A Review-aware Graph Contrastive Learning Framework for Recommendation

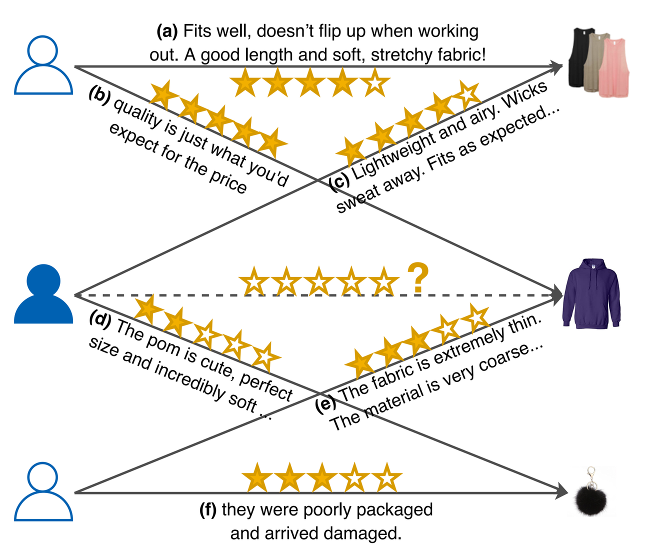
Jie Shuai[1], Kun Zhang[1], Le Wu[1,2,*], Peijie Sun[1], Richang Hong[1,2], Meng Wang[1,2], Yong Li[3]

[1] Key Laboratory of Knowledge Engineering with Big Data, Hefei University of Technology
[2] Institute of Artificial Intelligence, Hefei Comprehensive National Science Center
[3] Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University

SIGIR 2022 Madrid

Introduction

- Most online e-commerce platforms allow users to explain the reasons behind their ratings in text reviews.
- Reviews help alleviate the sparse issue in the recommender systems.
- Current review-based recommender systems can be mainly categorized into two types.
  - User and item representation learning enhanced by historical reviews
  - Interaction modeling enhanced by target reviews (Each target review is the review of a user to an item, which corresponds to the interaction)
- We argue that review-based recommendation data naturally forms a user-item bipartite graph with edge features.
  - Edge features
    - Numerical ratings
    - Textual reviews



- Challenges
  - How to explore this unique graph structure?
  - The rich graph patterns are hard to learn by sparse ratings. Can we develop self-supervised signals to boost graph learning?

The Problem Definition

- User-item bipartite graph with featured edges $\mathcal{G} =< \mathcal{U} \cup \mathcal{V}, \mathcal{E} >$
  - Nodes
    - a user set $\mathcal{U}$ ($|\mathcal{U}| = \text{M}$)
    - an item set $\mathcal{V}$ ($|\mathcal{V}| = \text{N}$)
  - Edges $\mathcal{E} = \{\boldsymbol{R}, \boldsymbol{E}\}$
    - A rating matrix $\boldsymbol{R} \in \mathcal{R}^{\text{M}\times\text{N}}$, each element r_{ij} represents the rating.
    - A review tensor $\boldsymbol{E} \in \mathbb{R}^{\text{M}\times\text{N}\times d}$, each vector $\boldsymbol{e}_{ij} \in \mathbb{R}^d$ represent the review feature.
- The goal is to the predict final rating matrix $\hat{\boldsymbol{R}} \in \mathcal{R}^{\text{M}\times\text{N}}$ with the graph $\mathcal{G}$.

The Proposed Model: RGCL

- RGCL-1: graph initialization
  - Nodes
    - Users and items are represented by free embeddings.
  - Edges
    - Ratings are regarded as edge labels.
    - Reviews are pre-encoded to feature vectors.
- RGCL-2: Review-aware Graph learning (RG)
  - User and item representation learning
    - Review-aware message passing

Influence from edges / Influence from neighbors

$$\boldsymbol{x}_{r;j\to i}^{(l)} = \frac{\overbrace{\underbrace{\sigma(\boldsymbol{w}_{r,1}^{(l)\top}\boldsymbol{e}_{ij})}\boldsymbol{W}_{r,1}^{(l)}\boldsymbol{e}_{ij}} + \overbrace{\underbrace{\sigma(\boldsymbol{w}_{r,2}^{(l)\top}\boldsymbol{e}_{ij})}\boldsymbol{W}_{r,2}^{(l)}\boldsymbol{v}_j^{(l-1)}}}{\sqrt{|\mathcal{N}_j||\mathcal{N}_i|}},$$

To tune the impacts by edges

    - Aggregation

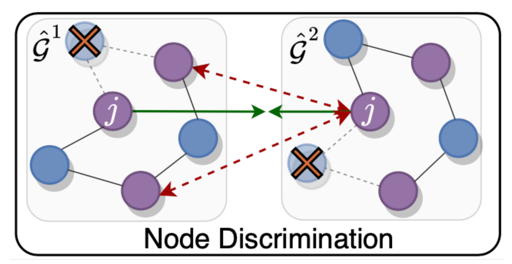
$$\boldsymbol{u}_i^{(l)} = \boldsymbol{W}^{(l)}\sum_{r\in\mathcal{R}}\sum_{k\in\mathcal{N}_{i,r}}\boldsymbol{x}_{r;k\to i}^{(l)}, \quad \boldsymbol{v}_j^{(l)} = \boldsymbol{W}^{(l)}\sum_{r\in\mathcal{R}}\sum_{k\in\mathcal{N}_{j,r}}\boldsymbol{x}_{r;k\to j}^{(l)}$$

$$\hat{\boldsymbol{u}}_i = \boldsymbol{u}_i^{(L)}, \quad \hat{\boldsymbol{v}}_j = \boldsymbol{v}_j^{(L)},$$

  - Interaction Modeling

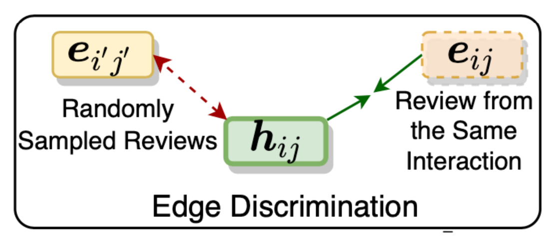
$$\hat{r}_{ij} = \boldsymbol{w}^\top\boldsymbol{h}_{ij}, \qquad \boldsymbol{h}_{ij} = \text{MLP}\left([\hat{\boldsymbol{u}}_i, \hat{\boldsymbol{v}}_j]\right),$$

- RGCL-3: graph contrastive learning
  - Node Discrimination (ND) for Node Representation Enhancement
    - Node dropping for data augmentation



$$\mathcal{L}_{nd}^{user} = \underbrace{-\mathbb{E}_{\mathcal{U}}\left[\log\left(F\left(\hat{\boldsymbol{u}}_i^1, \hat{\boldsymbol{u}}_i^2\right)\right)\right]}_{\text{Positive pairs}} + \underbrace{\mathbb{E}_{\mathcal{U}\times\mathcal{U}'}\left[\log\left(F\left(\hat{\boldsymbol{u}}_i^1, \hat{\boldsymbol{u}}_{i'}^2\right)\right)\right]}_{\text{Negative pairs}}$$

  - Edge Discrimination (ED) for target review alignment

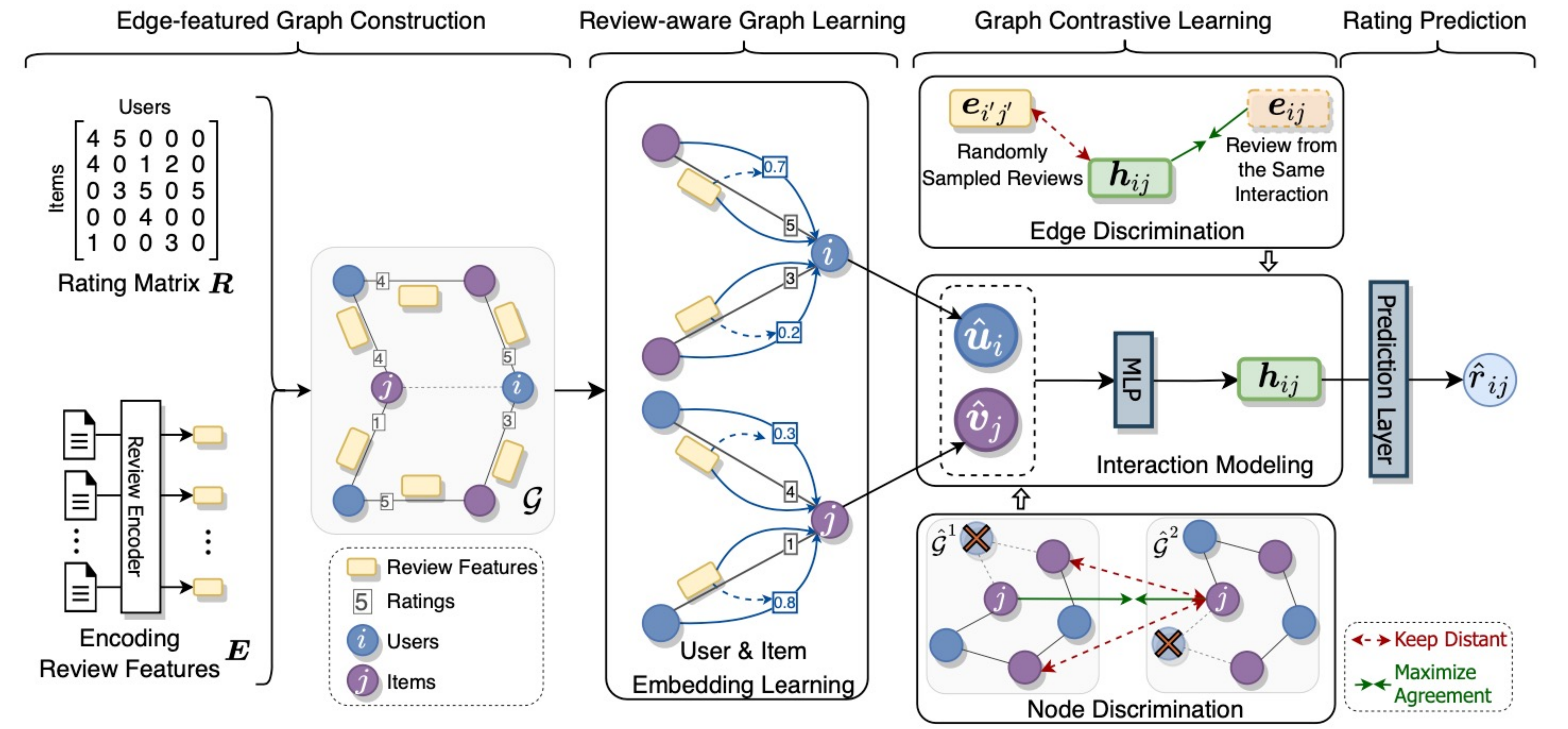


$$\mathcal{L}_{ed} = \underbrace{-\mathbb{E}_{\mathcal{E}}\left[\log\left(F\left(\boldsymbol{h}_{ij}, \boldsymbol{e}_{ij}\right)\right)\right]}_{\text{Positive pairs}} + \underbrace{\mathbb{E}_{\mathcal{E}\times\mathcal{E}'}\left[\log\left(F\left(\boldsymbol{h}_{ij}, \boldsymbol{e}_{i'j'}\right)\right)\right]}_{\text{Negative pairs}}$$

- Model optimization

$$\mathcal{L} = \frac{1}{|\mathcal{S}|}\sum_{(i,j)\in\mathcal{S}}(\hat{r}_{ij} - r_{ij})^2 + \alpha\mathcal{L}_{ed} + \beta\mathcal{L}_{nd}$$

The Framwork of Our Proposed RGCL



Experiments

Table 1: Statistics of datasets

Datasets	#Users	#Items	#Reviews	Density
Digital_Music	5,541	3,568	64,706	0.330%
Toys_and_Games	19,412	11,924	167,597	0.072%
Clothing	39,387	23,033	278,677	0.031%
CDs_and_Vinly	75,258	64,443	1,097,592	0.023%
Yelp	8,423	3,742	88,647	0.281%

Table 2: Overall comparison in terms of the MSE

Method	Digital_Music	Toys_and_Games	Clothing	CDs_and_Vinly	Yelp
(1) SVD	0.8523±4e-4	0.8086±1e-3	1.1167±1e-3	0.8662±2e-4	1.1939±1e-3
(2) NCF	0.8403±5e-3	0.8078±2e-3	1.1094±1e-3	0.8781±1e-3	1.1896±4e-3
(3) DeepCoNN	0.8378±1e-3	0.8028±7e-4	1.1184±2e-3	0.8621±1e-3	1.1877±1e-3
(4) NARRE	0.8172±1e-3	0.7962±1e-3	1.1064±1e-3	0.8495±1e-3	1.1862±1e-3
(5) DAML	0.8237±2e-3	0.7936±4e-3	1.1065±2e-3	0.8483±1e-3	1.1793±1e-3
(6) SDNet	0.8331±3e-3	0.8006±1e-3	1.1080±1e-3	0.8654±5e-4	1.1837±3e-3
(7) TransNets	0.8273±5e-3	0.7980±1e-2	1.1141±5e-3	0.8440±1e-3	1.1855±2e-3
(8) GC-MC	0.8090±1e-3	0.7986±5e-4	1.1088±1e-3	0.8404±1e-3	1.1737±1e-3
(9) RMG	0.8074±1e-3	0.7901±1e-3	1.1064±2e-3	0.8425±8e-4	1.1705±1e-3
(10) SSG	0.8218±2e-3	0.8064±1e-3	1.1228±1e-3	0.8458±1e-3	1.1807±1e-3
(11) RG	0.8037±2e-3 (0.5%)	0.7853±8e-4 (0.6%)	1.1024±9e-4 (0.4%)	0.8360±1e-3 (0.5%)	1.1692±2e-3 (0.1%)
(12) RG+ND	0.7780±2e-3 (3.6%)	0.7831±1e-3 (0.9%)	1.0925±3e-4 (1.3%)	0.8240±6e-4 (2.0%)	1.1625±1e-3 (0.7%)
(13) RG+ED	0.7810±3e-3 (3.3%)	0.7797±9e-4 (1.3%)	1.0891±8e-4 (1.6%)	0.8244±1e-3 (1.9%)	1.1636±1e-3 (0.6%)
(14) RGCL	**0.7735±4e-3 (4.2%)**	**0.7771±1e-4 (1.6%)**	**1.0858±1e-3 (1.9%)**	**0.8180±7e-4 (2.7%)**	**1.1609±8e-4 (0.8%)**

Ablation study 1

Table 3: MSE comparison of different components of RG

Models	Digital_Music	Toys_and_Games	Clothing
GC-MC	0.8145±1e-3	0.8034±5e-4	1.1088±1e-3
RGAT	0.8248±2e-3	0.8093±3e-3	1.1158±2e-3
RG (w/o review)	0.8077±1e-3	0.7927±1e-3	1.1038±1e-3
RG (w/o weight)	0.8074±2e-3	0.7901±1e-3	1.1064±1e-3
RG	0.8037±2e-3	0.7853±8e-4	1.1024±9e-4

Ablation study 2

Table 5: MSE of our CL constraints on different recommendation models.

Models	Digital_Music	Toys_and_Games	Clothing
NARRE	0.8172±1e-3	0.7962±1e-3	1.1064±1e-3
+ED	0.8018±4e-3 (1.9%)	0.7910±1e-3 (0.7%)	1.0961±1e-3 (0.9%)
GC-MC	0.8090±1e-3	0.7986±5e-4	1.1088±1e-3
+ED	0.7822±2e-3 (3.3%)	0.7848±2e-3 (1.7%)	1.0922±2e-3 (1.5%)
+ND+ED	0.7803±1e-3 (3.5%)	0.7825±1e-3 (2.0%)	1.0901±6e-4 (1.7%)

Conclusion

- We borrowed the natural user-item graph structure for review-based recommendation and unified review and graph in one model.
- We developed self-supervised signals to boost the embedding learning and interaction modeling based on reviews.
- Extensive experiments demonstrated the effectiveness of our proposed RGCL.

Code   Weixin