

# Graph Neural Network for Recommendations

Wenqi Fan

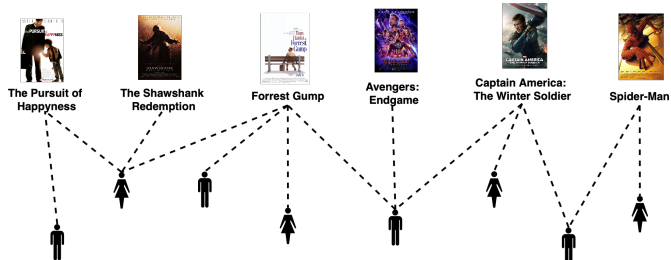
The Hong Kong Polytechnic University

<https://wenqifan03.github.io>, [wenqifan@polyu.edu.hk](mailto:wenqifan@polyu.edu.hk)

Tutorial website: <https://deeprs-tutorial.github.io>



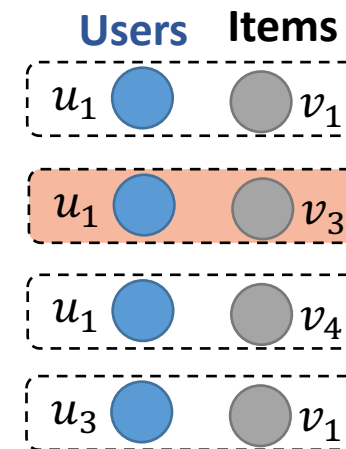
# A General Paradigm



users

|   | items |   |   |  |
|---|-------|---|---|--|
| 1 | 0     | 1 | 1 |  |
| 0 | 1     | 0 | 0 |  |
| 1 | 1     | 0 | 0 |  |
| 1 | 0     | 0 | 1 |  |

0/1 Interaction matrix

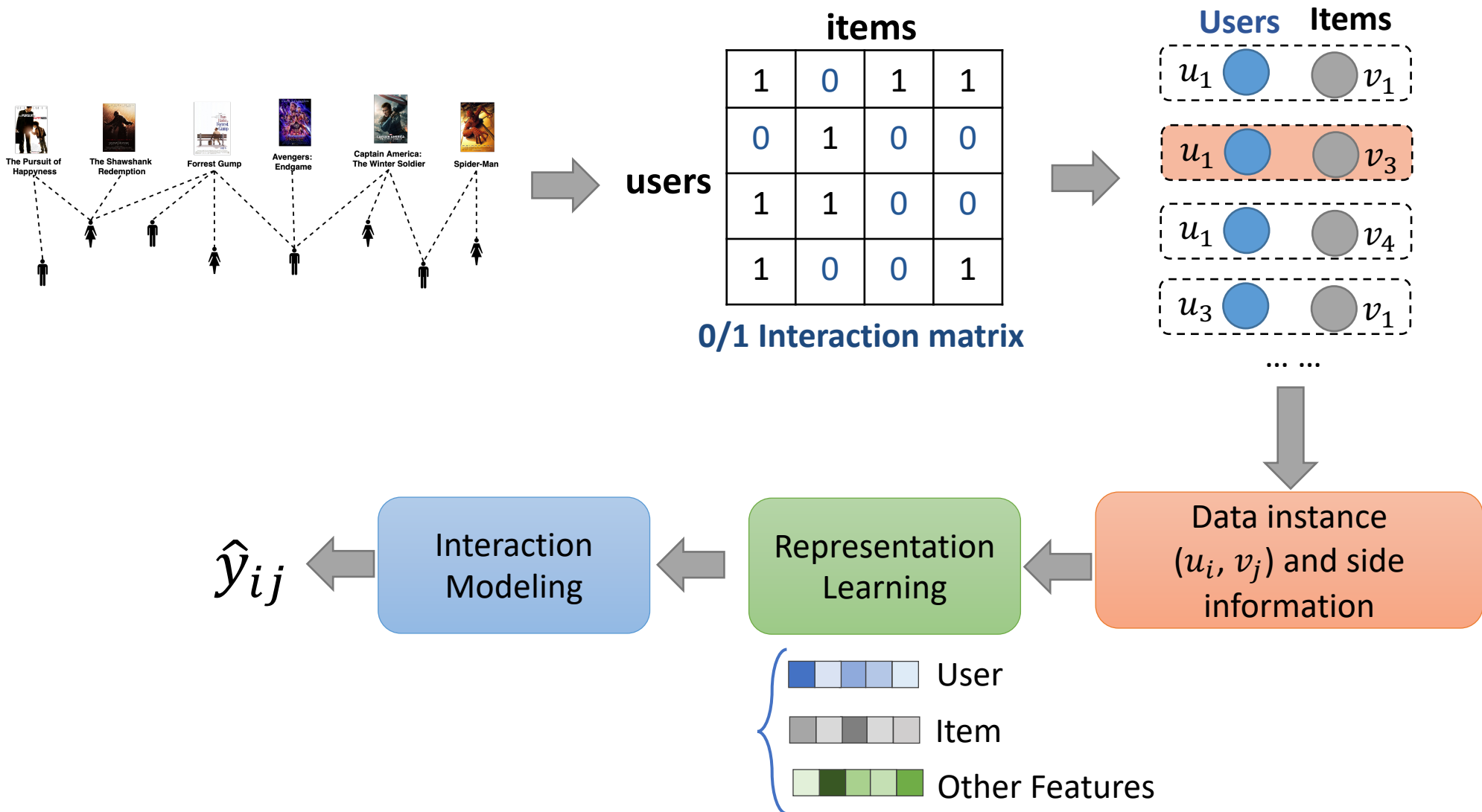


... ..

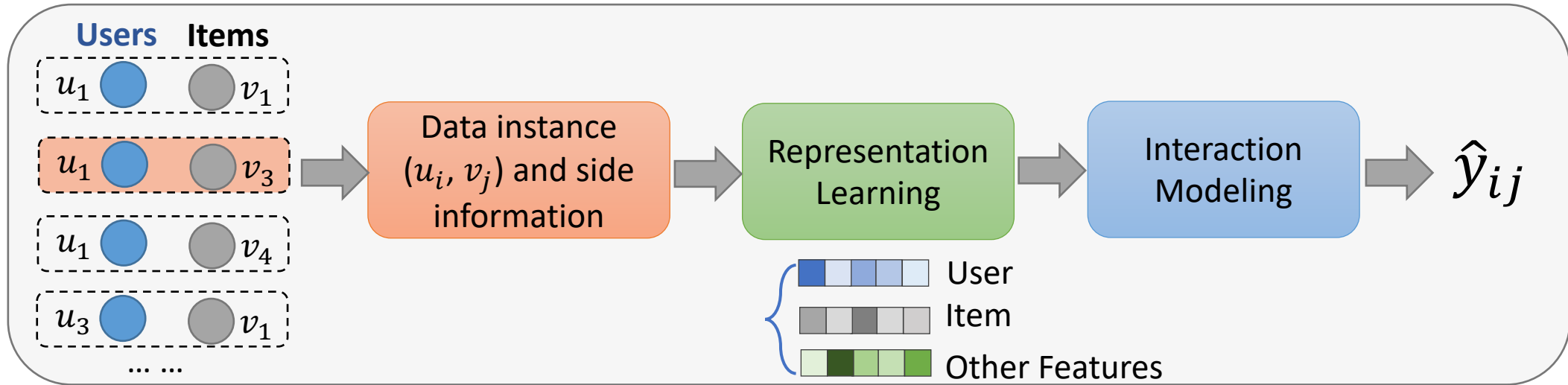


Data instance  
 $(u_i, v_j)$  and side  
 information

# A General Paradigm



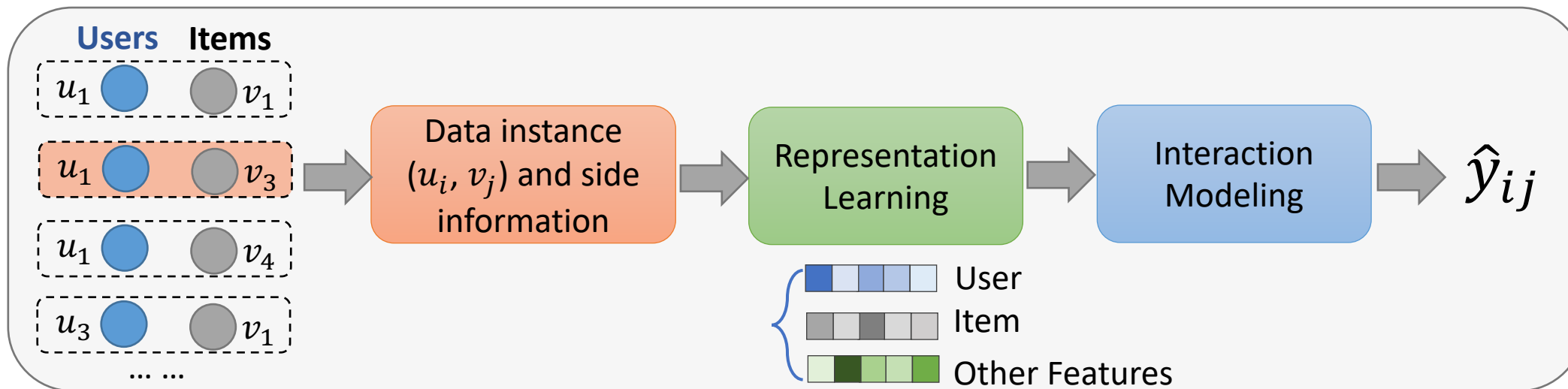
# A General Paradigm



## Information Isolated Island Issue

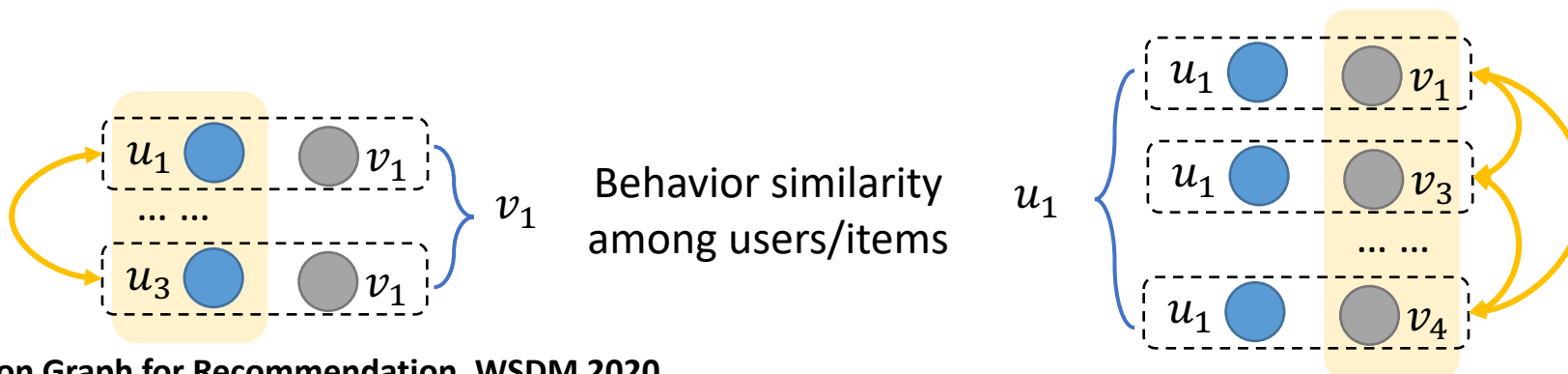
ignore implicit/explicit relationships among instances (**High-order Connectivity**)

# A General Paradigm



## Information Isolated Island Issue

ignore implicit/explicit relationships among instances (**High-order Connectivity**)



# Data as Graphs



**Most of the data in RS has essentially a graph structure**

- E-commerce, Content Sharing, Social Networking ...

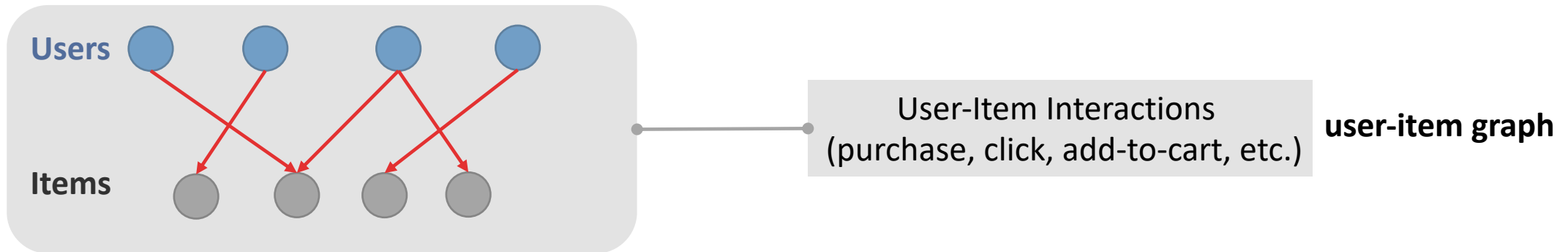
**The world is more closely connected than you might think!**

# Data as Graphs

Most of the data in RS has essentially a graph structure

- E-commerce, Content Sharing, Social Networking ...

**The world is more closely connected than you might think!**

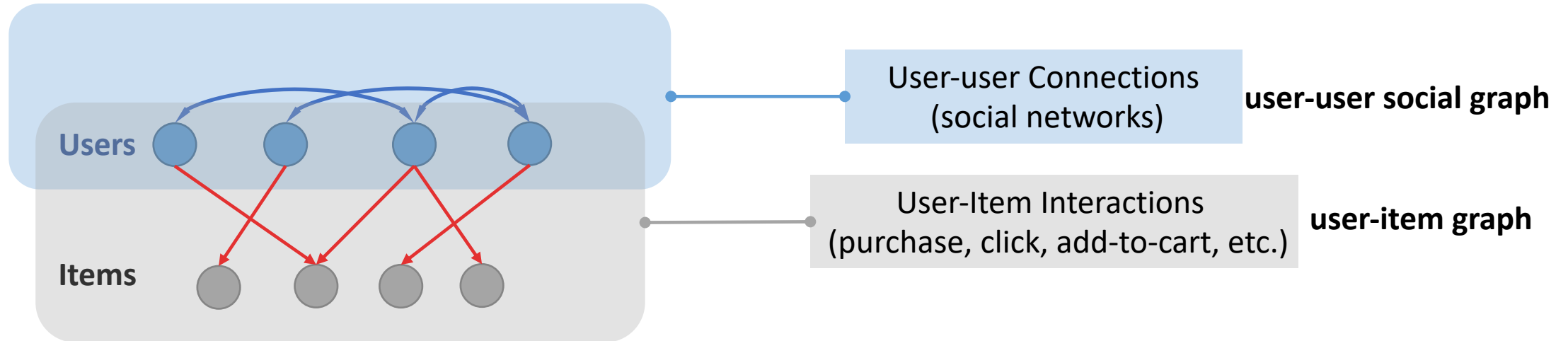


# Data as Graphs

Most of the data in RS has essentially a graph structure

- E-commerce, Content Sharing, Social Networking ...

**The world is more closely connected than you might think!**



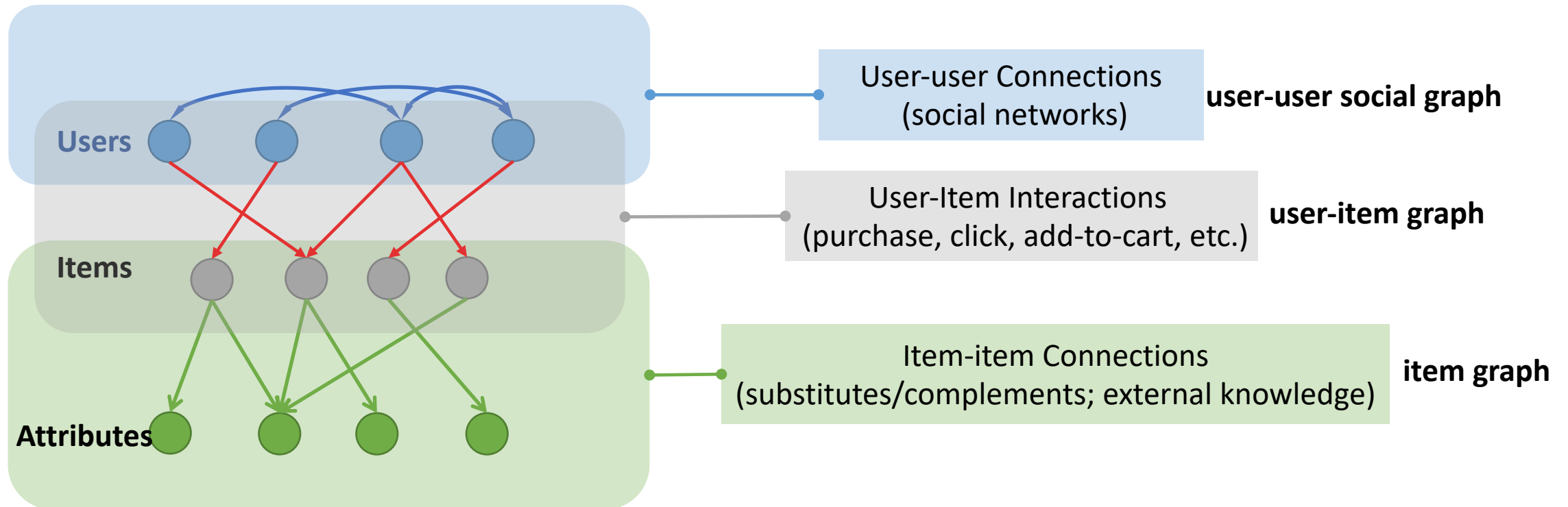


# Data as Graphs

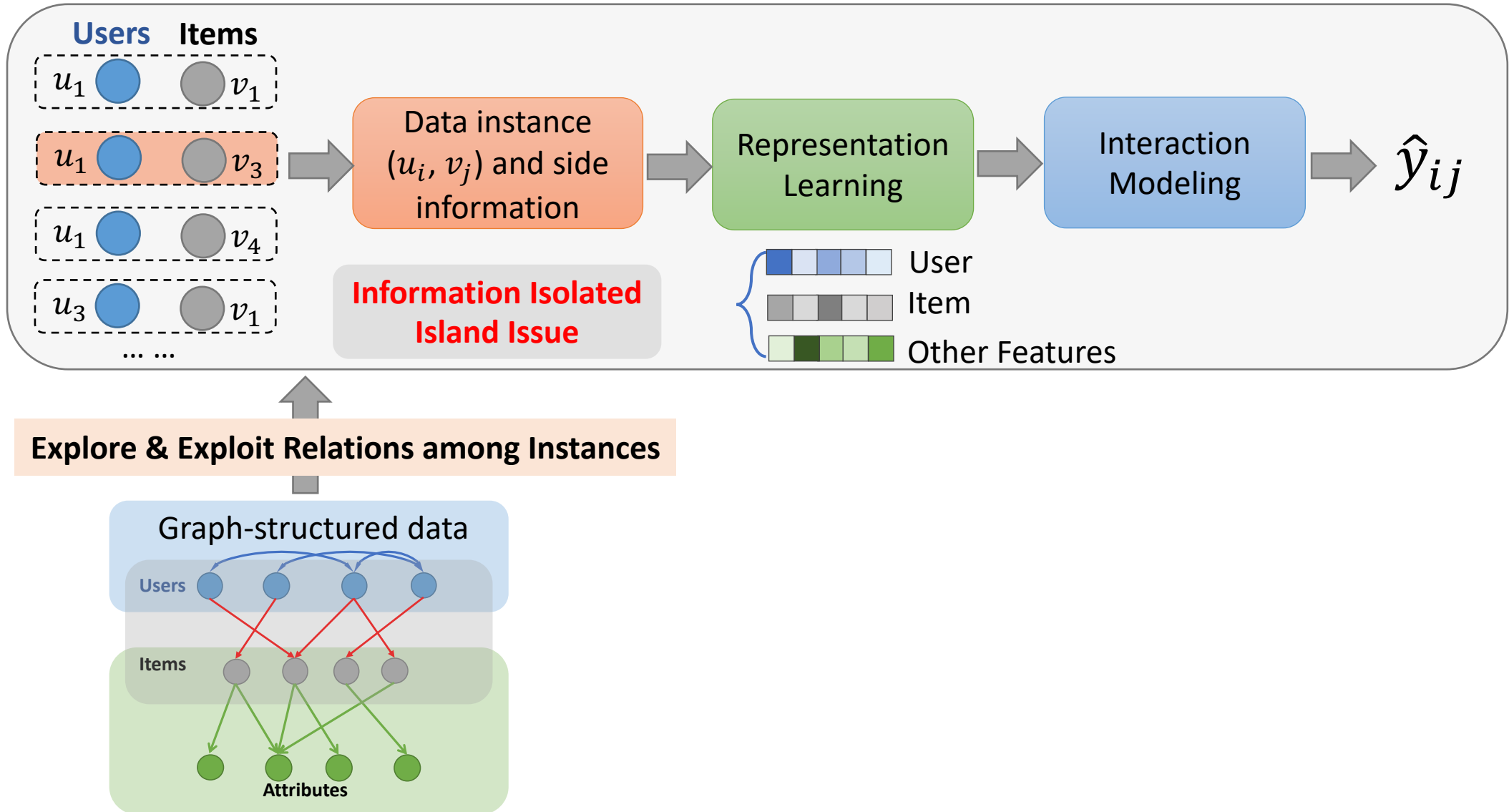
Most of the data in RS has essentially a graph structure

- E-commerce, Content Sharing, Social Networking ...

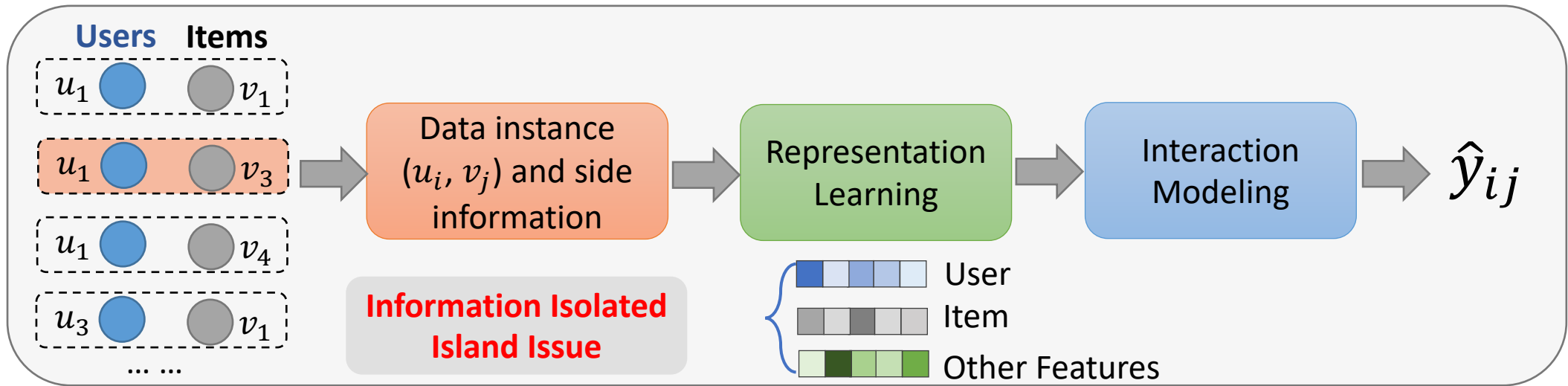
**The world is more closely connected than you might think!**



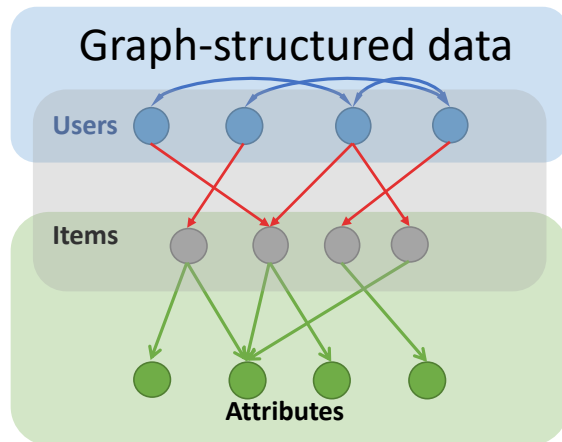
# How to solve such issue?



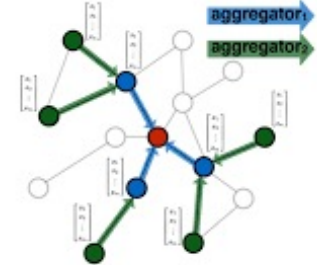
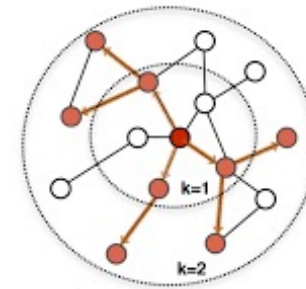
# How to solve such issue?



Explore & Exploit Relations among Instances

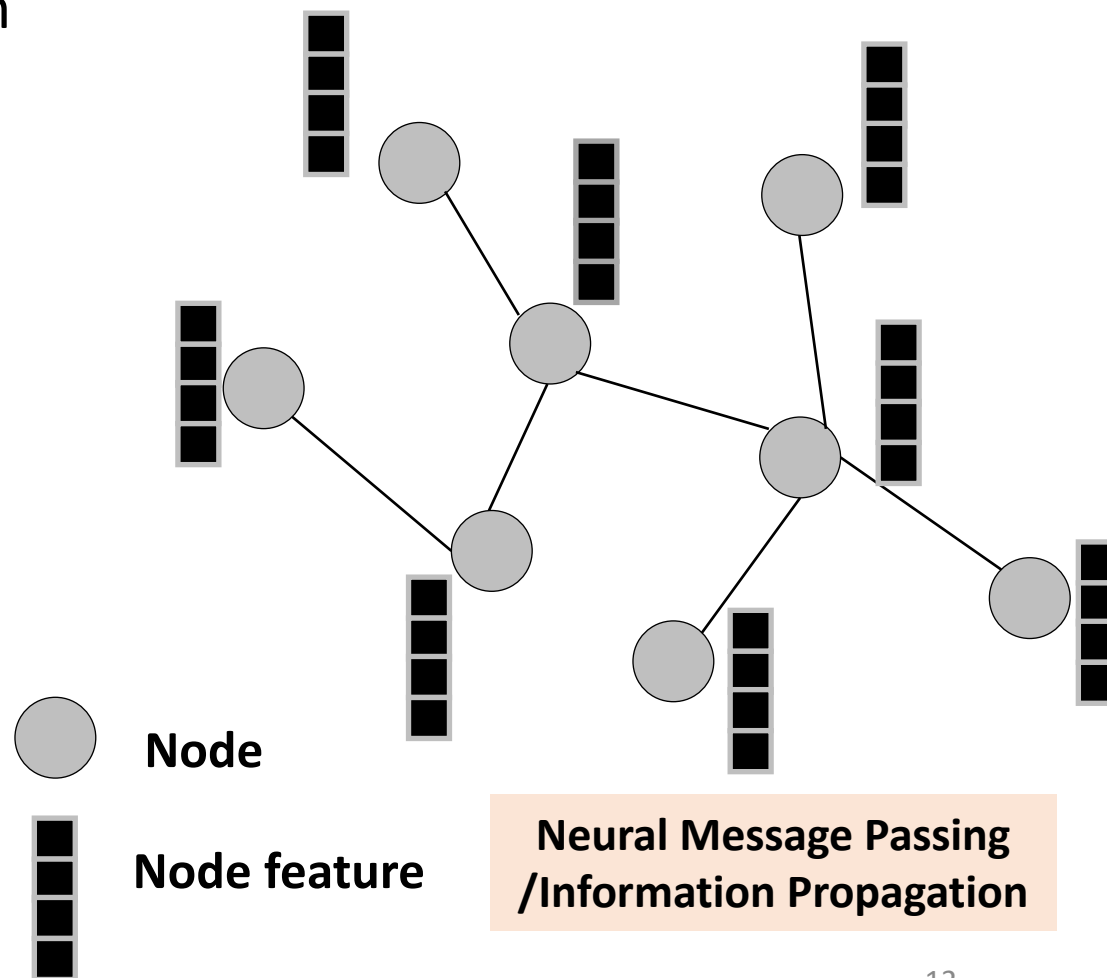


Graph Neural Networks (GNNs)



# Graph Neural Networks (GNNs)

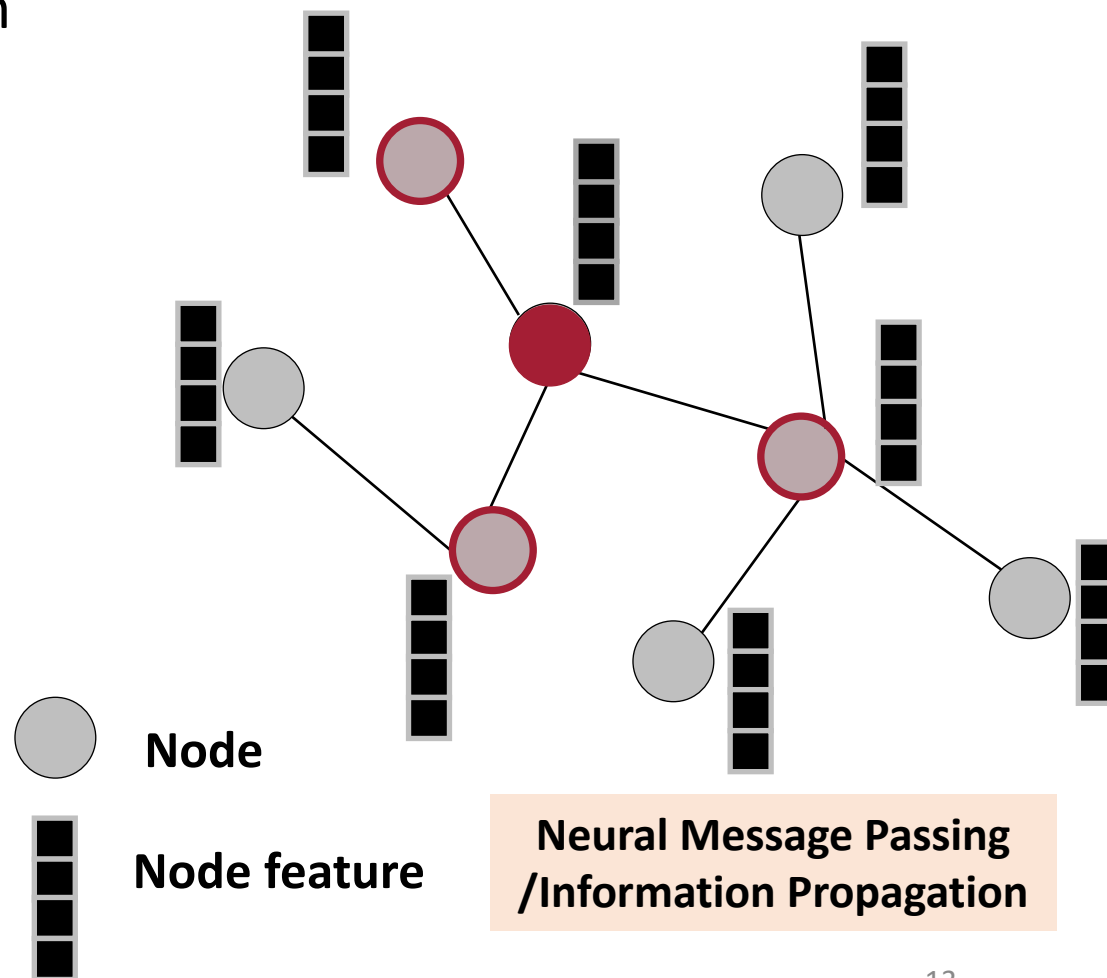
**Key idea:** Generate node embeddings via using neural networks to aggregate information from local neighborhoods.



# Graph Neural Networks (GNNs)

**Key idea:** Generate node embeddings via using neural networks to aggregate information from local neighborhoods.

1. Model a local structural information (neighborhood) of a node;

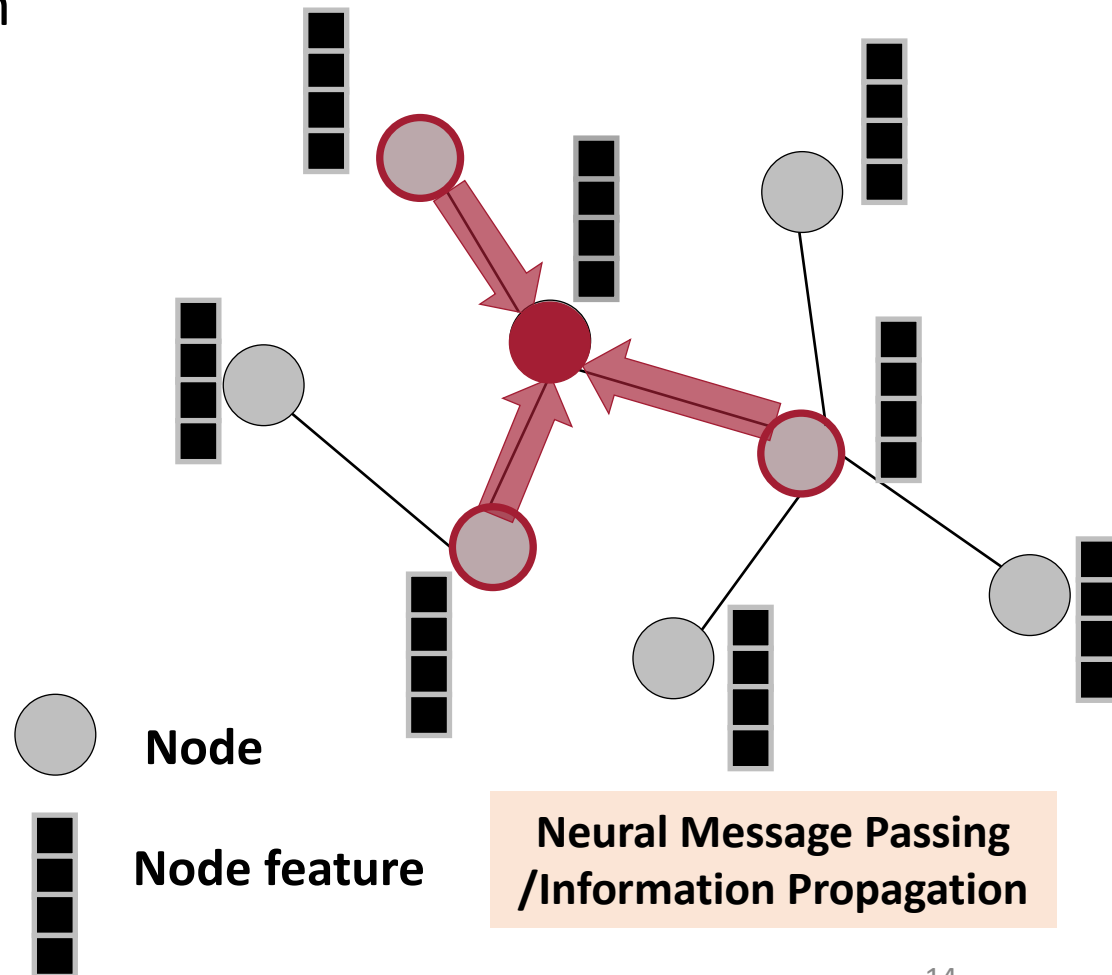


# Graph Neural Networks (GNNs)

**Key idea:** Generate node embeddings via using neural networks to aggregate information from local neighborhoods.

1. Model a local structural information (neighborhood) of a node;
2. Aggregation operation;
3. Representation update.

GNNs can naturally integrate node feature and the topological structure for graph-structured data.



# Graph Neural Networks (GNNs)

**Basic approach:** Average neighbor messages and apply a neural network.

$$\mathbf{h}_v^0 = \mathbf{x}_v$$

Initial 0-th layer embeddings are equal to node  $v$ 's features

$$\mathbf{h}_v^k = \sigma \left( \mathbf{w}_1^k \sum_{u \in N(v)} \frac{\mathbf{h}_u^{k-1}}{\sqrt{|N(u)|}} + \mathbf{w}_2^k \mathbf{h}_v^{k-1} \right)$$

k-th layer embedding of node  $v$

$$\mathbf{z}_v = \mathbf{h}_v^k$$

Embedding after k layers of neighborhood aggregation.

# Graph Neural Networks (GNNs)

**Basic approach:** Average neighbor messages and apply a neural network.

$$\mathbf{h}_v^0 = \mathbf{x}_v$$

Initial 0-th layer embeddings are equal to node  $v$ 's features

Non-linearity (e.g., ReLU or tanh)

$$\mathbf{h}_v^k = \sigma \left( \mathbf{W}_1^k \sum_{u \in N(v)} \frac{\mathbf{h}_u^{k-1}}{\sqrt{|N(u)|}} + \mathbf{W}_2^k \mathbf{h}_v^{k-1} \right)$$

$\mathbf{h}_v^k$ : k-th layer embedding of node  $v$   
 $\sigma$ : Non-linearity (e.g., ReLU or tanh)  
 $\mathbf{W}_1^k, \mathbf{W}_2^k$ : trainable matrices (i.e., what we learn)  
 $\sum_{u \in N(v)} \frac{\mathbf{h}_u^{k-1}}{\sqrt{|N(u)|}}$ : Average of neighbor's previous layer embeddings  
 $\mathbf{h}_v^{k-1}$ : Previous layer embedding of node  $v$

$$\mathbf{z}_v = \mathbf{h}_v^k$$

Embedding after  $k$  layers of neighborhood aggregation.



# Graph Neural Network (GNN)



- Simple neighborhood aggregation:

$$\mathbf{h}_v^k = \sigma \left( \mathbf{W}_1^k \sum_{u \in N(v)} \frac{\mathbf{h}_u^{k-1}}{\sqrt{|N(u)|}} + \mathbf{W}_2^k \mathbf{h}_v^{k-1} \right)$$

- GraphSAGE:

- GAT:

# Graph Neural Network (GNN)



- Simple neighborhood aggregation:

$$\mathbf{h}_v^k = \sigma \left( \mathbf{W}_1^k \sum_{u \in N(v)} \frac{\mathbf{h}_u^{k-1}}{\sqrt{|N(u)|}} + \mathbf{W}_2^k \mathbf{h}_v^{k-1} \right)$$

- GraphSAGE:

$$\mathbf{h}_v^k = \sigma \left( [\mathbf{W}_1^k \cdot \text{AGG}(\{\mathbf{h}_u^{k-1}, \forall u \in N(u)\}), \mathbf{W}_2^k \cdot \mathbf{h}_v^k] \right)$$

Generalized Aggregation: mean, pooling, LSTM

- GAT:

# Graph Neural Network (GNN)

- Simple neighborhood aggregation:

$$\mathbf{h}_v^k = \sigma \left( \mathbf{W}_1^k \sum_{u \in N(v)} \frac{\mathbf{h}_u^{k-1}}{\sqrt{|N(u)|}} + \mathbf{W}_2^k \mathbf{h}_v^{k-1} \right)$$

- GraphSAGE:

$$\mathbf{h}_v^k = \sigma \left( [\mathbf{W}_1^k \cdot \text{AGG}(\{\mathbf{h}_u^{k-1}, \forall u \in N(u)\}), \mathbf{W}_2^k \cdot \mathbf{h}_v^k] \right)$$

Generalized Aggregation: mean, pooling, LSTM

- GAT:

$$\mathbf{h}_v^k = \sigma \left( \sum_{u \in N(v)} \alpha_{v,u} \mathbf{W}^k \mathbf{h}_u^{k-1} \right)$$

Learned attention weights

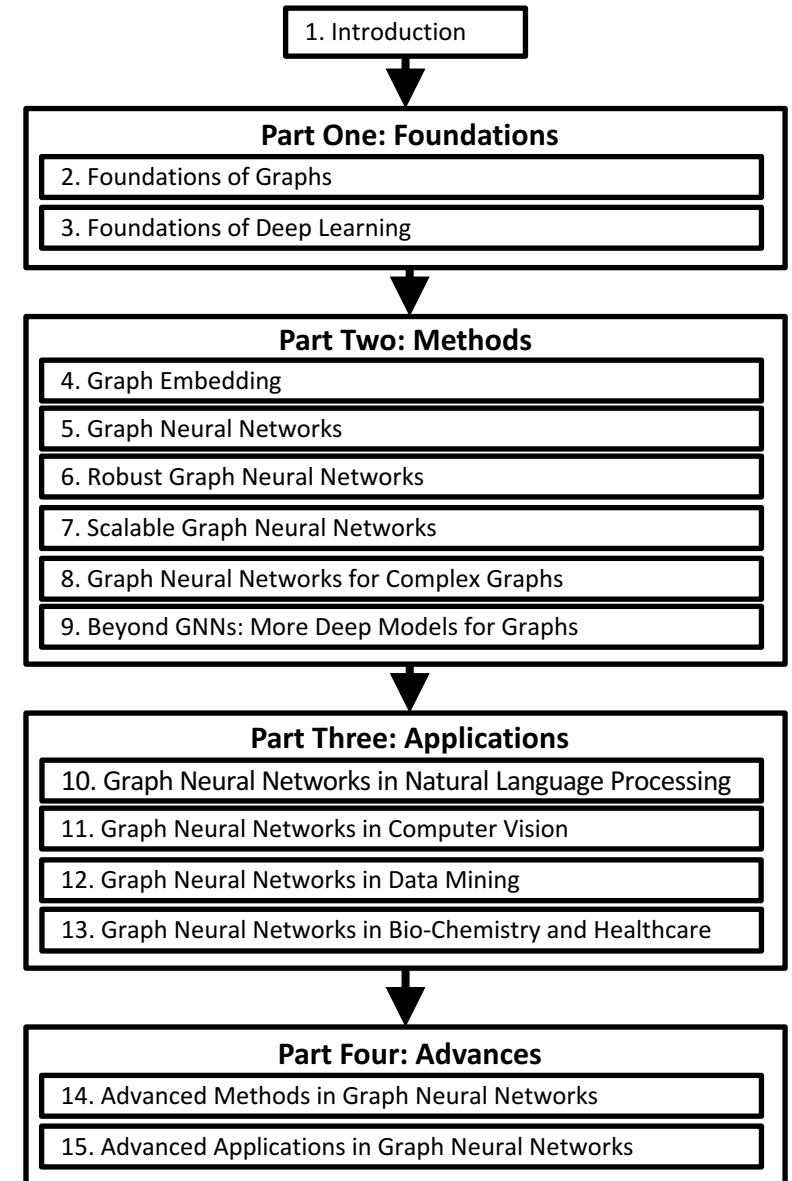
# Book: Deep Learning on Graphs



[https://cse.msu.edu/~mayao4/dlg\\_book/](https://cse.msu.edu/~mayao4/dlg_book/)



Yao Ma and Jiliang Tang, MSU



# GNNs based Recommendation



## ■ Collaborative Filtering

- Graph Convolutional Neural Networks for Web-Scale Recommender Systems (KDD'18)
- Graph Convolutional Matrix Completion (KDD'18 Deep Learning Day )
- Neural Graph Collaborative Filtering (SIGIR'19)
- LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation (SIGIR'20)

## ■ Collaborative Filtering with Side Information (Users/Items)

### □ Social Recommendation (Users)

- Graph Neural Network for Social Recommendation (WWW'19)
- A Neural Influence Diffusion Model for Social Recommendation (SIGIR'19)
- A Graph Neural Network Framework for Social Recommendations (TKDE'20)

### □ Knowledge-graph-aware Recommendation (Items)

- Knowledge Graph Convolutional Networks for Recommender Systems with Label Smoothness Regularization (KDD'19 and WWW'19)
- KGAT: Knowledge Graph Attention Network for Recommendation (KDD'19)

# GNNs based Recommendation



## ■ Collaborative Filtering

- Graph Convolutional Neural Networks for Web-Scale Recommender Systems (KDD'18)
- Graph Convolutional Matrix Completion (KDD'18 Deep Learning Day )
- Neural Graph Collaborative Filtering (SIGIR'19)
- LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation (SIGIR'20)

## ■ Collaborative Filtering with Side Information (Users/Items)

### □ Social Recommendation (Users)

- Graph Neural Network for Social Recommendation (WWW'19)
- A Neural Influence Diffusion Model for Social Recommendation (SIGIR'19)
- A Graph Neural Network Framework for Social Recommendations (TKDE'20)

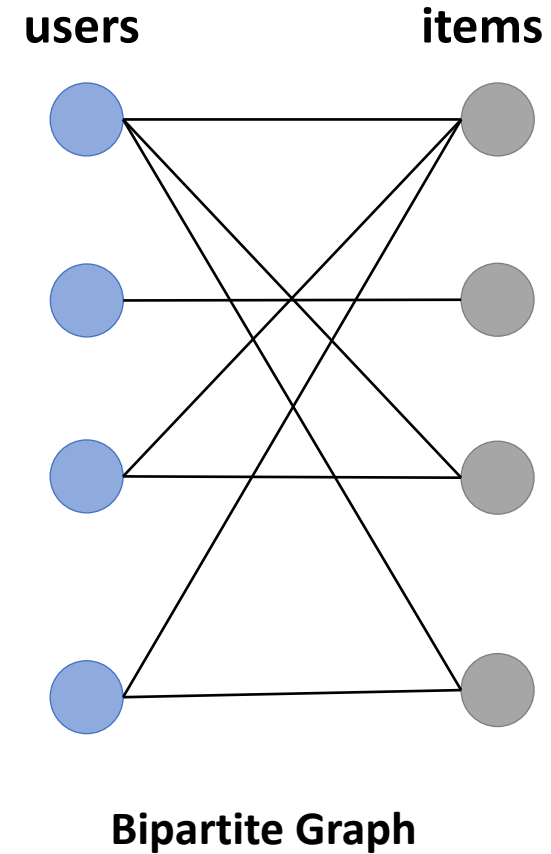
### □ Knowledge-graph-aware Recommendation (Items)

- Knowledge Graph Convolutional Networks for Recommender Systems with Label Smoothness Regularization (KDD'19 and WWW'19)
- KGAT: Knowledge Graph Attention Network for Recommendation (KDD'19)

# Interactions as Bipartite Graph

|       | items |   |   |   |
|-------|-------|---|---|---|
| users | 1     | 0 | 1 | 1 |
|       | 0     | 1 | 0 | 0 |
|       | 1     | 1 | 0 | 0 |
|       | 1     | 0 | 0 | 1 |

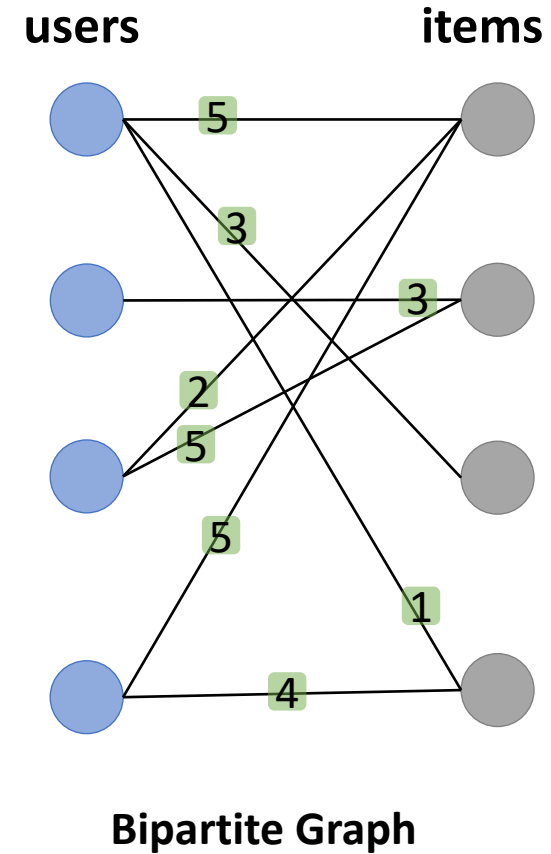
0/1 Interaction matrix



# Interactions as Bipartite Graph

|       | items |   |   |   |
|-------|-------|---|---|---|
| users | 5     | 0 | 3 | 1 |
|       | 0     | 3 | 0 | 0 |
|       | 2     | 5 | 0 | 0 |
|       | 5     | 0 | 0 | 4 |

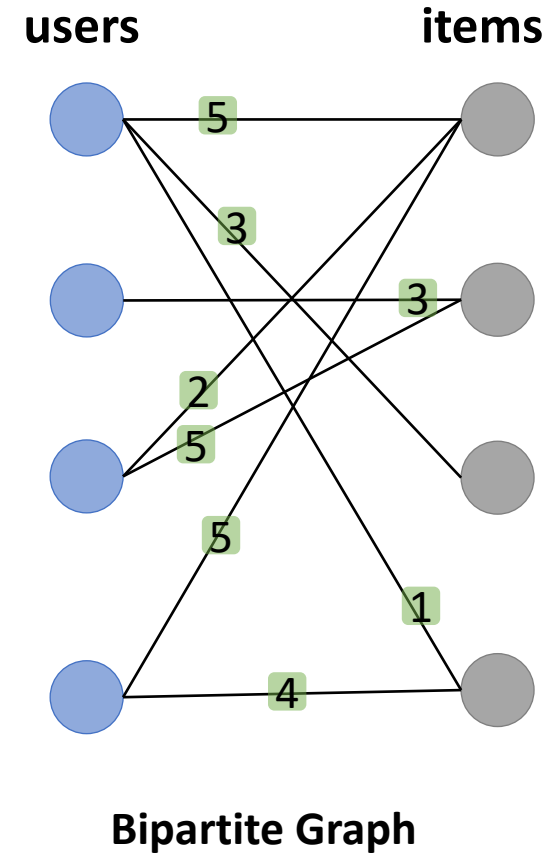
Weighted interaction matrix





## User representation learning

Aggregate for each rating:  $\mu_{i,r} = \sum_{j \in \mathcal{N}_{i,r}} \frac{1}{c_{ij}} W_r x_j$

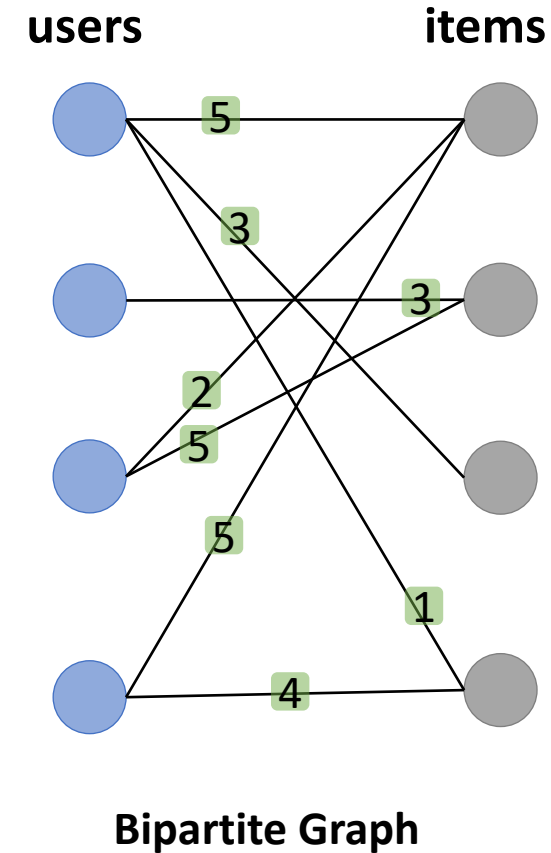


## User representation learning

Aggregate for each rating:  $\mu_{i,r} = \sum_{j \in \mathcal{N}_{i,r}} \frac{1}{c_{ij}} W_r x_j$

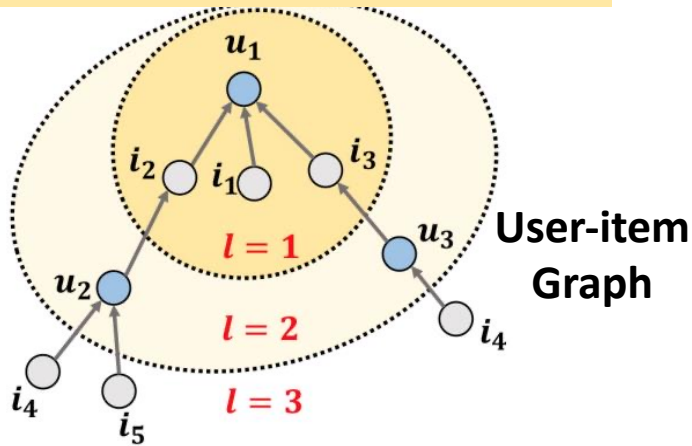
$$u_i = \mathbf{W} \cdot \sigma(\text{accum}(u_{i,1}, \dots, u_{i,R}))$$

Item representation learning in a similar way



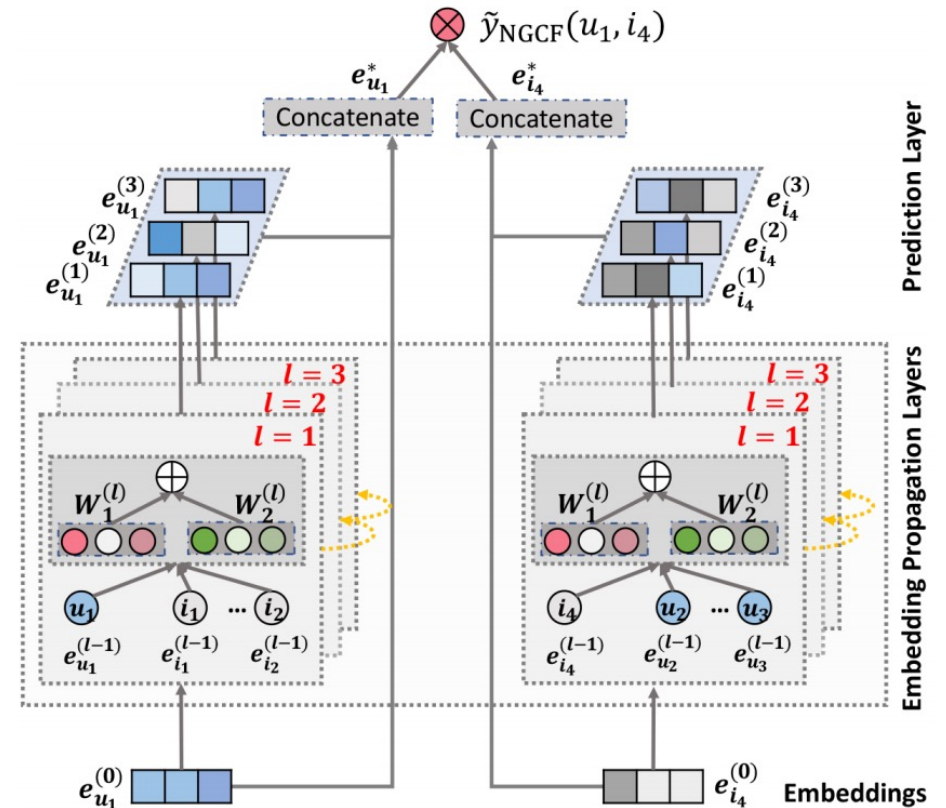
# NGCF

## High-order Connectivity for $u_1$



## Embedding Propagation, inspired by GNNs

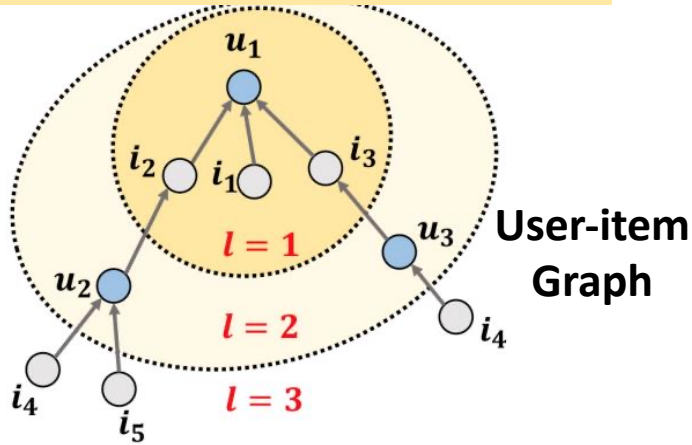
- Propagate embeddings recursively on the user-item graph
- Construct information flows in the embedding space



# NGCF



## High-order Connectivity for $u_1$



## Embedding Propagation, inspired by GNNs

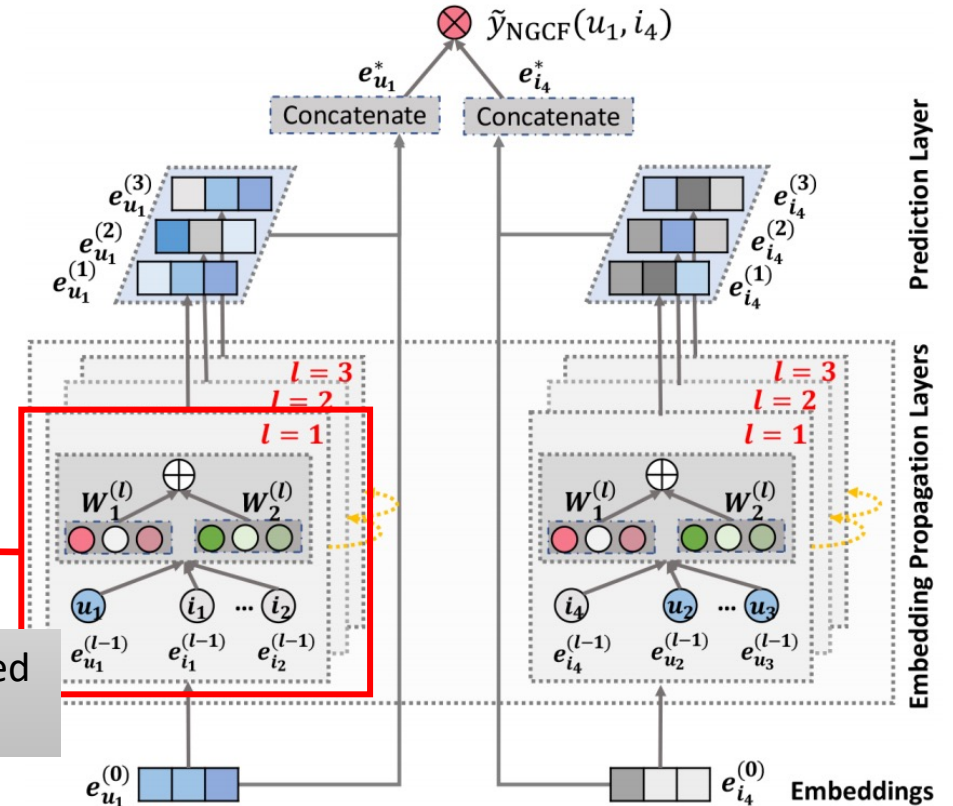
- Propagate embeddings recursively on the user-item graph
- Construct information flows in the embedding space

$$\mathbf{e}_u^{(l)} = \text{LeakyReLU}\left(\mathbf{m}_{u \leftarrow u}^{(l)} + \sum_{i \in \mathcal{N}_u} \mathbf{m}_{u \leftarrow i}^{(l)}\right)$$

$$\begin{cases} \mathbf{m}_{u \leftarrow i}^{(l)} = p_{ui} \left( \mathbf{W}_1^{(l)} \mathbf{e}_i^{(l-1)} + \mathbf{W}_2^{(l)} (\mathbf{e}_i^{(l-1)} \odot \mathbf{e}_u^{(l-1)}) \right) \\ \mathbf{m}_{u \leftarrow u}^{(l)} = \mathbf{W}_1^{(l)} \mathbf{e}_u^{(l-1)} \end{cases}$$

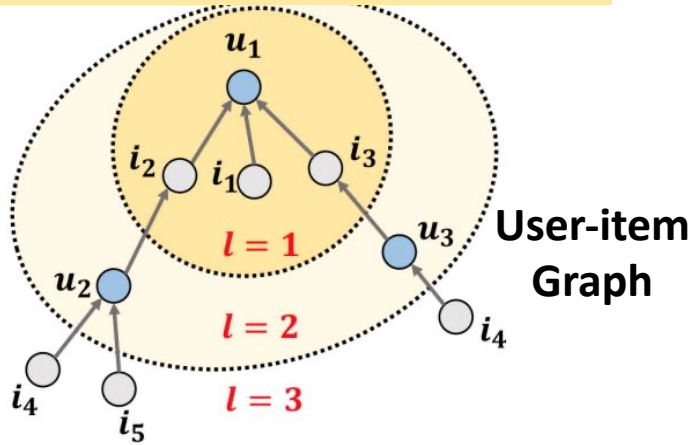
Self-connections

collaborative signal: message passed from interacted items to  $u$



# NGCF

## High-order Connectivity for $u_1$



## Embedding Propagation, inspired by GNNs

- Propagate embeddings recursively on the user-item graph
- Construct information flows in the embedding space

$$\mathbf{e}_u^* = \mathbf{e}_u^{(0)} \parallel \dots \parallel \mathbf{e}_u^{(L)}, \quad \mathbf{e}_i^* = \mathbf{e}_i^{(0)} \parallel \dots \parallel \mathbf{e}_i^{(L)},$$

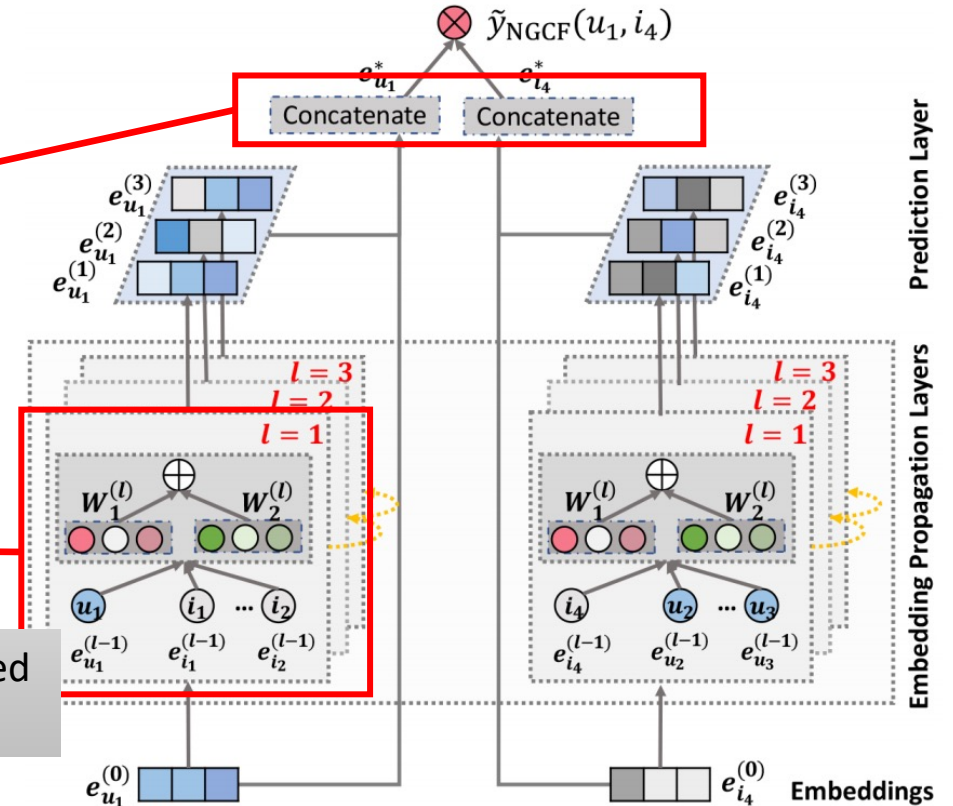
$$\mathbf{e}_u^{(l)} = \text{LeakyReLU}\left(\mathbf{m}_{u \leftarrow u}^{(l)} + \sum_{i \in \mathcal{N}_u} \mathbf{m}_{u \leftarrow i}^{(l)}\right),$$

$$\begin{cases} \mathbf{m}_{u \leftarrow i}^{(l)} = p_{ui} \left( \mathbf{W}_1^{(l)} \mathbf{e}_i^{(l-1)} + \mathbf{W}_2^{(l)} (\mathbf{e}_i^{(l-1)} \odot \mathbf{e}_u^{(l-1)}) \right) \\ \mathbf{m}_{u \leftarrow u}^{(l)} = \mathbf{W}_1^{(l)} \mathbf{e}_u^{(l-1)} \end{cases}$$

Self-connections

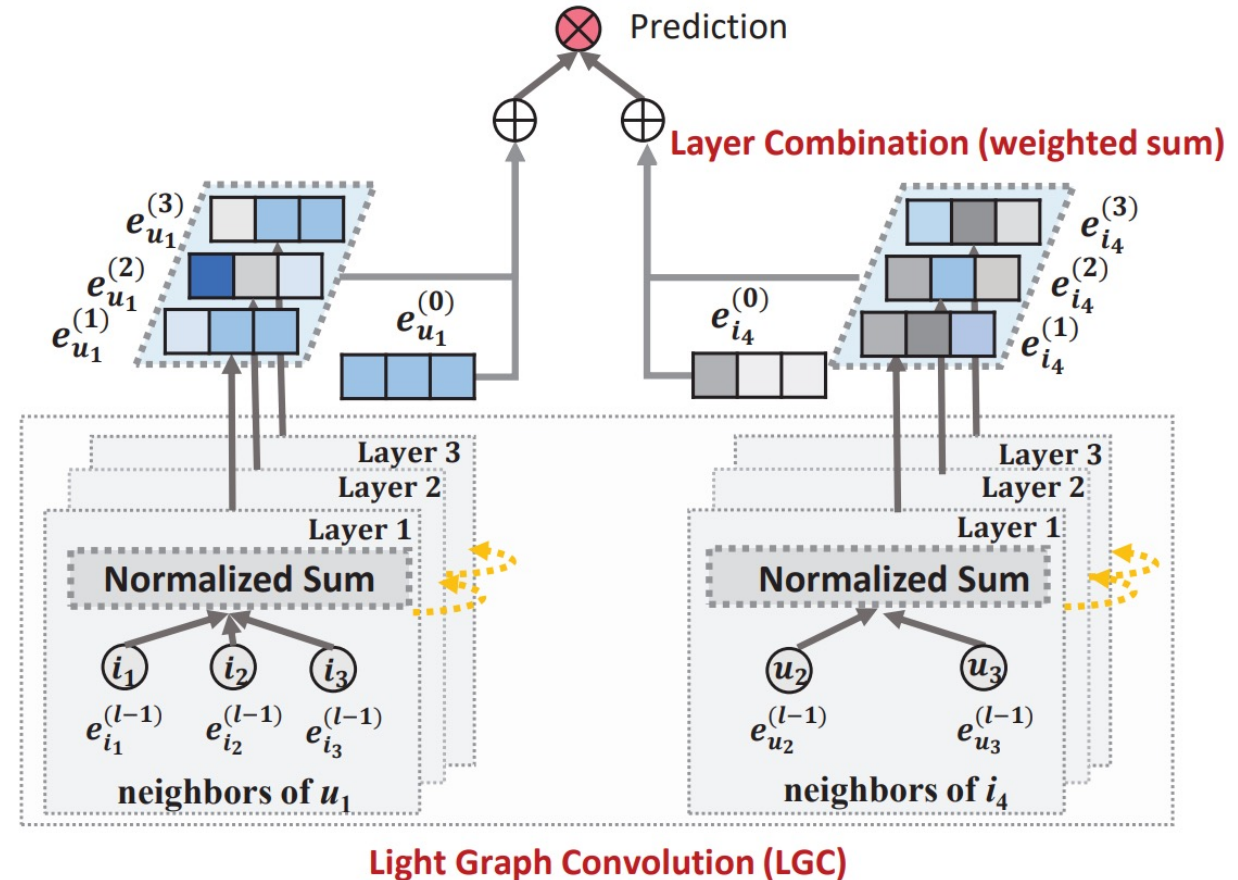
collaborative signal: message passed from interacted items to  $u$

Different layers



# LightGCN

## Simplifying GCN for recommendation



discard feature transformation and nonlinear activation

# LightGCN

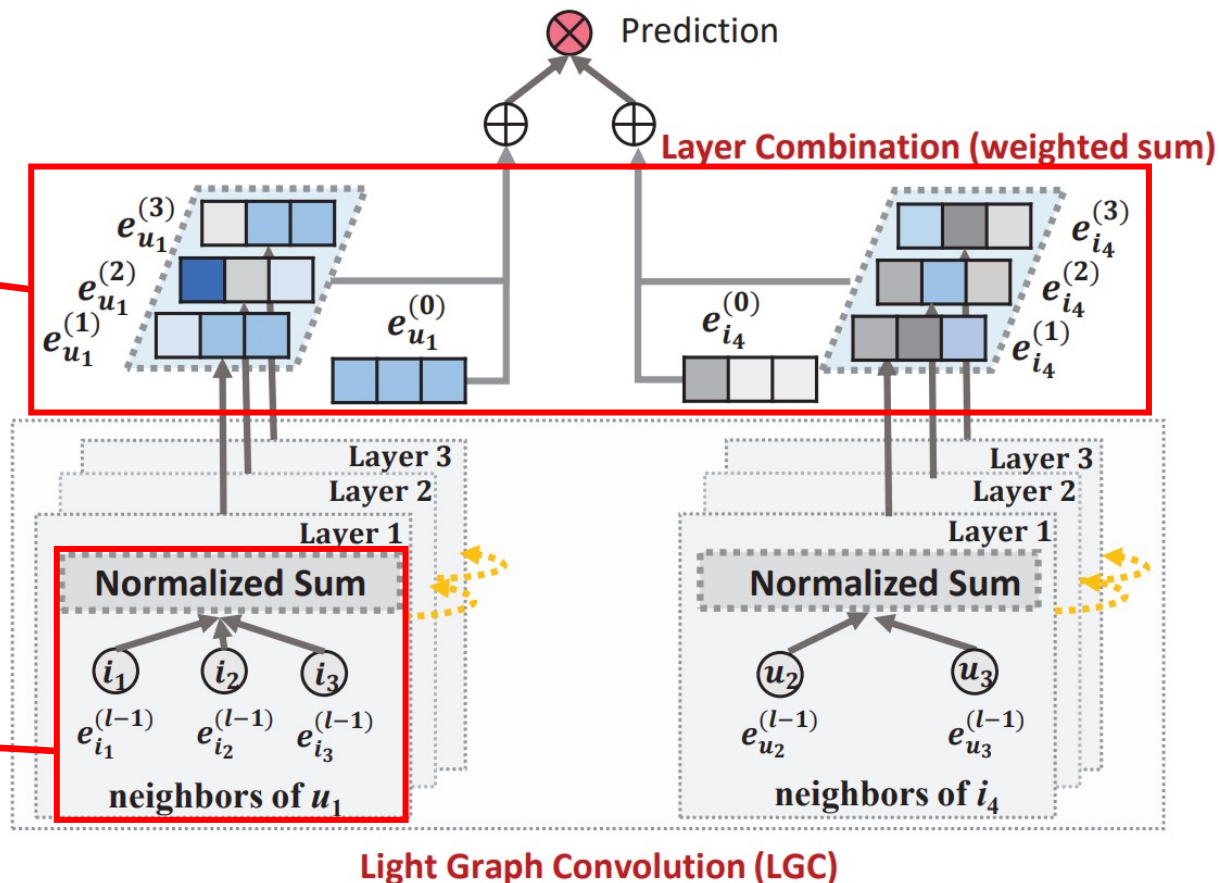


## Simplifying GCN for recommendation

$$\mathbf{e}_u = \sum_{k=0}^K \alpha_k \mathbf{e}_u^{(k)}; \quad \mathbf{e}_i = \sum_{k=0}^K \alpha_k \mathbf{e}_i^{(k)},$$

$$\mathbf{e}_u^{(k+1)} = \sum_{i \in \mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u|} \sqrt{|\mathcal{N}_i|}} \mathbf{e}_i^{(k)},$$

$$\mathbf{e}_i^{(k+1)} = \sum_{u \in \mathcal{N}_i} \frac{1}{\sqrt{|\mathcal{N}_i|} \sqrt{|\mathcal{N}_u|}} \mathbf{e}_u^{(k)}.$$



discard feature transformation and nonlinear activation

# GNN based Recommendation



## ■ Collaborative Filtering

- Graph Convolutional Neural Networks for Web-Scale Recommender Systems (KDD'18)
- Graph Convolutional Matrix Completion (KDD'18 Deep Learning Day )
- Neural Graph Collaborative Filtering (SIGIR'19)
- LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation (SIGIR'20)

## ■ Collaborative Filtering with Side Information (Users/Items)

### □ Social Recommendation (Users)

- Graph Neural Network for Social Recommendation (WWW'19)
- A Neural Influence Diffusion Model for Social Recommendation (SIGIR'19)
- A Graph Neural Network Framework for Social Recommendations (TKDE'20)

### □ Knowledge-graph-aware Recommendation (Items)

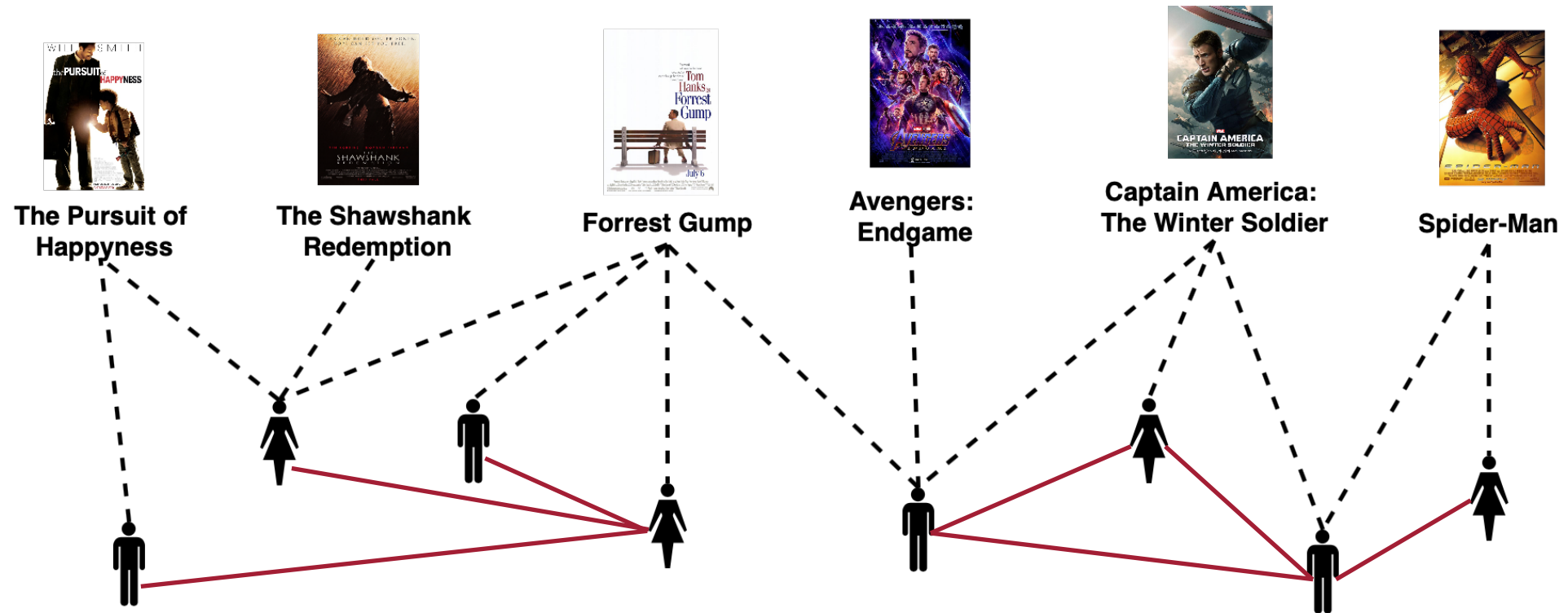
- Knowledge Graph Convolutional Networks for Recommender Systems with Label Smoothness Regularization (KDD'19 and WWW'19)
- KGAT: Knowledge Graph Attention Network for Recommendation (KDD'19)



# Social Recommendation

## Side information about users: social networks

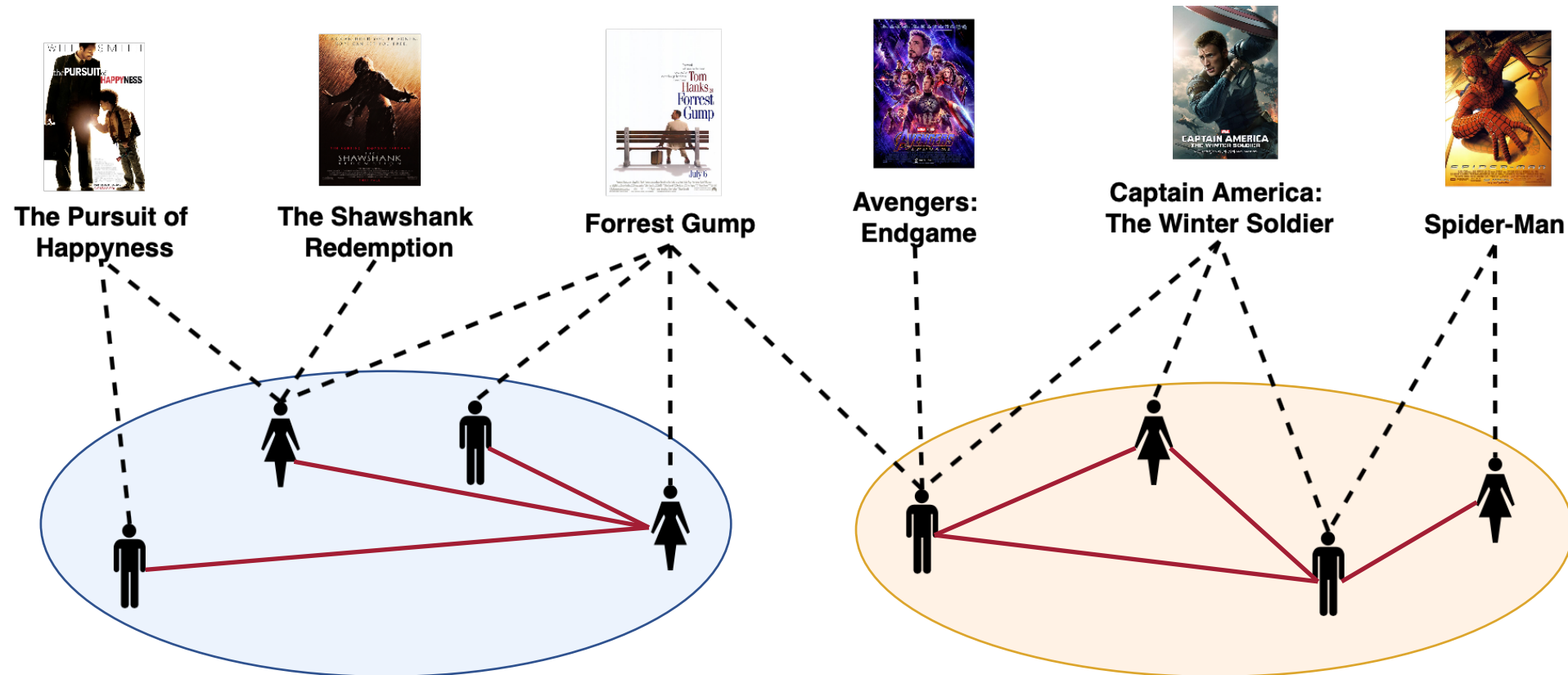
- ❑ Users' preferences **are similar to or influenced by** the people around them (nearer neighbours)  
[Tang et. al, 2013]



# Social Recommendation

## Side information about users: social networks

- ❑ Users' preferences **are similar to or influenced by** the people around them (nearer neighbours)  
[Tang et. al, 2013]

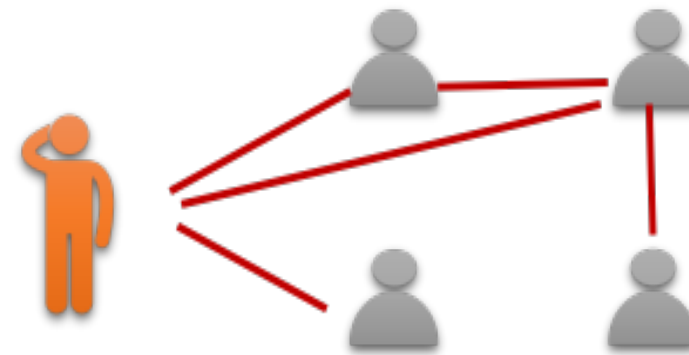


# GraphRec

## Graph Data in Social Recommendation



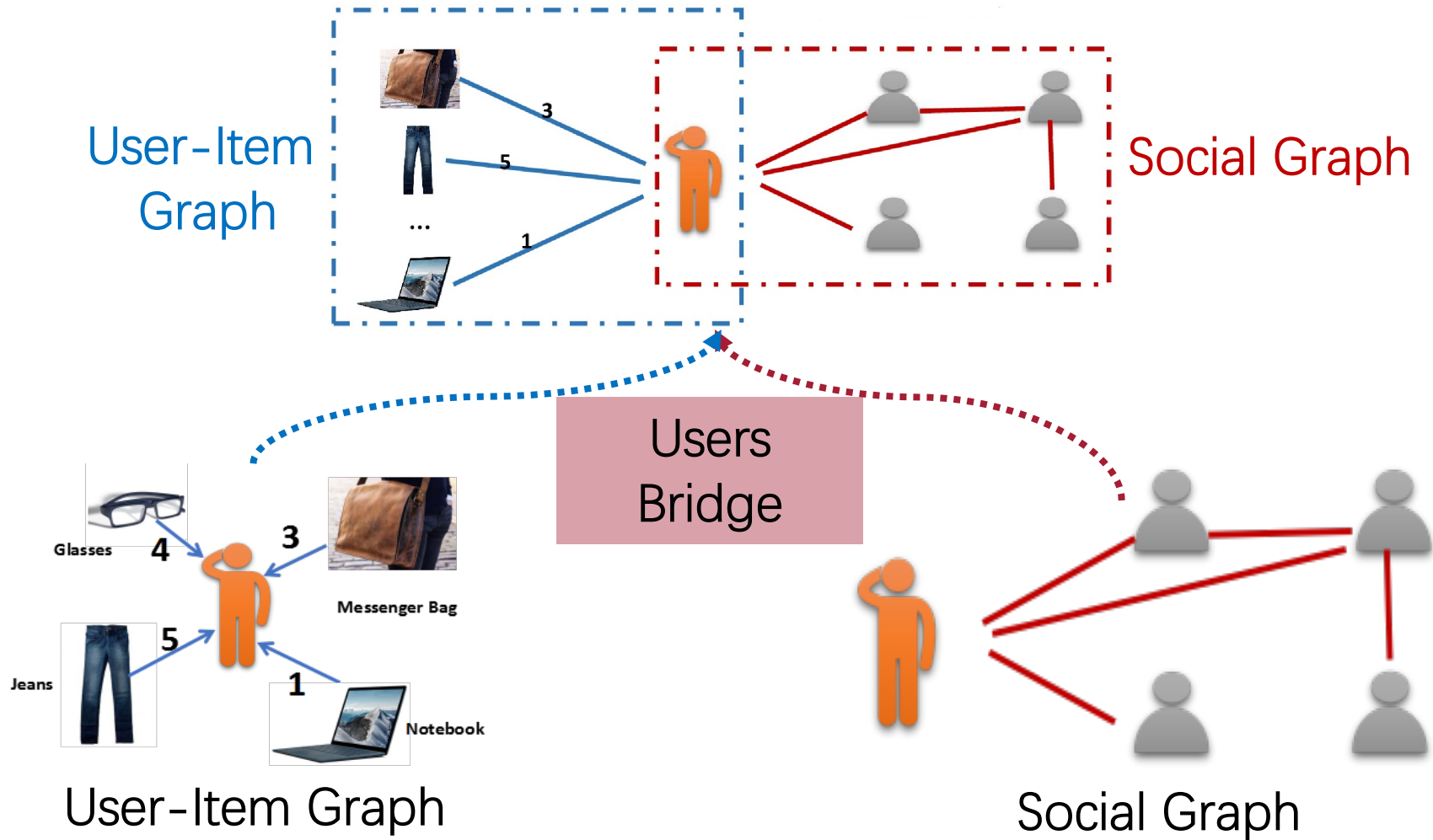
User-Item Graph



Social Graph

# GraphRec

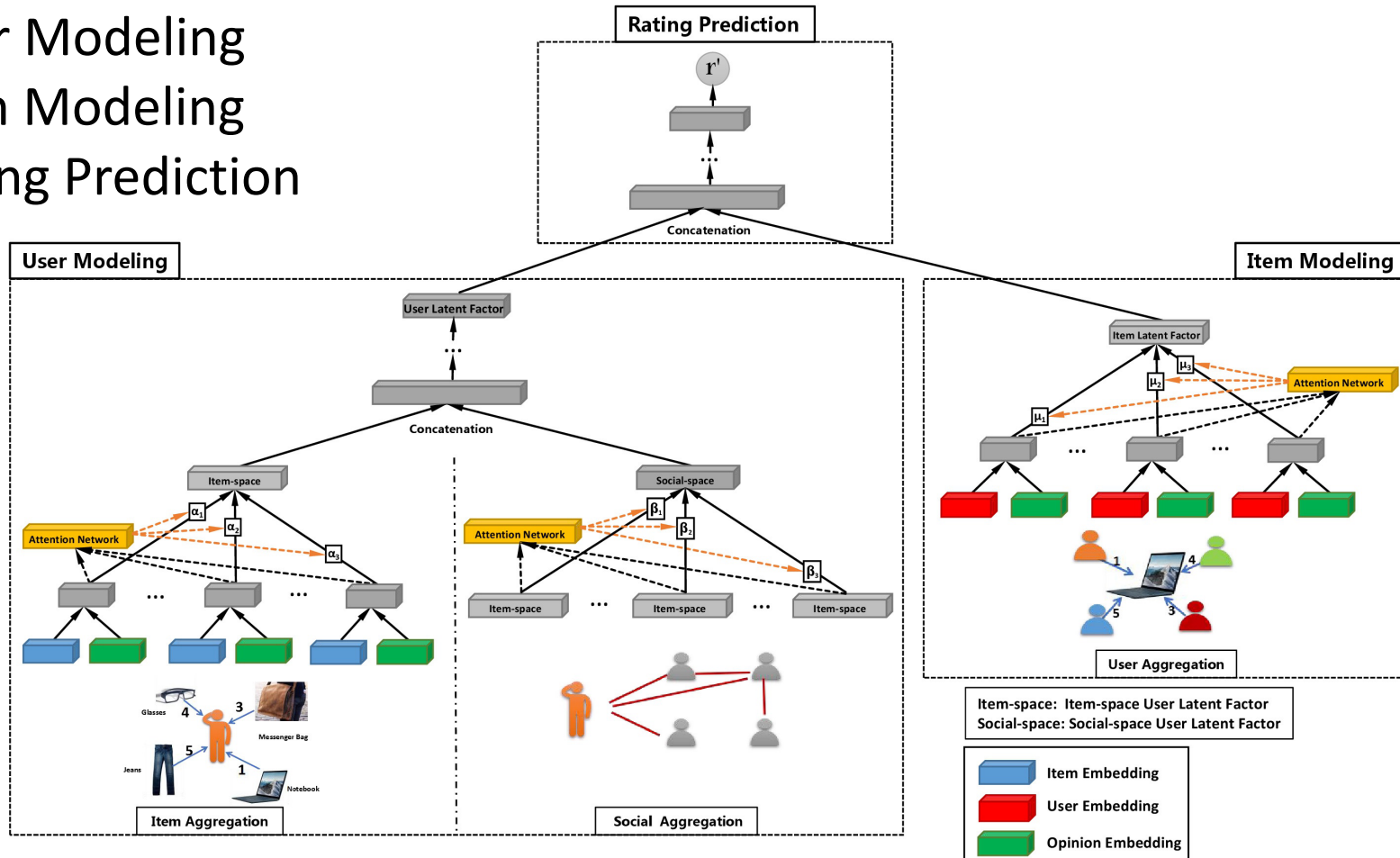
## Graph Data in Social Recommendation



# GraphRec

Three Components:

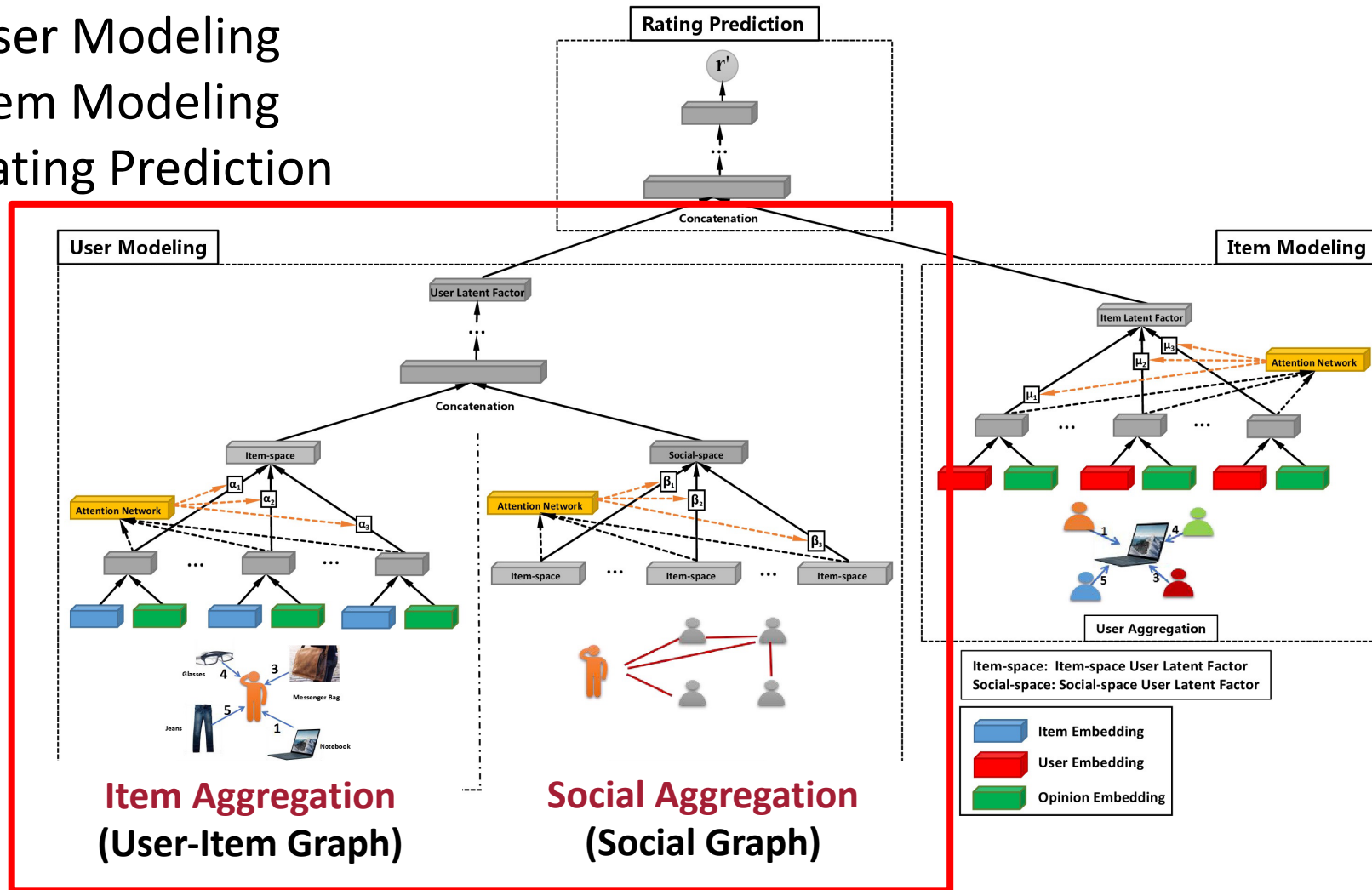
- ❑ User Modeling
- ❑ Item Modeling
- ❑ Rating Prediction



# GraphRec

Three Components:

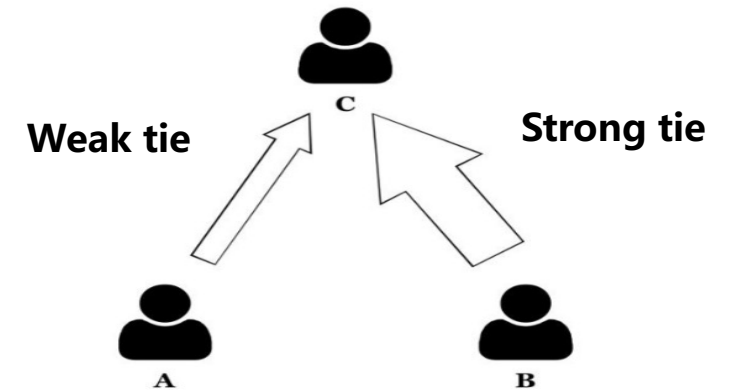
- ❑ User Modeling
- ❑ Item Modeling
- ❑ Rating Prediction



# GraphRec: User Modeling

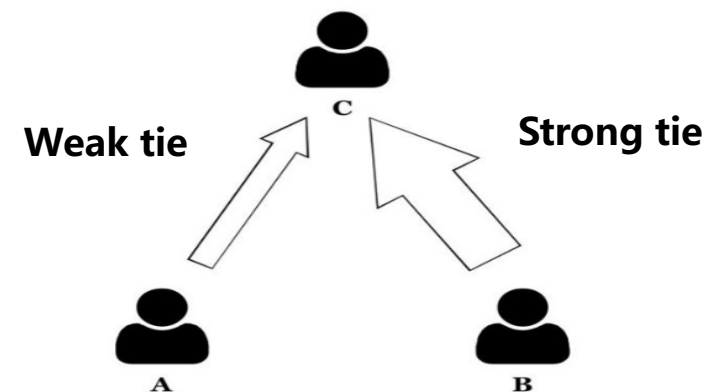


- ❑ Social Aggregation in user-user social graph
- ❑ Users are likely to share more similar tastes with strong ties than weak ties.



# GraphRec: User Modeling

- Social Aggregation in user-user social graph
- Users are likely to share more similar tastes with strong ties than weak ties.

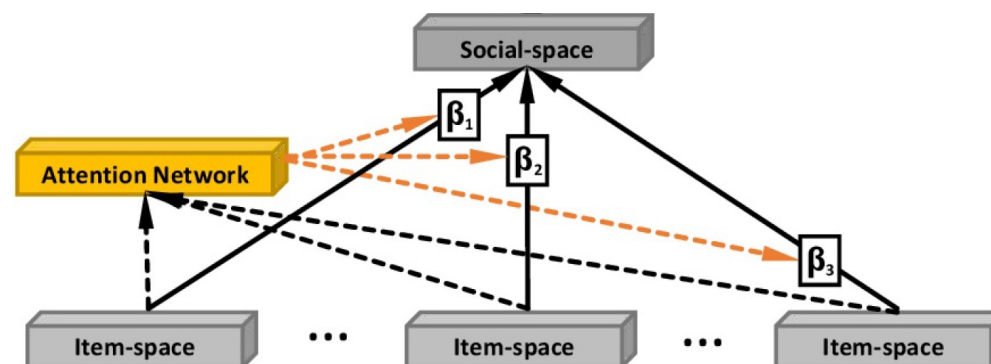


Attention network to differentiate the importance weight.

Aggregating item-space users messages from social neighbors

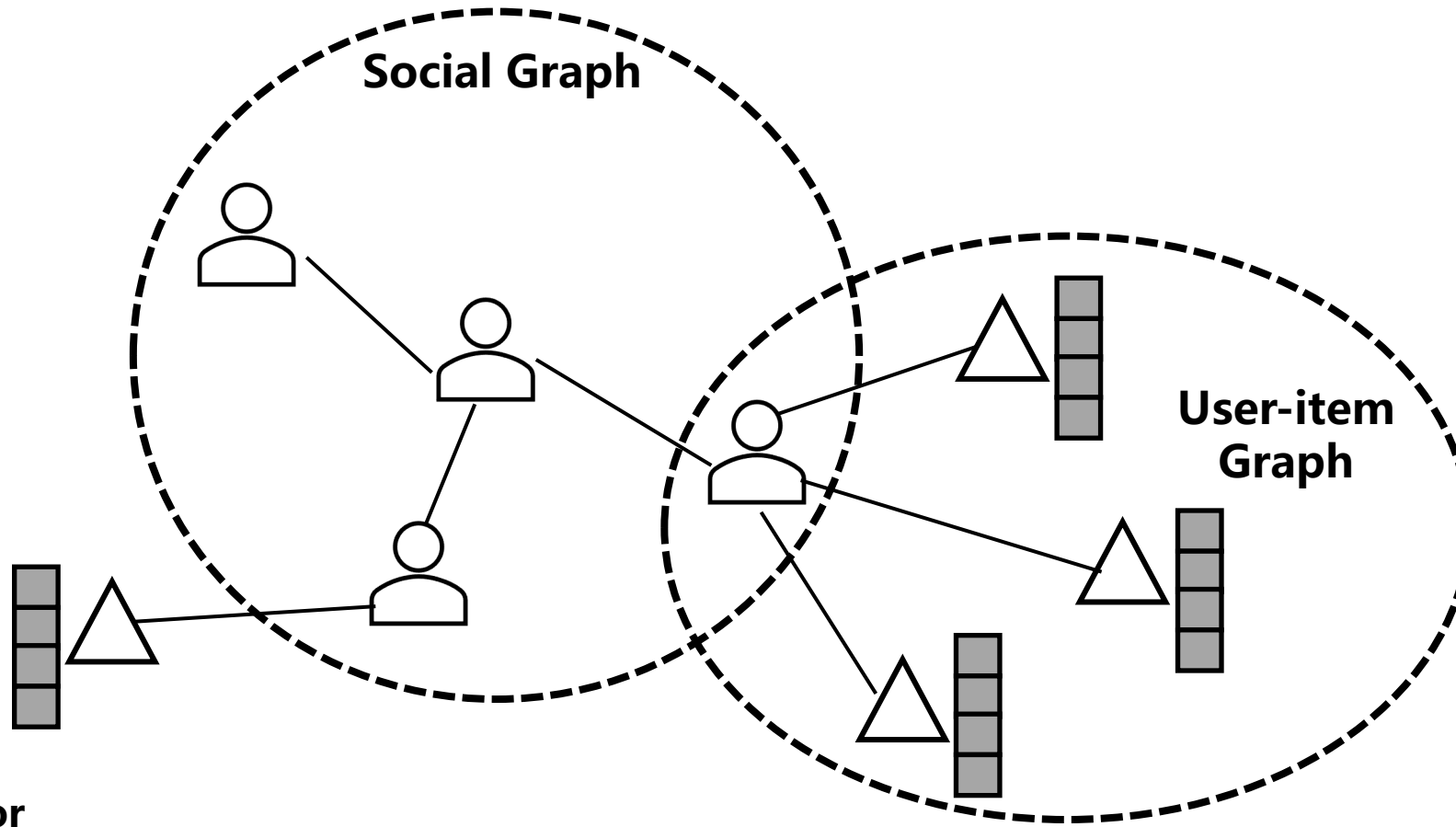
$$\mathbf{h}_i^S = \sigma(\mathbf{W} \cdot \left\{ \sum_{o \in N(i)} \beta_{io} \mathbf{h}_o^I \right\} + \mathbf{b})$$

attentive weight





# User Modeling: Social Aggregation



 Social-space user latent factor

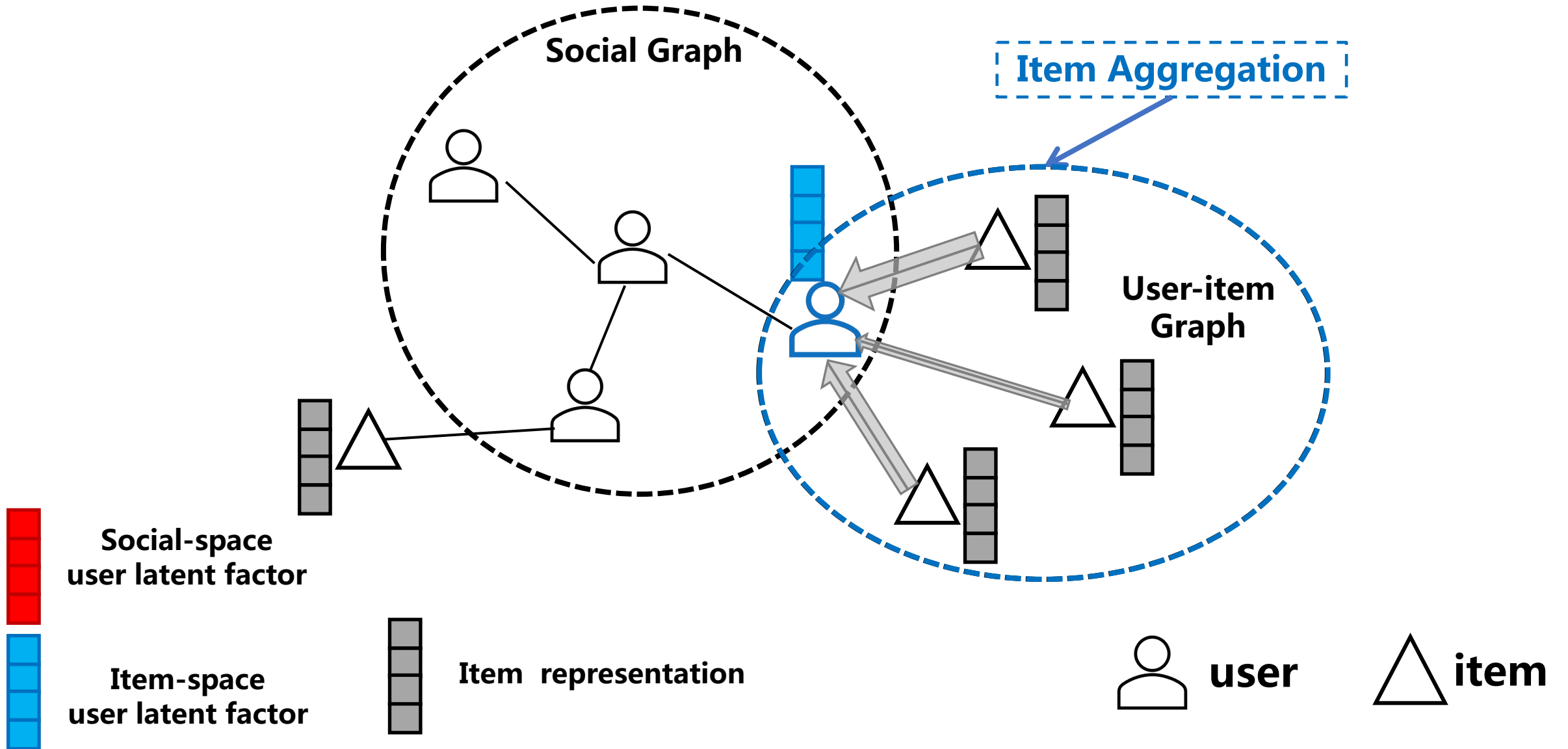
 Item-space user latent factor

 Item representation

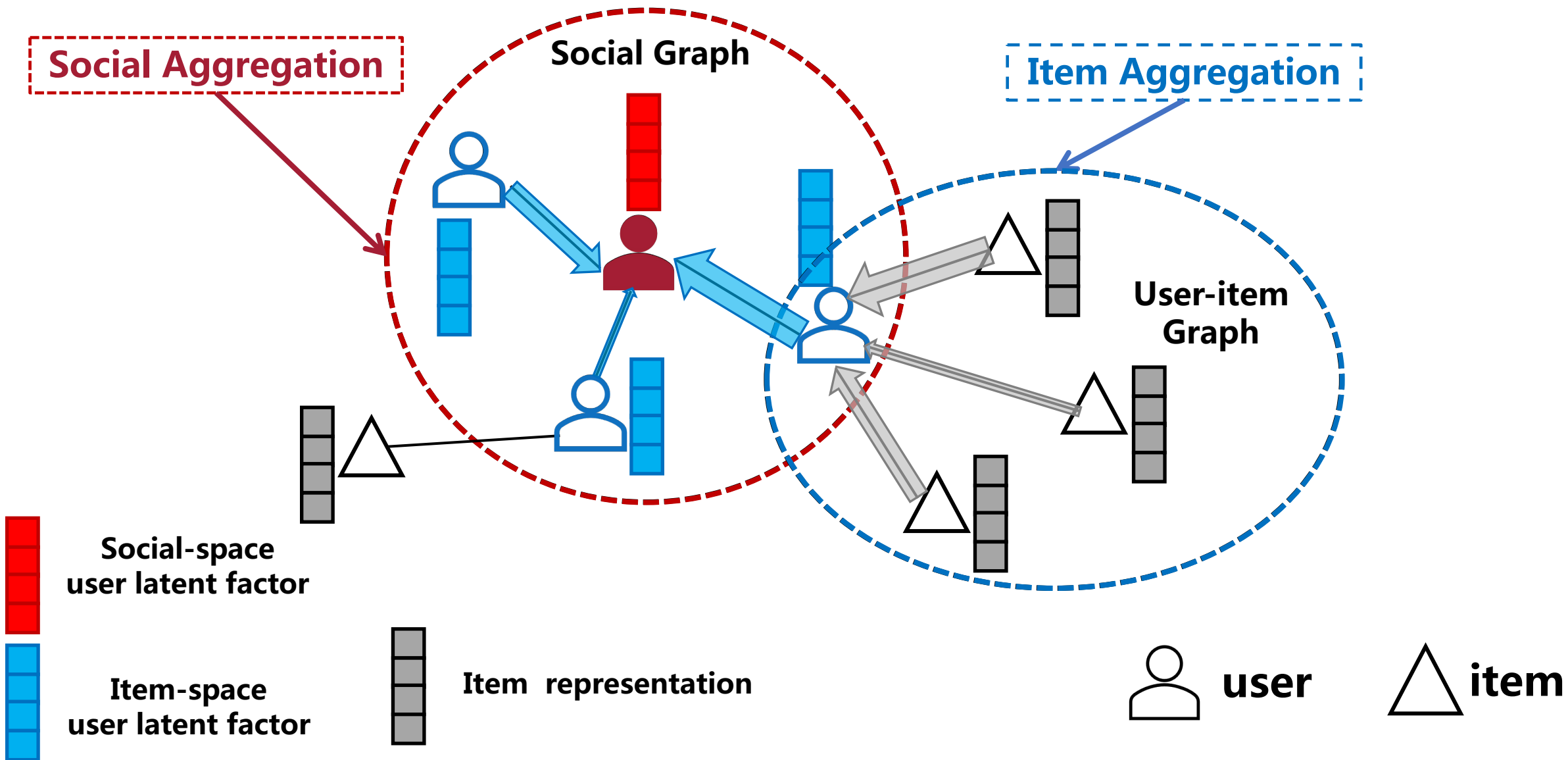
 user

 item

# User Modeling: Social Aggregation



# User Modeling: Social Aggregation



# GNNs based Recommendation



## ■ Collaborative Filtering

- Graph Convolutional Neural Networks for Web-Scale Recommender Systems (KDD'18)
- Graph Convolutional Matrix Completion (KDD'18 Deep Learning Day )
- Neural Graph Collaborative Filtering (SIGIR'19)
- LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation (SIGIR'20)

## ■ Collaborative Filtering with Side Information (Users/Items)

### □ Social Recommendation (Users)

- Graph Neural Network for Social Recommendation (WWW'19)
- A Neural Influence Diffusion Model for Social Recommendation (SIGIR'19)
- A Graph Neural Network Framework for Social Recommendations (TKDE'20)

### □ Knowledge-graph-aware Recommendation (Items)

- Knowledge Graph Convolutional Networks for Recommender Systems with Label Smoothness Regularization (KDD'19 and WWW'19)
- KGAT: Knowledge Graph Attention Network for Recommendation (KDD'19)

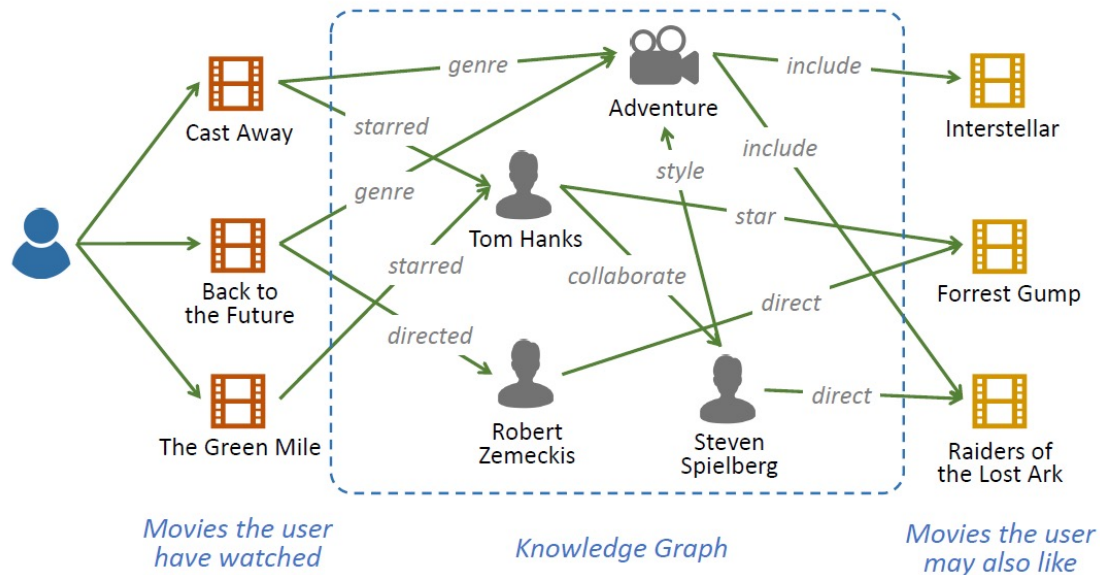
# KGCN (WWW'19)

## Side information about items: Knowledge Graph (KG)

### Heterogeneous Graph:

- Nodes: entities (Items)
- Edges: relations

**Triples: (head, relation, tail)**



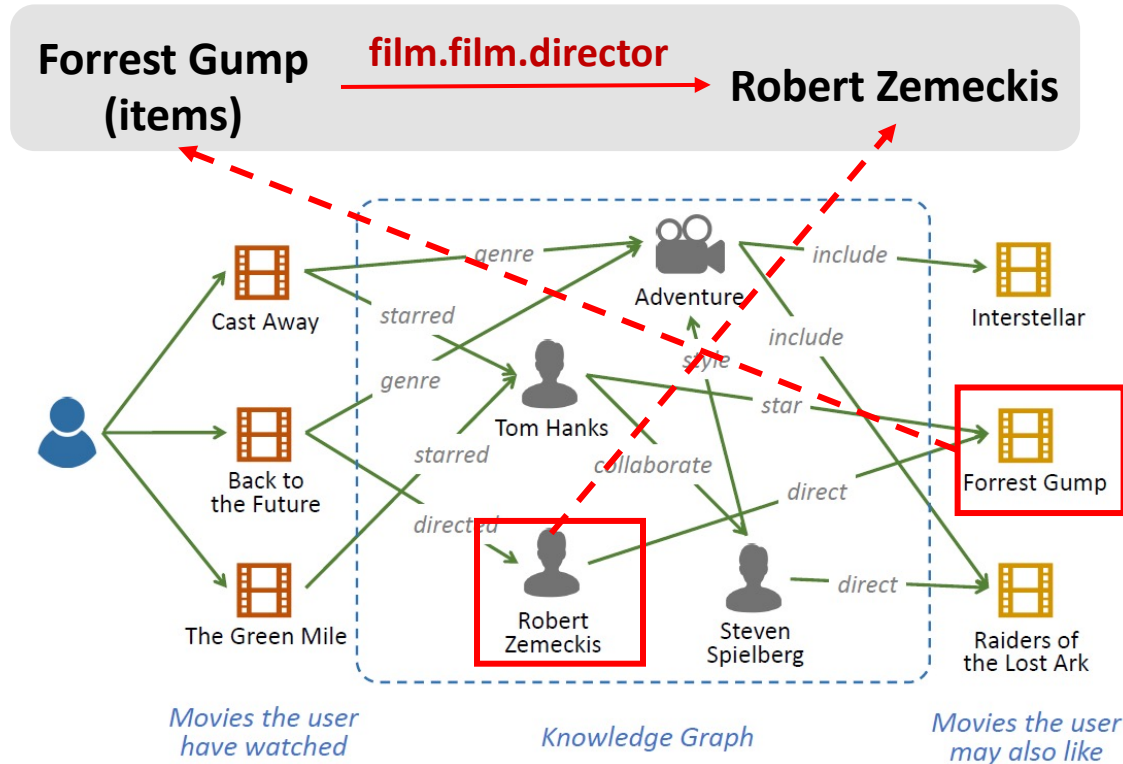
# KGCN (WWW'19)

## Side information about items: Knowledge Graph (KG)

### Heterogeneous Graph:

- Nodes: entities (Items)
- Edges: relations

Triples: (head, relation, tail)



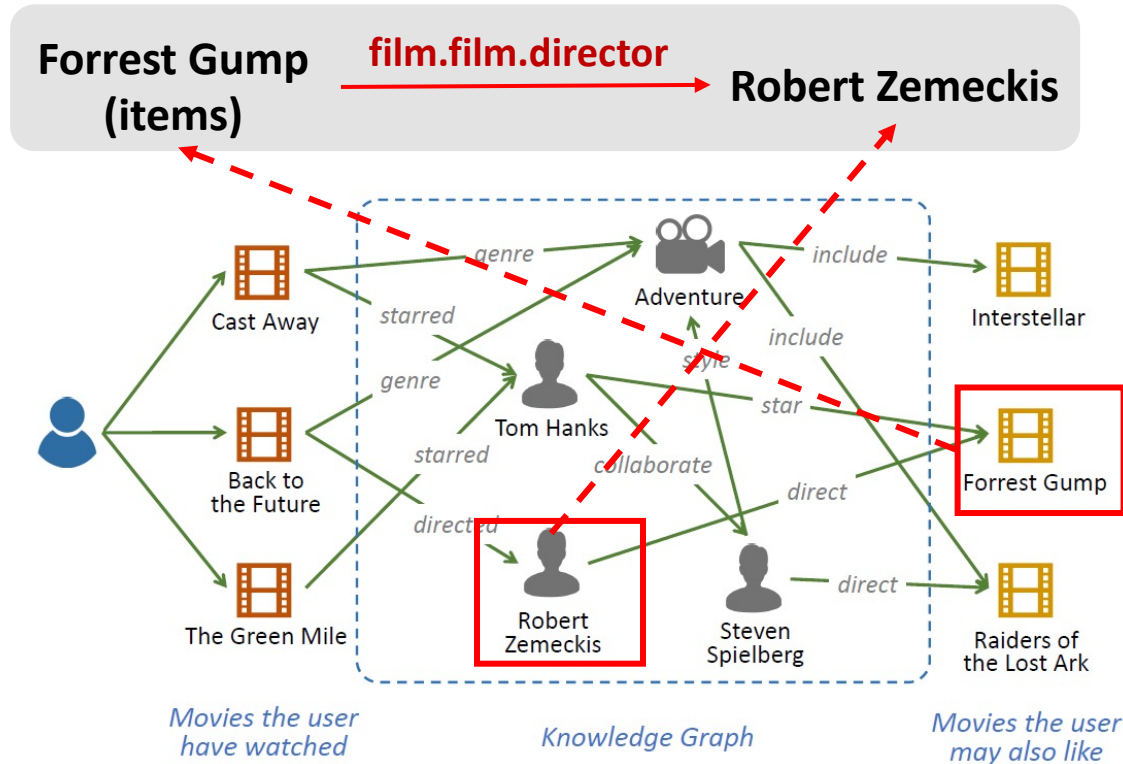
# KGCN (WWW'19)

## Side information about items: Knowledge Graph (KG)

### Heterogeneous Graph:

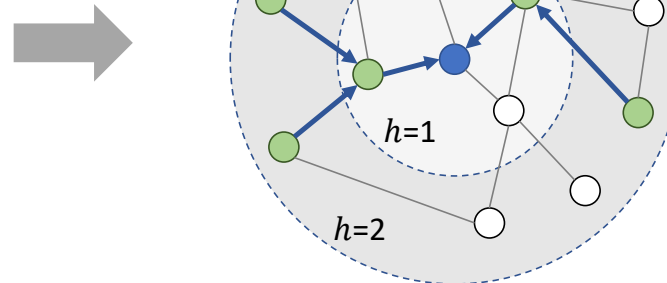
- Nodes: entities (Items)
- Edges: relations

Triples: (head, relation, tail)



$$\hat{y}_{uv} = f(\mathbf{u}, \mathbf{v}^u)$$

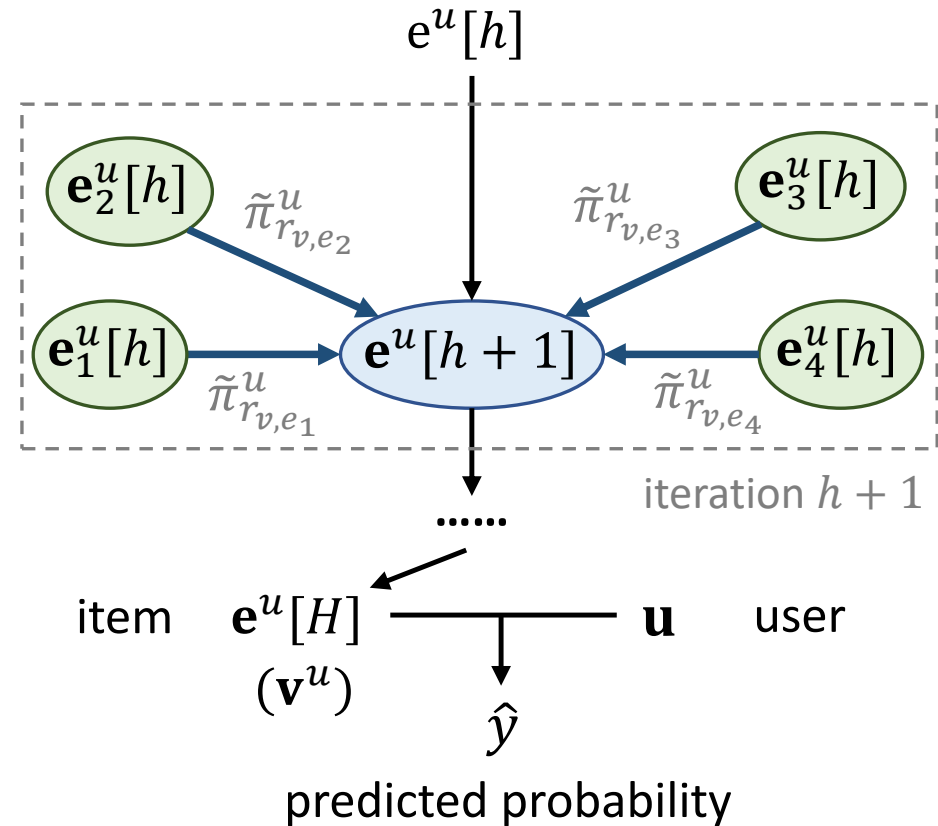
GNNs?



# KGCN (WWW'19)



- Representation Aggregation of neighboring entities



Transform a heterogeneous KG into a user-personalized weighted graph



# KGCN (WWW'19)

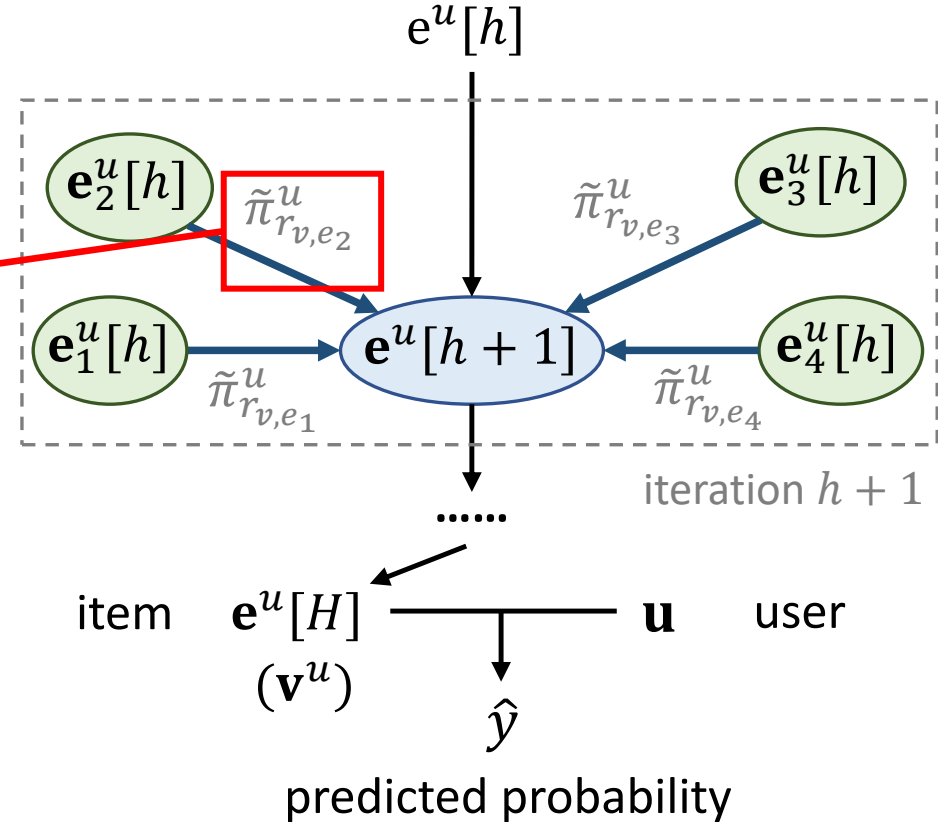


- Representation Aggregation of neighboring entities

$$\pi_r^u = g(\mathbf{u}, \mathbf{r}) \quad \text{user-specific relation (e.g., inner product)}$$

$$\tilde{\pi}_{r,v,e}^u = \frac{\exp(\pi_{r,v,e}^u)}{\sum_{e \in \mathcal{N}(v)} \exp(\pi_{r,v,e}^u)}$$

Normalized



Transform a heterogeneous KG into a user-personalized weighted graph

# KGCN (WWW'19)



- Representation Aggregation of neighboring entities

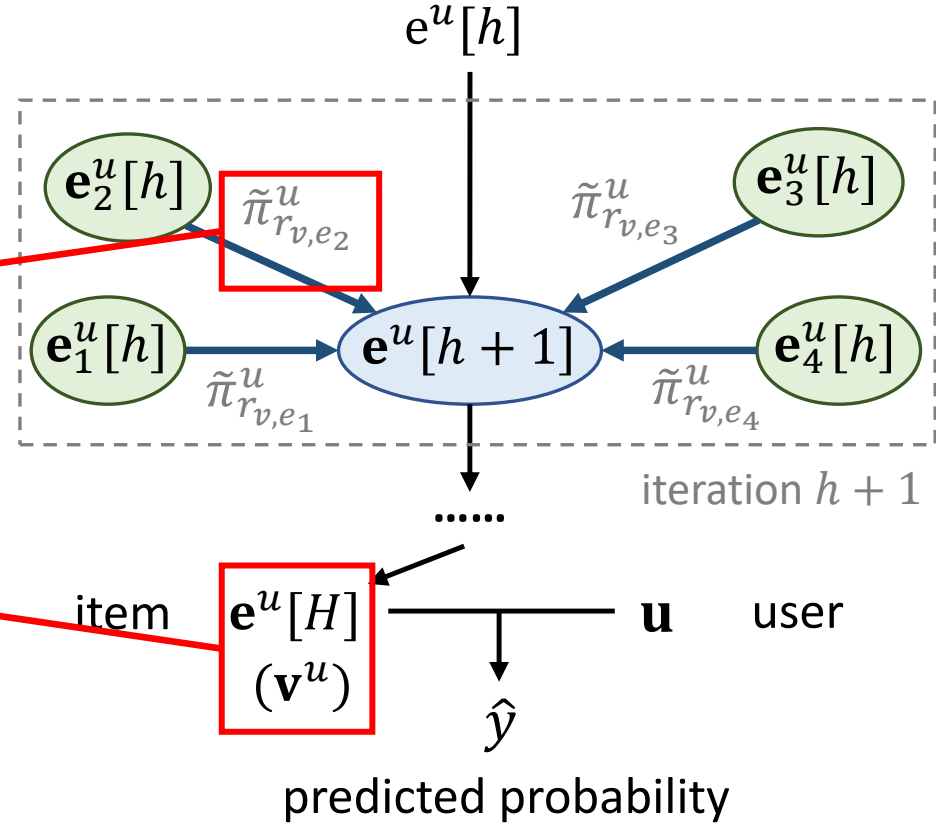
$$\pi_r^u = g(\mathbf{u}, \mathbf{r}) \quad \text{user-specific relation (e.g., inner product)}$$

$$\tilde{\pi}_{r,v,e}^u = \frac{\exp(\pi_{r,v,e}^u)}{\sum_{e \in \mathcal{N}(v)} \exp(\pi_{r,v,e}^u)}$$

Normalized

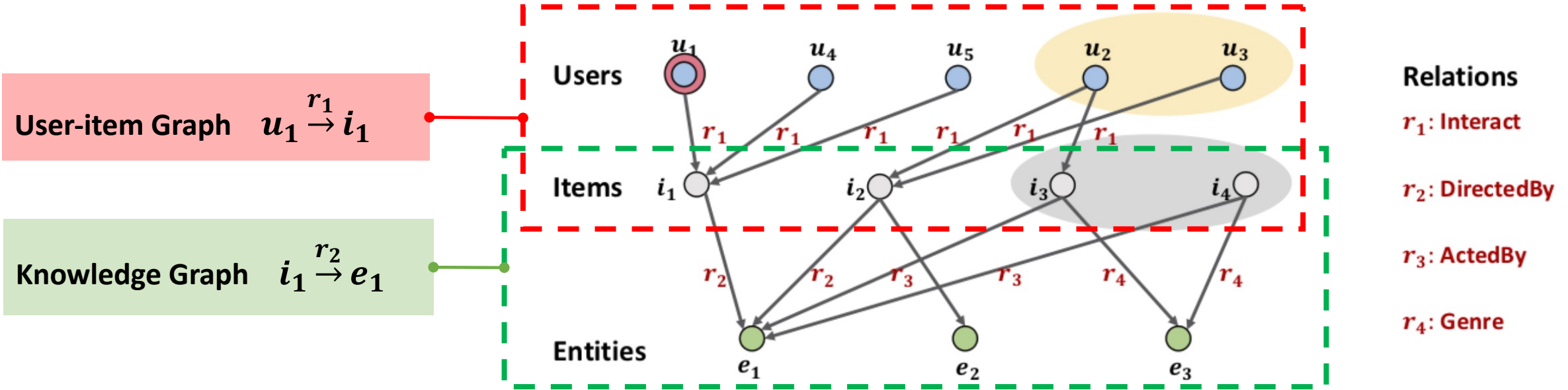
$$\mathbf{v}_{\mathcal{N}(v)}^u = \sum_{e \in \mathcal{N}(v)} \tilde{\pi}_{r,v,e}^u \mathbf{e}_e$$

$$\hat{y}_{uv} = f(\mathbf{u}, \mathbf{v}^u)$$

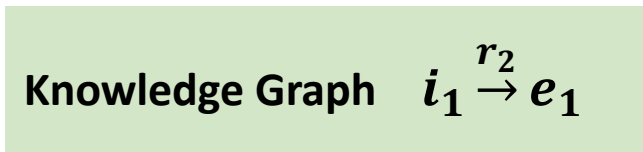


Transform a heterogeneous KG into a user-personalized weighted graph

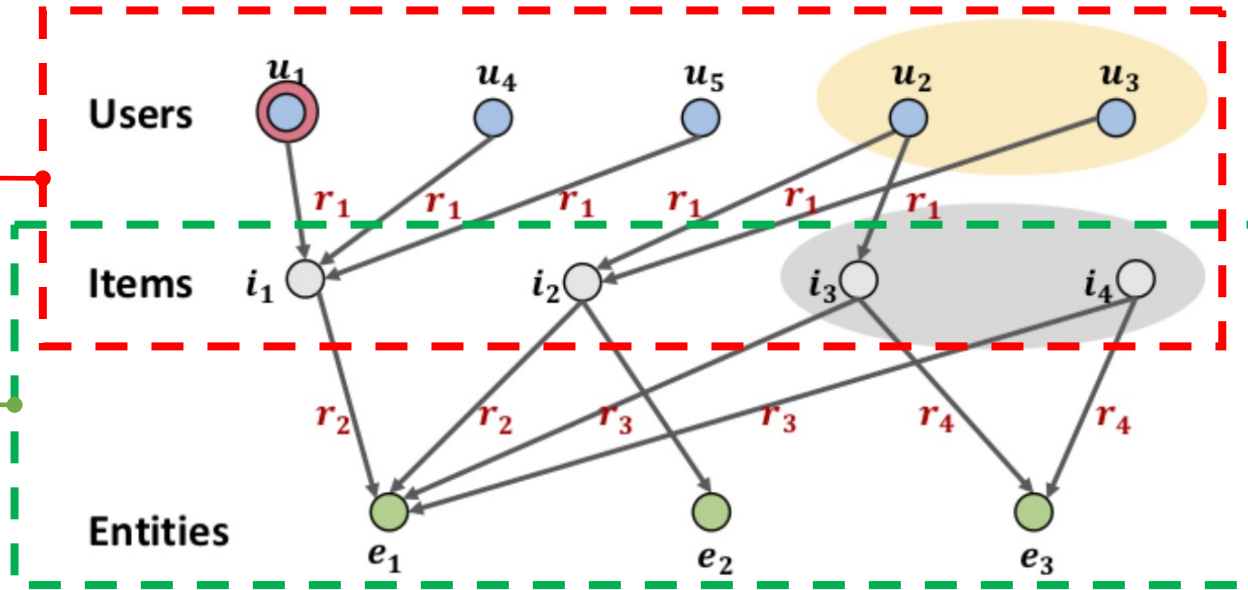
# KGAT



# KGAT



Collaborative Knowledge Graph (CKG)



Relations

$r_1$ : Interact

$r_2$ : DirectedBy

$r_3$ : ActedBy

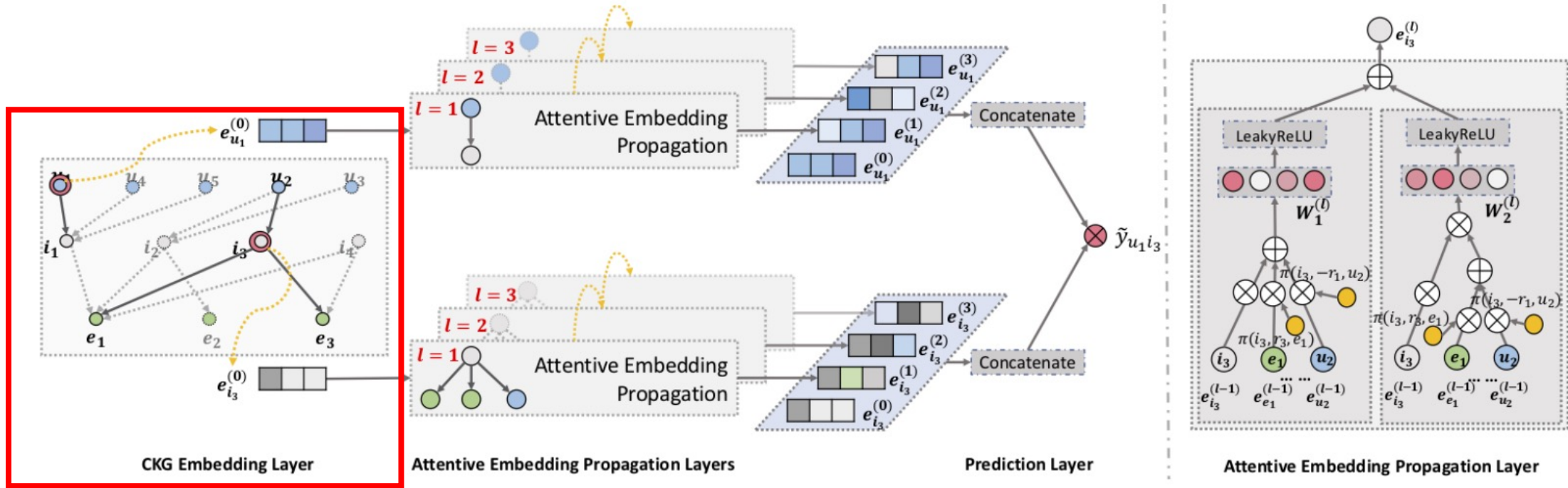
$r_4$ : Genre

To fully exploit high-order relations in CKG e.g., the long-range connectivities:

$$u_1 \xrightarrow{r_1} i_1 \xrightarrow{r_2} e_1 \xrightarrow{-r_2} i_2 \xrightarrow{-r_1} \{u_2, u_3\}$$

$$u_1 \xrightarrow{r_1} i_1 \xrightarrow{r_2} e_1 \xrightarrow{-r_3} \{i_3, i_4\}$$

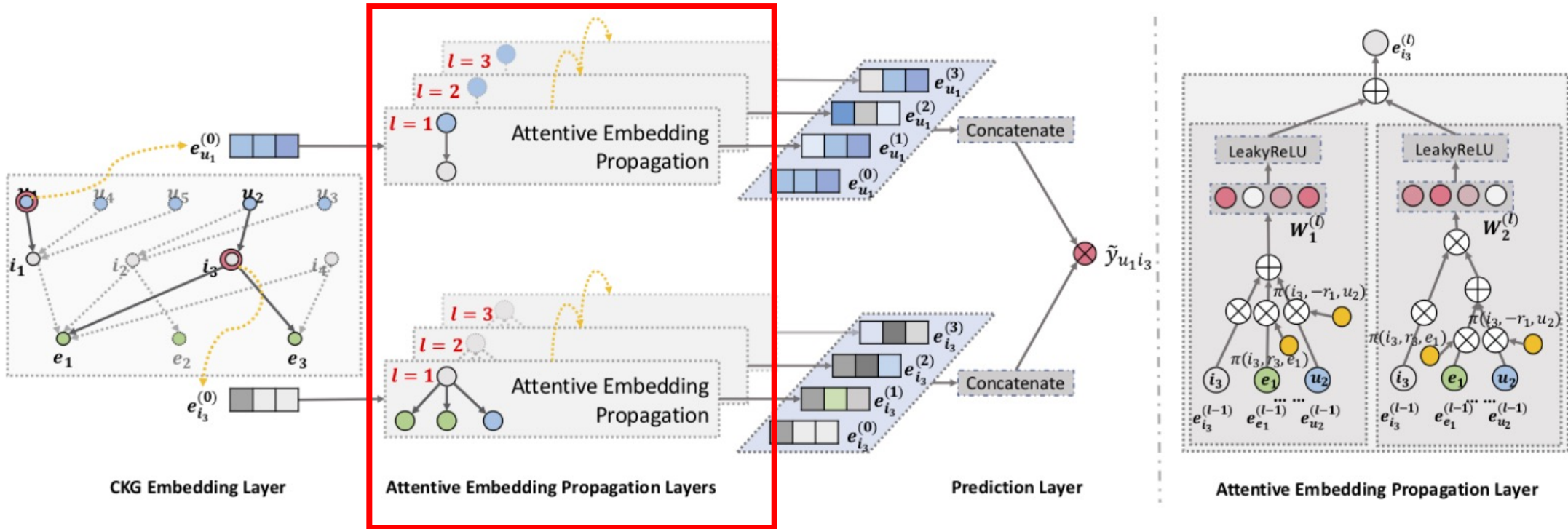
# KGAT



$$g(h, r, t) = \|\mathbf{W}_r \mathbf{e}_h + \mathbf{e}_r - \mathbf{W}_r \mathbf{e}_t\|_2^2$$

$$\mathcal{L}_{\text{KG}} = \sum_{(h, r, t, t') \in \mathcal{T}} -\ln \sigma(g(h, r, t') - g(h, r, t))$$

# KGAT

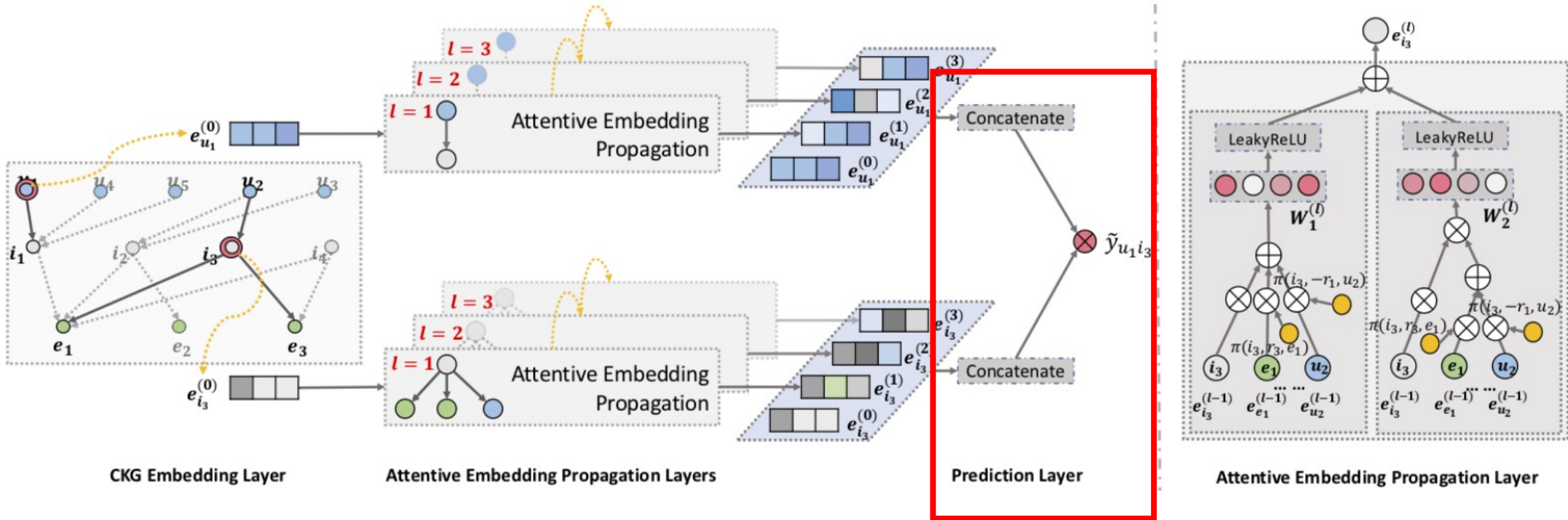


Information Propagation: 
$$\mathbf{e}_{N_h} = \sum_{(h,r,t) \in N_h} \pi(h,r,t) \mathbf{e}_t$$

Knowledge-aware Attention: 
$$\pi(h,r,t) = (\mathbf{W}_r \mathbf{e}_t)^\top \tanh(\mathbf{W}_r \mathbf{e}_h + \mathbf{e}_r)$$

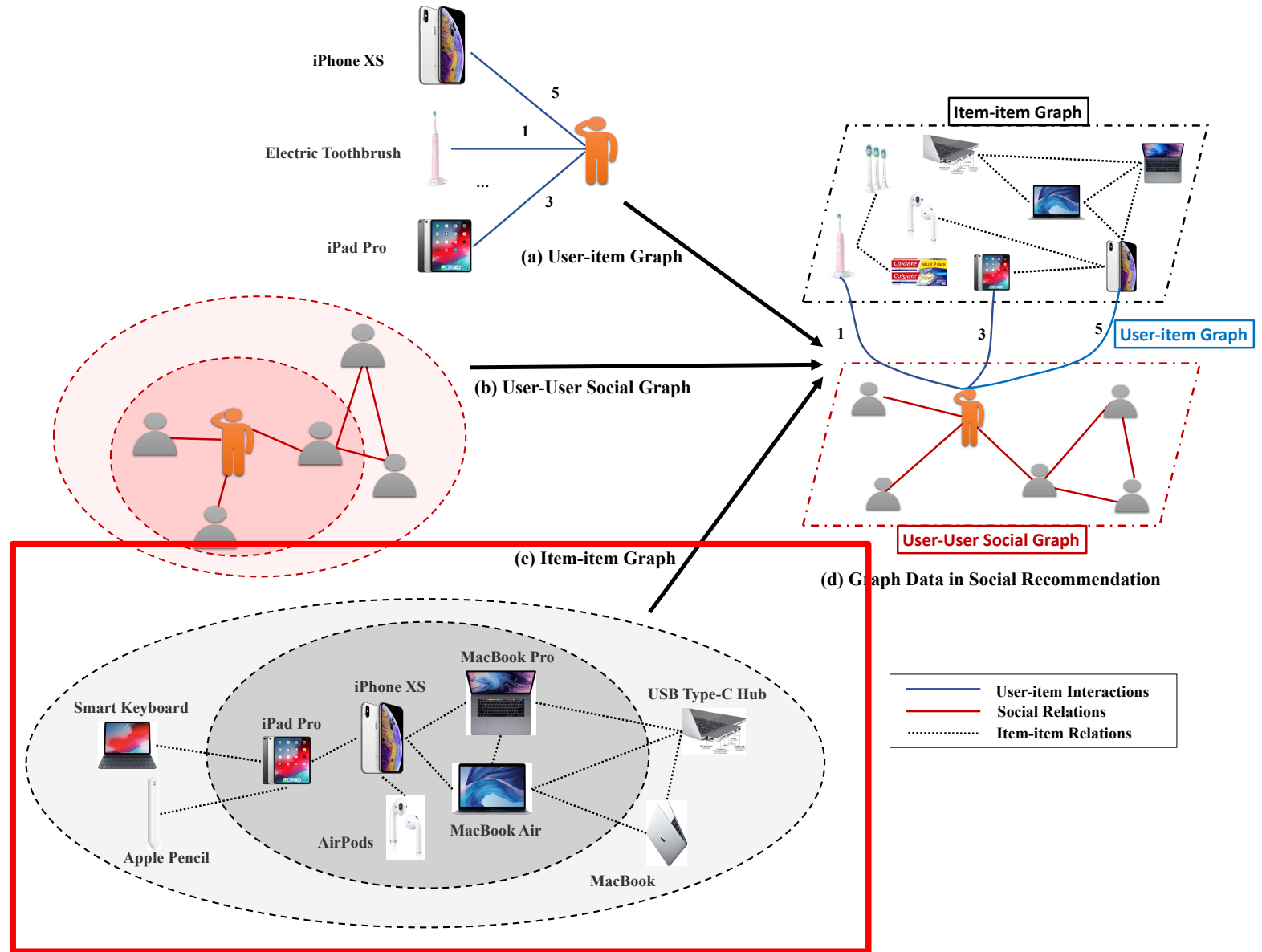
Information Aggregation: 
$$f_{\text{Bi-Interaction}} = \text{LeakyReLU}(\mathbf{W}_1(\mathbf{e}_h + \mathbf{e}_{N_h})) + \text{LeakyReLU}(\mathbf{W}_2(\mathbf{e}_h \odot \mathbf{e}_{N_h})),$$

# KGAT



$$\mathbf{e}_u^* = \mathbf{e}_u^{(0)} \parallel \dots \parallel \mathbf{e}_u^{(L)}, \quad \mathbf{e}_i^* = \mathbf{e}_i^{(0)} \parallel \dots \parallel \mathbf{e}_i^{(L)} \quad \hat{y}(u, i) = \mathbf{e}_u^{*T} \mathbf{e}_i^*$$

# GraphRec+





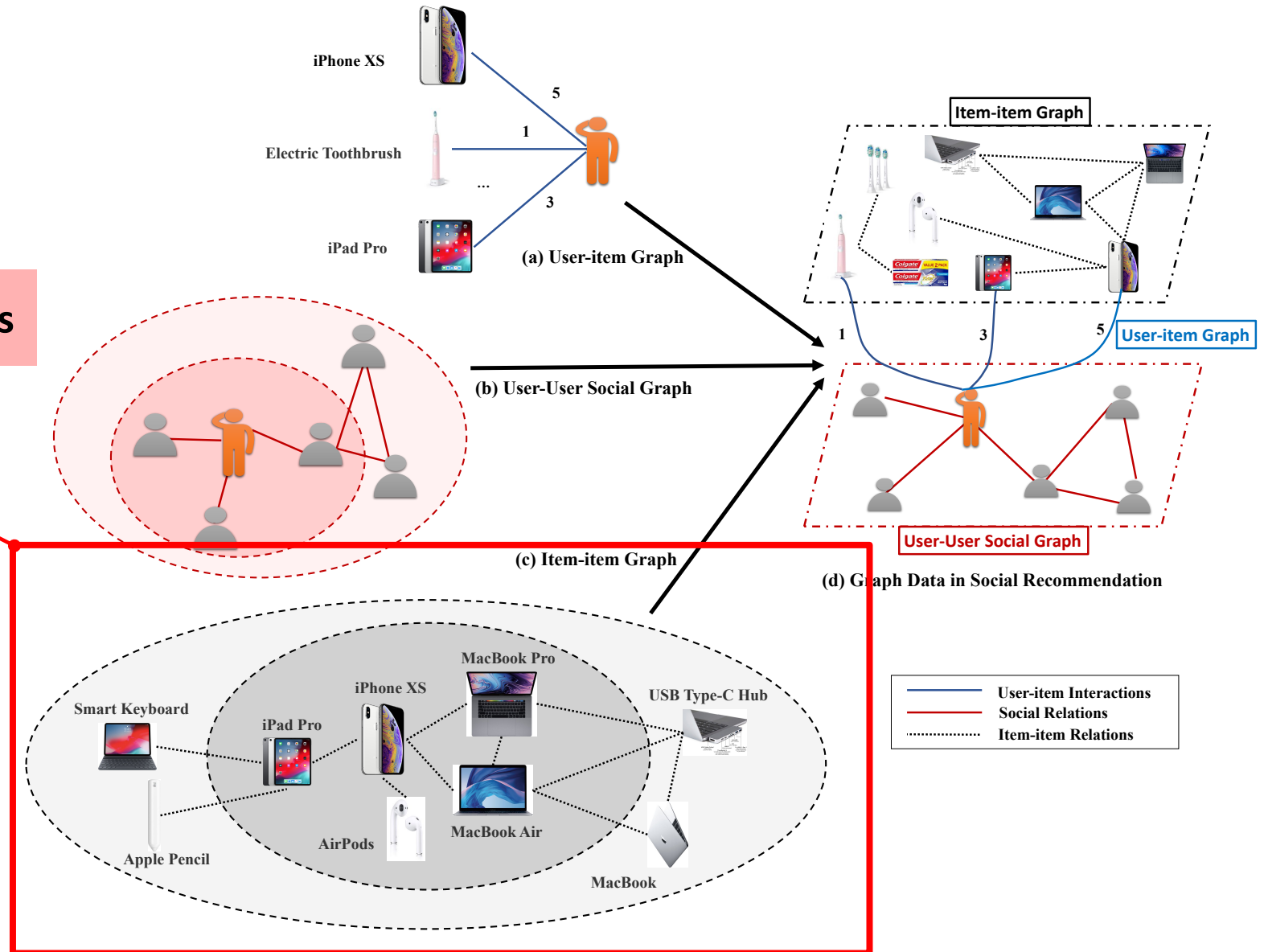
# GraphRec+

## Item-item Graph

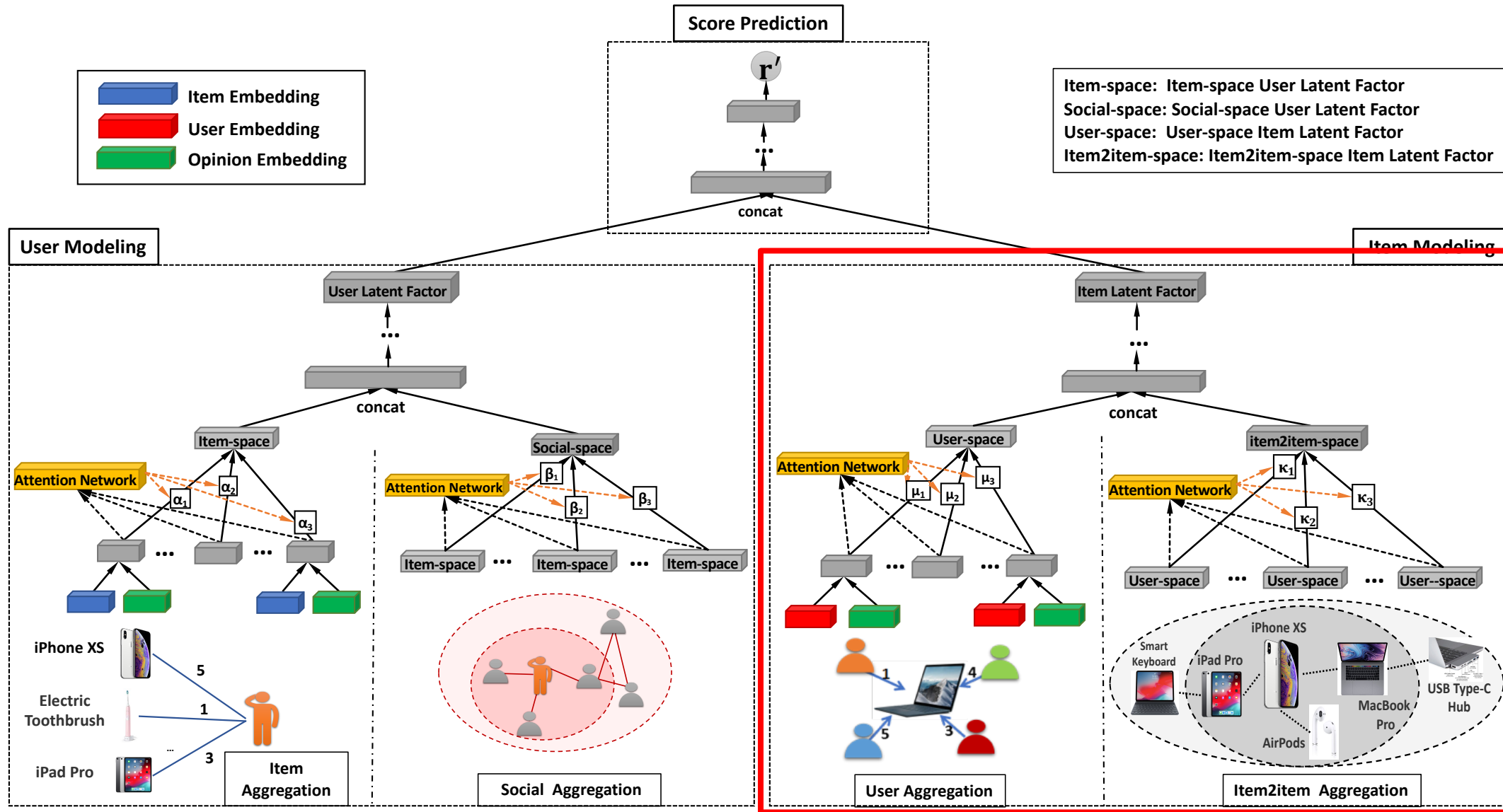
### Substitutable and Complementary Items

E.g.,

- 'users who bought A also bought B'
- 'users who viewed A also viewed B'



# GraphRec+



# Conclusion: Future Directions



## ● Depth

When the deeper GNNs can help in recommender systems?

# Conclusion: Future Directions



## ○ Depth

When the deeper GNNs can help in recommender systems?

## ○ Security (Data Poisoning Attack & Defense)

### ➤ Edge

user-item interactions

social relations

knowledge graph

### ➤ Node (users/items) Features

### ➤ Local Graph Structure

# Conclusion: Future Directions

## Depth

When the deeper GNNs can help in recommender systems?

## Security (Data Poisoning Attack & Defense)

### ➤ Edge

user-item interactions

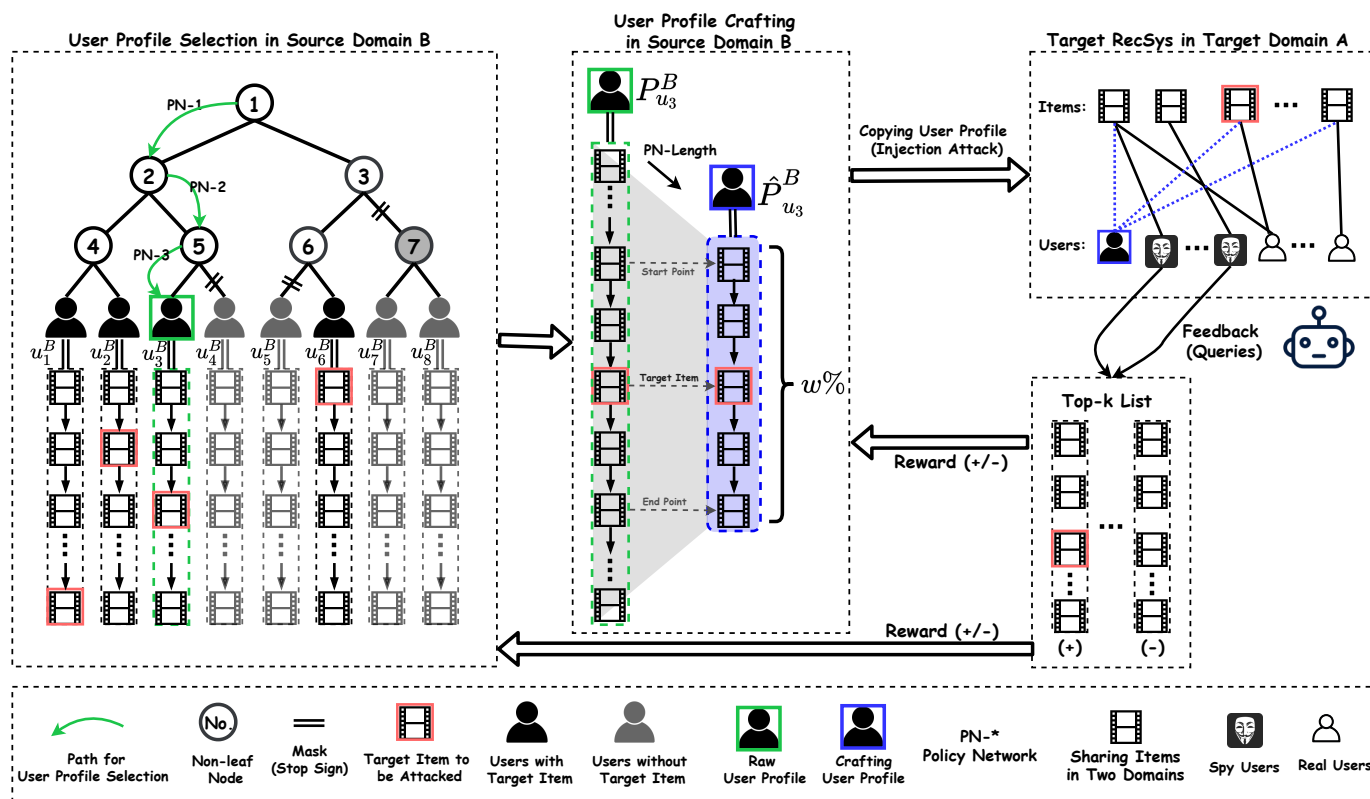
social relations

knowledge graph

### ➤ Node (users/items) Features

### ➤ Local Graph Structure

## CopyAttack



- **Wenqi Fan**, Computing, The Hong Kong Polytechnic University
- [wenqifan@polyu.edu.hk](mailto:wenqifan@polyu.edu.hk)
- <https://wenqifan03.github.io>



I am actively recruiting self-motivated Ph.D. students, Master, and Research Assistants. Visiting scholars and interns are also welcome. Send me an email with your CV if you are interested.