot \mathbf{z}_i^{\Phi} + \mathbf{b}),$$

$$\beta_{\Phi_i} = \frac{\exp(w_{\Phi_i})}{\sum_{i=1}^{P} \exp(w_{\Phi_i})},$$

Semantic-Level Aggregating

$$\mathbf{Z} = \sum_{i=1}^{P} \beta_{\Phi_i} \cdot \mathbf{Z}_{\Phi_i}.$$

Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, Philip S. Yu
Heterogeneous Graph Attention Network
WWW,2019

Baselines

- Deepwalk
- Esim
- Metapath2vec
- GCN
- GAT
- HAN_{nd}
- HAN_{sem}
- HAN

Table 2: Statistics of the datasets.

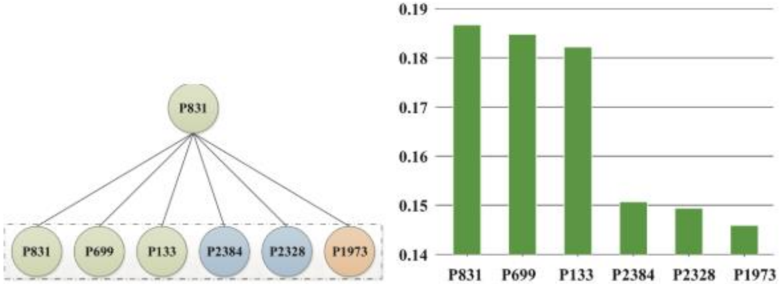
Dataset	Relations(A-B)	Number of A	Number of B	Number of A-B	Feature	Training	Validation	Test	Meta-paths
DBLP	Paper-Author	14328	4057	19645	334	800	400	2857	*APA*
	Paper-Conf	14328	20	14328					*APCPA*
	Paper-Term	14327	8789	88420					*APTPA*
IMDB	Movie-Actor	4780	5841	14340	1232	300	300	2687	*MAM*
	Movie-Director	4780	2269	4780					*MDM*
ACM	Paper-Author	3025	5835	9744	1830	600	300	2125	*PAP*
	Paper-Subject	3025	56	3025					*PSP*

Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, Philip S. Yu
Heterogeneous Graph Attention Network
WWW,2019

Table 3: Qantitative results (%) on the node classification task.

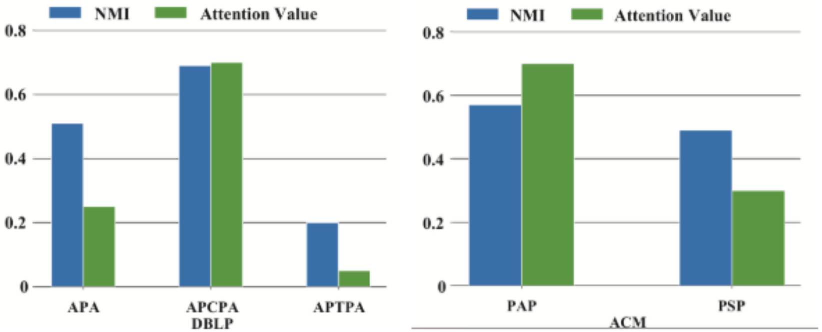
Datasets	Metrics	Training	DeepWalk	ESim	metapath2vec	HERec	GCN	GAT	HAN_{nd}	HAN_{sem}	HAN
ACM	Macro-F1	20%	77.25	77.32	65.09	66.17	86.81	86.23	88.15	89.04	**89.40**
		40%	80.47	80.12	69.93	70.89	87.68	87.04	88.41	89.41	**89.79**
		60%	82.55	82.44	71.47	72.38	88.10	87.56	87.91	**90.00**	89.51
		80%	84.17	83.00	73.81	73.92	88.29	87.33	88.48	90.17	**90.63**
	Micro-F1	20%	76.92	76.89	65.00	66.03	86.77	86.01	87.99	88.85	**89.22**
		40%	79.99	79.70	69.75	70.73	87.64	86.79	88.31	89.27	**89.64**
		60%	82.11	82.02	71.29	72.24	88.12	87.40	87.68	**89.85**	89.33
		80%	83.88	82.89	73.69	73.84	88.35	87.11	88.26	89.95	**90.54**
DBLP	Macro-F1	20%	77.43	91.64	90.16	91.68	90.79	90.97	91.17	92.03	**92.24**
		40%	81.02	92.04	90.82	92.16	91.48	91.20	91.46	92.08	**92.40**
		60%	83.67	92.44	91.32	92.80	91.89	90.80	91.78	92.38	**92.80**
		80%	84.81	92.53	91.89	92.34	92.38	91.73	91.80	92.53	**93.08**
	Micro-F1	20%	79.37	92.73	91.53	92.69	91.71	91.96	92.05	92.99	**93.11**
		40%	82.73	93.07	92.03	93.18	92.31	92.16	92.38	93.00	**93.30**
		60%	85.27	93.39	92.48	93.70	92.62	91.84	92.69	93.31	**93.70**
		80%	86.26	93.44	92.80	93.27	93.09	92.55	92.69	93.29	**93.99**
IMDB	Macro-F1	20%	40.72	32.10	41.16	41.65	45.73	49.44	49.78	**50.87**	50.00
		40%	45.19	31.94	44.22	43.86	48.01	50.64	52.11	50.85	**52.71**
		60%	48.13	31.68	45.11	46.27	49.15	51.90	51.73	52.09	**54.24**
		80%	50.35	32.06	45.15	47.64	51.81	52.99	52.66	51.60	**54.38**
	Micro-F1	20%	46.38	35.28	45.65	45.81	49.78	55.28	54.17	55.01	**55.73**
		40%	49.99	35.47	48.24	47.59	51.71	55.91	56.39	55.15	**57.97**
		60%	52.21	35.64	49.09	49.88	52.29	56.44	56.09	56.66	**58.32**
		80%	54.33	35.59	48.81	50.99	54.61	56.97	56.38	56.49	**58.51**

Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, Philip S. Yu
Heterogeneous Graph Attention Network
WWW,2019



(a) Meta-path based neighbors of P831 (b) Attention values of P831's neighbors

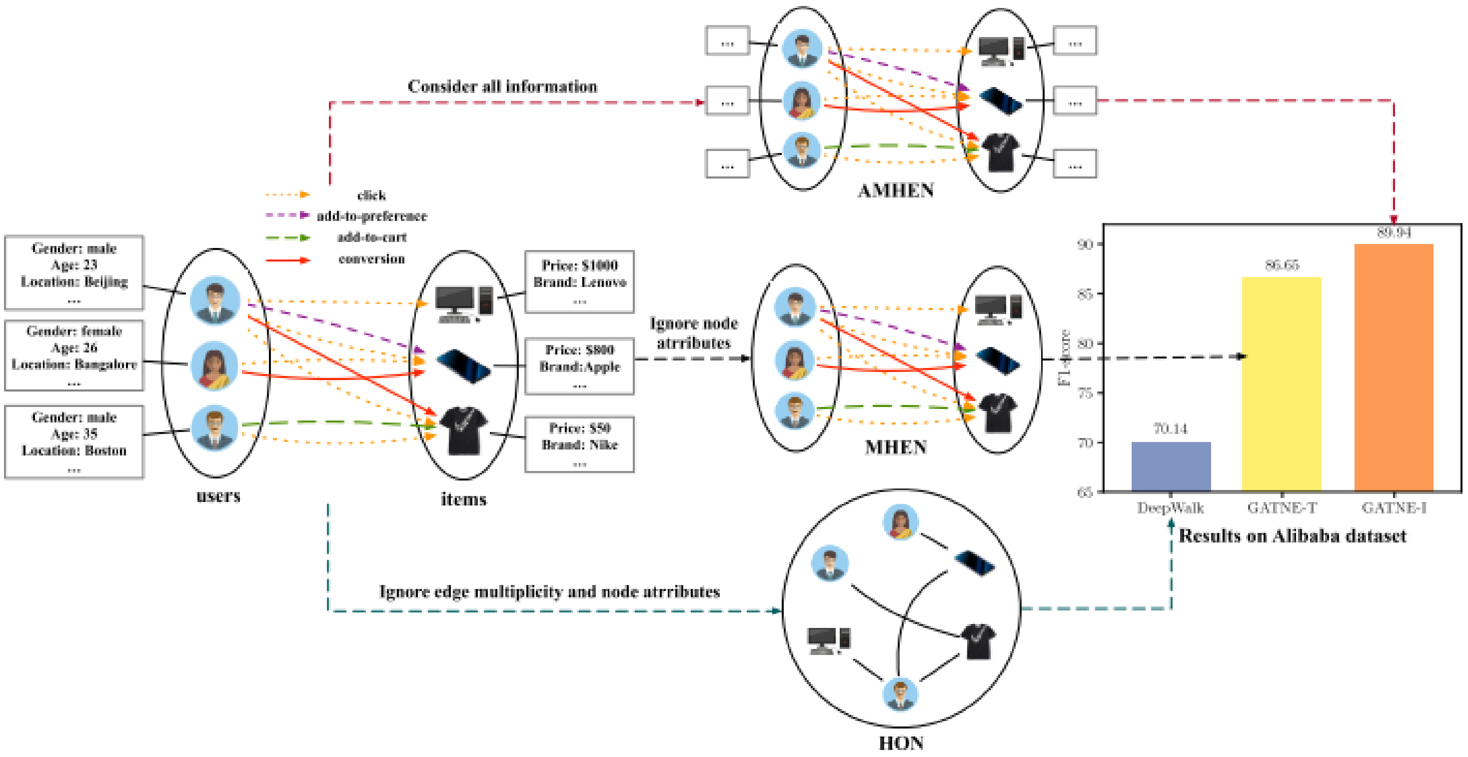
Figure 4: Meta-path based neighbors of node P831 and corresponding attention values (Different colors mean different classes, e.g., *green* means Data Mining, *blue* means Database, *orange* means Wireless Communication).



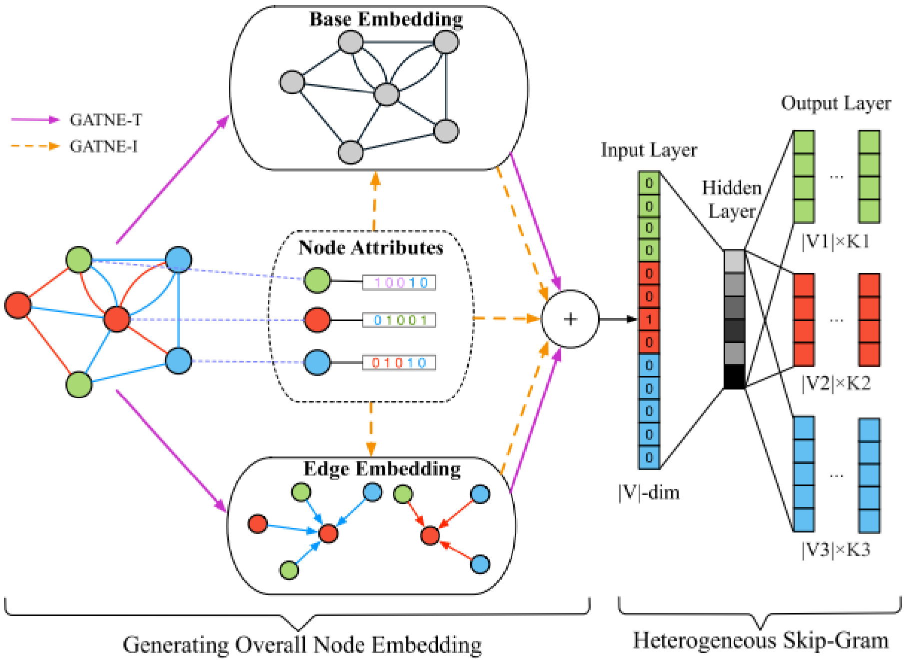
(a) NMI values on DBLP (b) NMI values on ACM

Figure 5: Performance of single meta-path and corresponding attention value.

Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, Philip S. Yu
Heterogeneous Graph Attention Network
WWW,2019



Yukuo Cen, Xu Zou, Jianwei Zhang, Hongxia Yang, Jingren Zhou, Jie Tang
Representation Learning for Attributed Multiplex Heterogeneous Network
KDD,2019



GATNE-T

$$\mathbf{u}_{i,r}^{(k)} = aggregator(\{\mathbf{u}_{j,r}^{(k-1)}, \forall v_j \in \mathcal{N}_{i,r}\}),$$

$$\mathbf{U}_i = (\mathbf{u}_{i,1}, \mathbf{u}_{i,2}, \ldots, \mathbf{u}_{i,m}).$$

$$\mathbf{a}_{i,r} = \mathrm{softmax}(\mathbf{w}_r^T \tanh(\mathbf{W}_r \mathbf{U}_i))^T,$$

$$\mathbf{v}_{i,r} = \mathbf{b}_i + \alpha_r \mathbf{M}_r^T \mathbf{U}_i \mathbf{a}_{i,r},$$

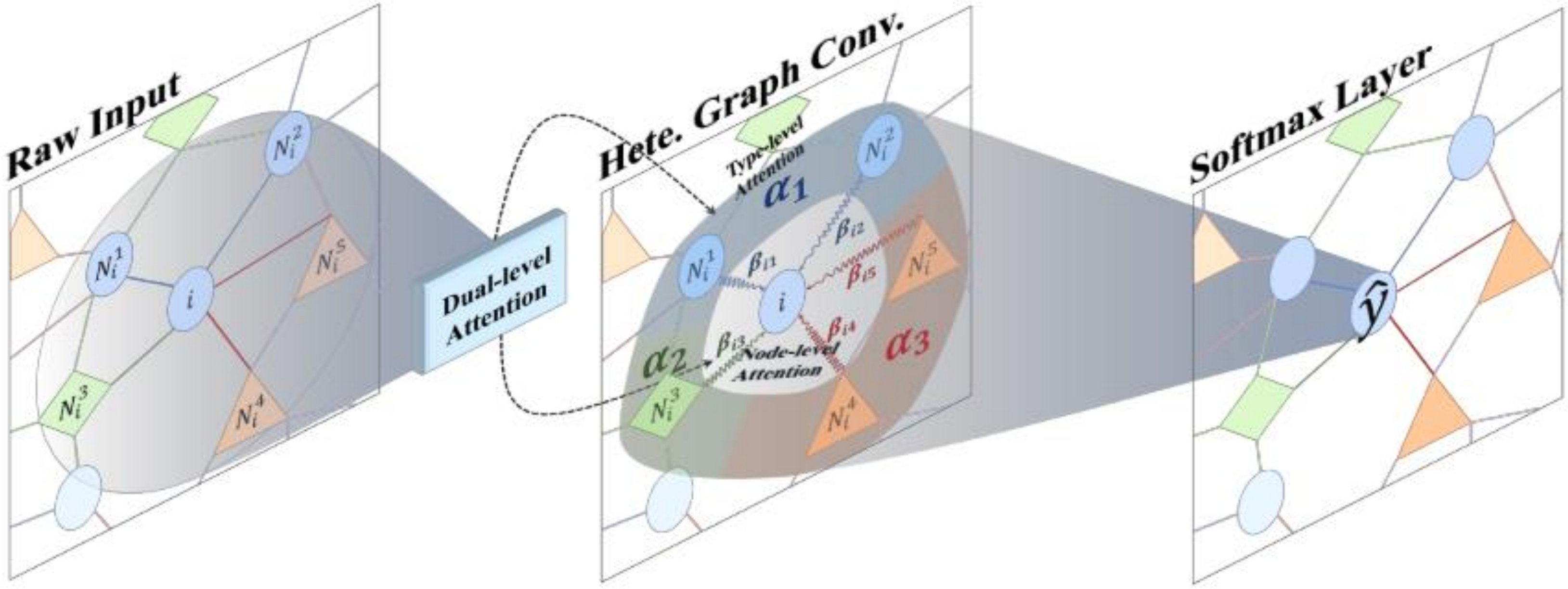
GATNE-I

$$\mathbf{v}_{i,r} = \mathbf{h}_z(\mathbf{x}_i) + \alpha_r \mathbf{M}_r^T \mathbf{U}_i \mathbf{a}_{i,r} + \beta_r \mathbf{D}_z^T \mathbf{x}_i,$$

Yukuo Cen, Xu Zou, Jianwei Zhang, Hongxia Yang, Jingren Zhou, Jie Tang
Representation Learning for Attributed Multiplex Heterogeneous Network
KDD,2019

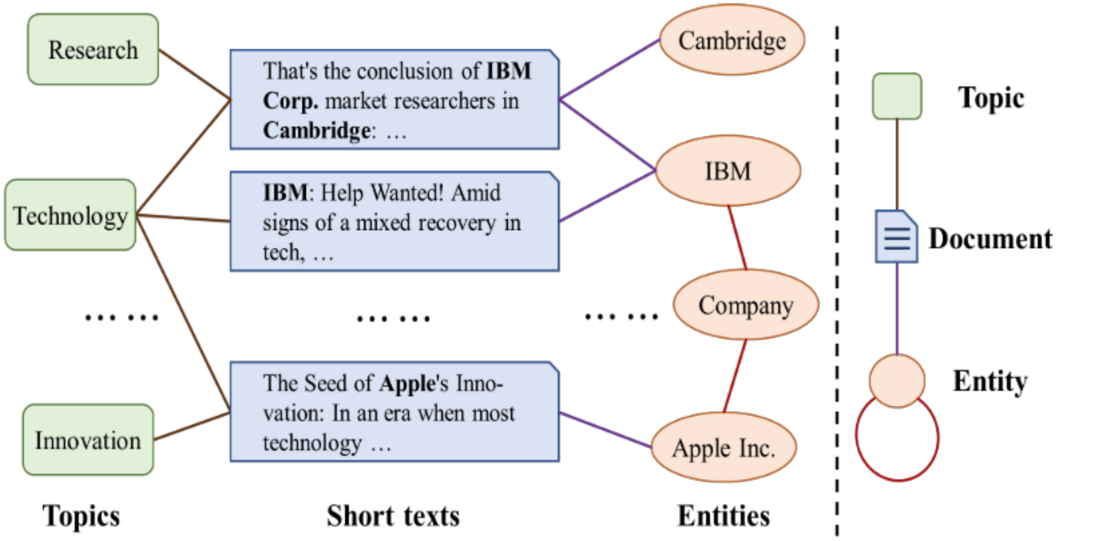
CONTENTS

- Introduction
- Models
- Applications
- Conclusion



Heterogeneous Graph **Dual-level Attention** **Semi-supervised Training**

Linmei Hu, Tianchi Yang, Chuan Shi, Houye Ji, Xiaoli Li
Heterogeneous Graph Attention Networks for Semi-supervised ShortText Classification.
EMNLP/IJCNLP (1),2019



(a) HIN for short texts (b) HIN schema

- **Nodes**

  Text: Document in the corpus.

  Topic: Mined by LDA.

  Entity: Recognized by Tagme.

- **Edges**

  Text-topic: The text is assigned to the Top P topics.

  Text-Entity: The text contains the entity.

  Entity-Entity: The similarity score between two entities

Linmei Hu, Tianchi Yang, Chuan Shi, Houye Ji, Xiaoli Li
Heterogeneous Graph Attention Networks for Semi-supervised ShortText Classification.
EMNLP/IJCNLP (1),2019

Type-level Attention

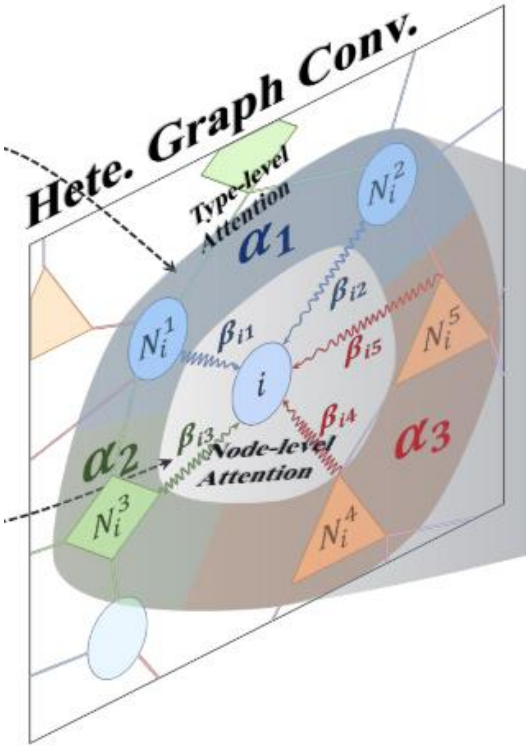
$$a_{\tau} = \sigma(\mu_{\tau}^{T} \cdot [h_v || h_{\tau}]), \qquad \alpha_{\tau} = \frac{\exp(a_{\tau})}{\sum_{\tau' \in \mathcal{T}} \exp(a_{\tau'})}.$$

Node-level Attention

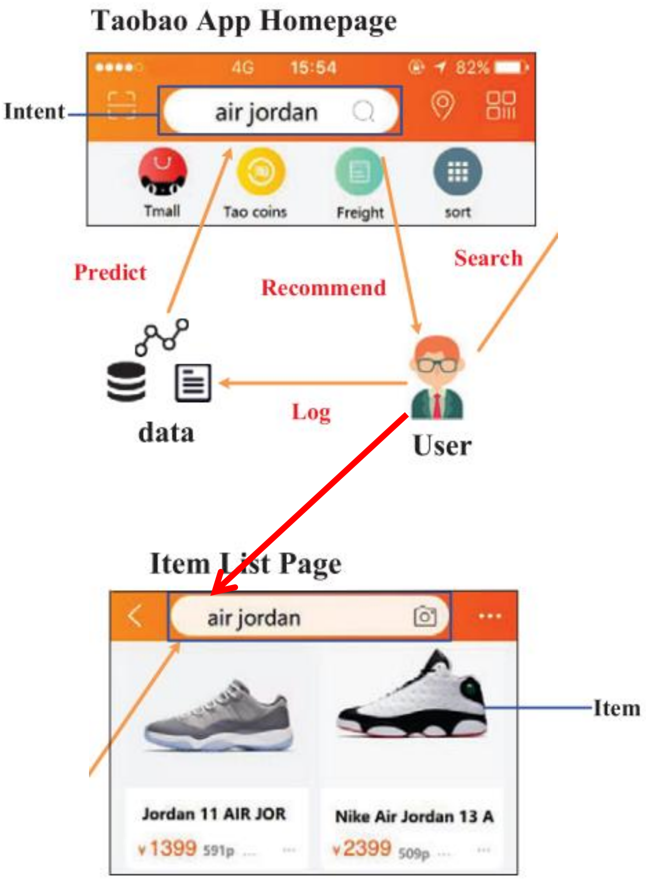
$$b_{vv'} = \sigma(\nu^{T} \cdot \alpha_{\tau'}[h_v || h_{v'}]), \quad \beta_{vv'} = \frac{\exp(b_{vv'})}{\sum_{i \in \mathcal{N}_v} \exp(b_{vi})}.$$

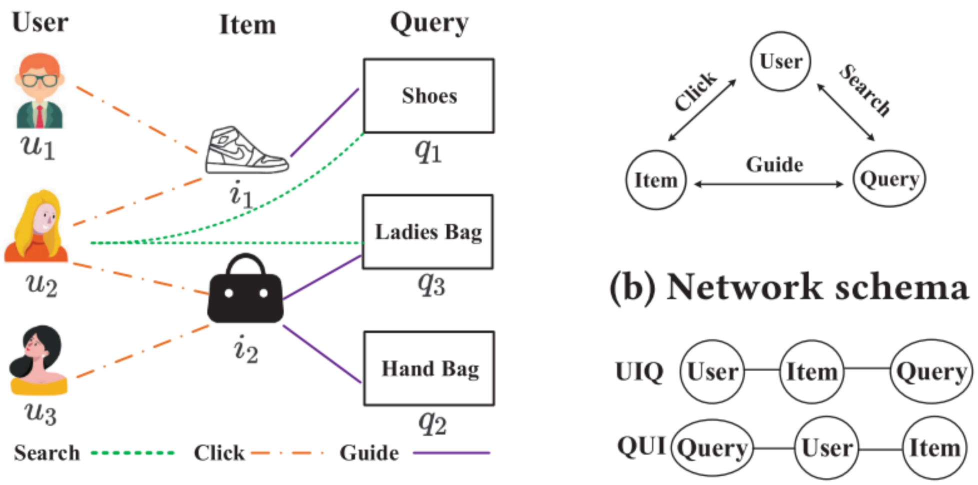
Dual-level Attention Based Hete. Graph Conv.

$$H^{(l+1)} = \sigma(\sum_{\tau \in \mathcal{T}} \mathcal{B}_{\tau} \cdot H_{\tau}^{(l)} \cdot W_{\tau}^{(l)}).$$



Linmei Hu, Tianchi Yang, Chuan Shi, Houye Ji, Xiaoli Li
Heterogeneous Graph Attention Networks for Semi-supervised ShortText Classification.
EMNLP/IJCNLP (1),2019





(a) Toy example

(b) Network schema

(c) Metapaths

Shaohua Fan, Junxiong Zhu, Xiaotian Han, Chuan Shi, Linmei Hu, Biyu Ma, Yongliang Li
Metapath-guided Heterogeneous Graph Neural Network for IntentRecommendation
KDD,2019

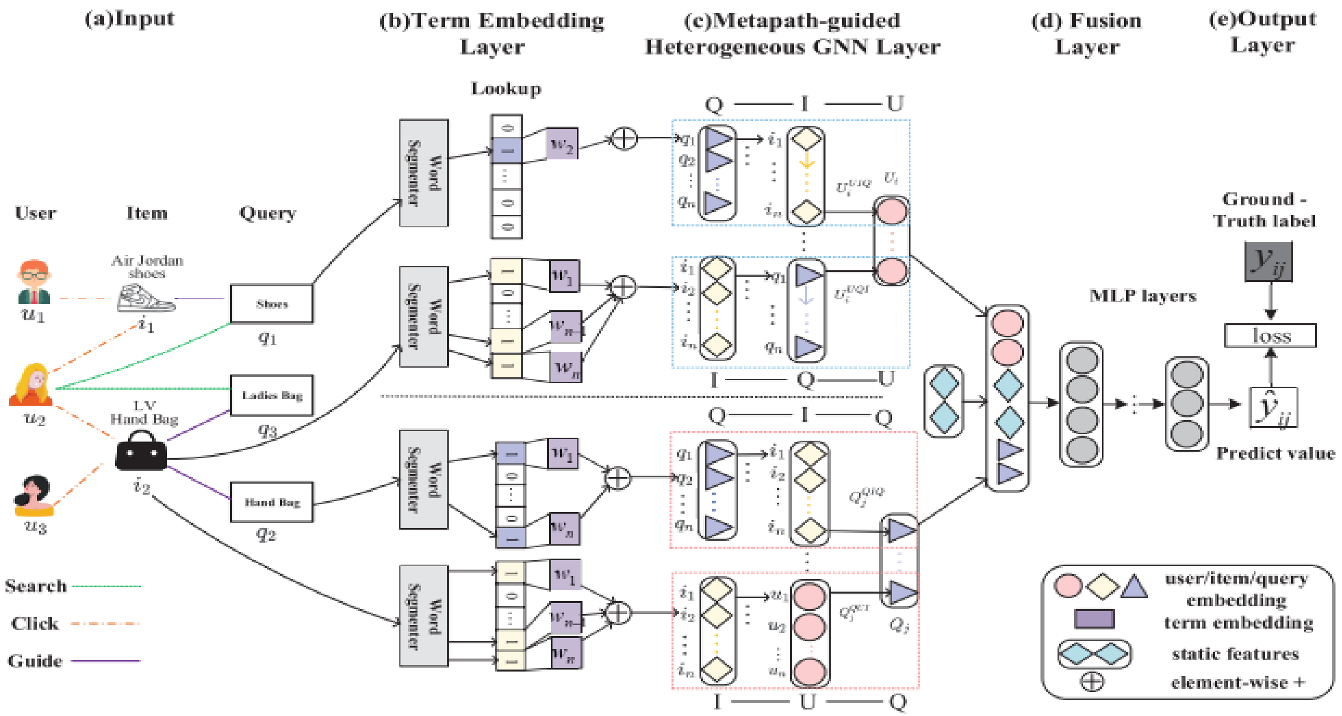
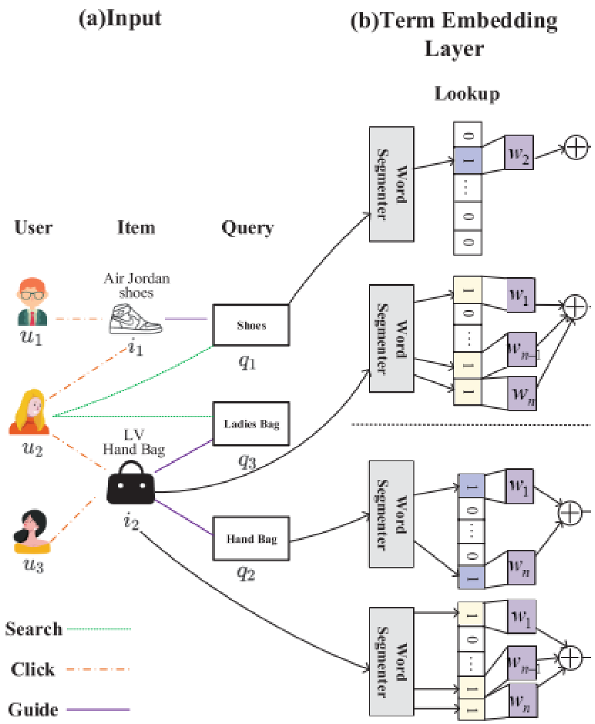


Figure 3: The framework of MEIRec.

Handle large and Dynamic data **Capture rich semantics** **Task-specific loss**

Shaohua Fan, Junxiong Zhu, Xiaotian Han, Chuan Shi, Linmei Hu, Biyu Ma, Yongliang Li
Metapath-guided Heterogeneous Graph Neural Network for IntentRecommendation
KDD,2019



Parameter share

Queries and items are constituted by the same term embedding

$$
\begin{array}{rccccc}
 & \{w_1, & w_2, & \cdots, & w_{n-1}, & w_n\} \\
q_2 = & (1, & 0, & \cdots, & 0, & 1) \\
i_2 = & (1, & 0, & \cdots, & 1, & 1).
\end{array}
$$

Reduce parameter space complexity

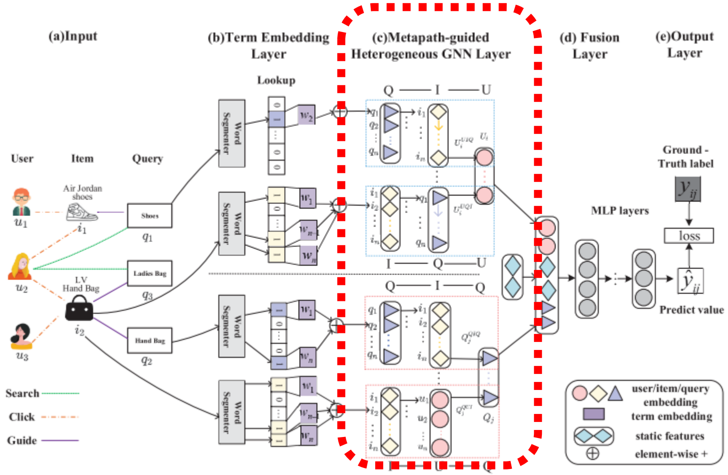
Traditional latent factor model:

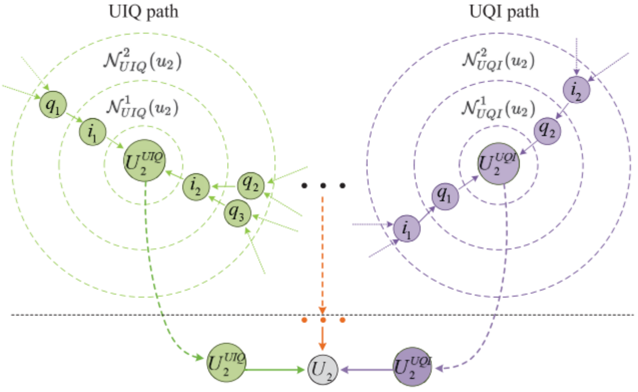
MEIRec:

New objects

The embedding of new objects (item, query) can be computed by trained term embedding in the testing phase

Shaohua Fan, Junxiong Zhu, Xiaotian Han, Chuan Shi, Linmei Hu, Biyu Ma, Yongliang Li
Metapath-guided Heterogeneous Graph Neural Network for IntentRecommendation
KDD,2019





Initial embedding

$$E_{q_2} = g(e_{w_1}, e_{w_n}), E_{i_2} = g(e_{w_1}, e_{w_{n-1}}, e_{w_n}),$$

Neighbor aggregation

$$I_j^{\mathrm{UIQ}} = g(E_{q_1}, E_{q_2}, \cdots),$$

$$U_i^{\mathrm{UIQ}} = g(I_1^{\mathrm{UIQ}}, I_2^{\mathrm{UIQ}}, \cdots),$$

Meta-path Aggregating

$$U_i = g(U_i^{\rho_1}, U_i^{\rho_2}, \cdots, U_i^{\rho_k}),$$

Shaohua Fan, Junxiong Zhu, Xiaotian Han, Chuan Shi, Linmei Hu, Biyu Ma, Yongliang Li
Metapath-guided Heterogeneous Graph Neural Network for IntentRecommendation
KDD,2019

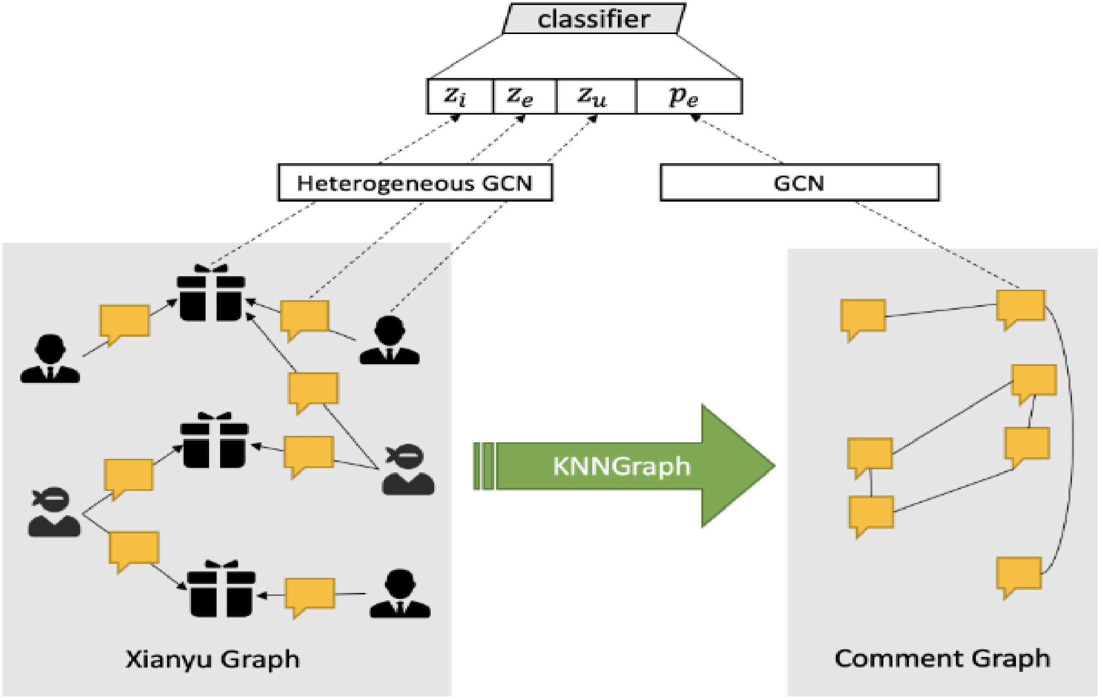
Spammers normally take the following two adversarial tricks to circumvent the anti-spam system:

- **Camouflage:** Using different expressions with similar meaning.
- **Deforming the comments:** Spammers replace some key-words in the comments with rarely used Chinese characters or typos deliberately.

Ao Li, Zhou Qin, Runshi Liu, Yiqun Yang, Dong Li
Spam Review Detection with Graph Convolutional Networks
CIKM,2019

Heterogeneous graph

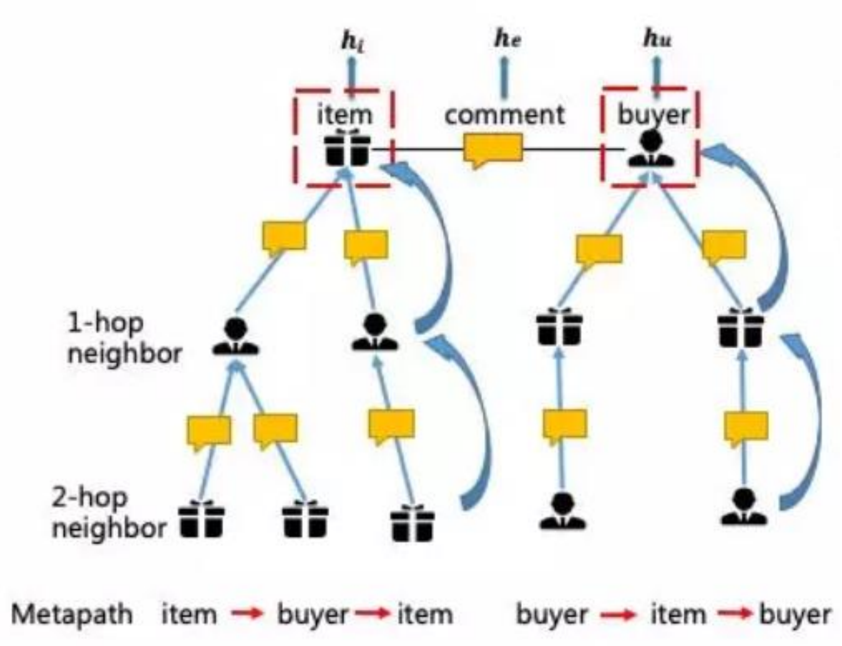
- Item-comment-user
- Metapaths
- 2-hop neighborhood aggregation



Homogeneous graph

- comment-comment
- Made by KNNGraph
- Smoothing process

Ao Li, Zhou Qin, Runshi Liu, Yiqun Yang, Dong Li
Spam Review Detection with Graph Convolutional Networks
CIKM,2019



$$h_e^l = \sigma\left(W_E^l \cdot AGG_E^l(h_e^{l-1}, h_{U(e)}^{l-1}, h_{I(e)}^{l-1})\right)$$

$$AGG_E^l\left(h_e^{l-1}, h_{U(e)}^{l-1}, h_{I(e)}^{l-1}\right) = concat\left(h_e^{l-1}, h_{U(e)}^{l-1}, h_{I(e)}^{l-1}\right)$$

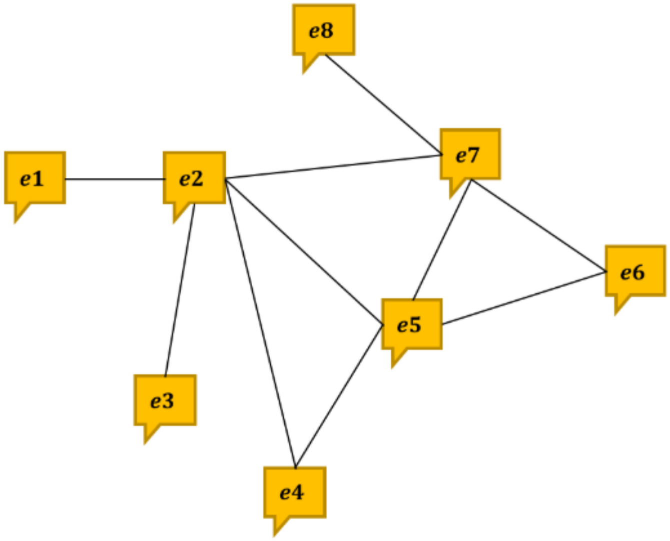
$$h_{N(u)}^l = \sigma\left(W_U^l \cdot AGG_U^l(\mathcal{H}_{IE}^{l-1})\right)$$

$$h_{N(i)}^l = \sigma\left(W_I^l \cdot AGG_I^l(\mathcal{H}_{UE}^{l-1})\right)$$

$$h_u^l = concat\left(V_U^l \cdot h_u^{l-1}, h_{N(u)}^l\right)$$

$$h_i^l = concat\left(V_I^l \cdot h_i^{l-1}, h_{N(i)}^l\right)$$

Ao Li, Zhou Qin, Runshi Liu, Yiqun Yang, Dong Li
Spam Review Detection with Graph Convolutional Networks
CIKM,2019



the Comment Graph is constructed as follows:

• Remove all the duplicated comments.
• Generate comments embeddings by the method described in.
• Obtain the similar comment pairs by employing the approx- imate KNN Graph algorithm.
• Remove comment pairs posted by same user or posted un- der same item, since the local context has been taken into consideration on Xianyu Graph.

Ao Li, Zhou Qin, Runshi Liu, Yiqun Yang, Dong Li
Spam Review Detection with Graph Convolutional Networks
CIKM,2019

CONTENTS

Thanks