Personalized News Recommendation with Knowledge-aware Interactive Matching

Tao Qi[1*], Fangzhao Wu[2], Chuhan Wu[1], Yongfeng Huang[1]

[1]Department of Electronic Engineering & BNRist, Tsinghua University, Beijing 100084, China

[2]Microsoft Research Asia, Beijing 100080, China

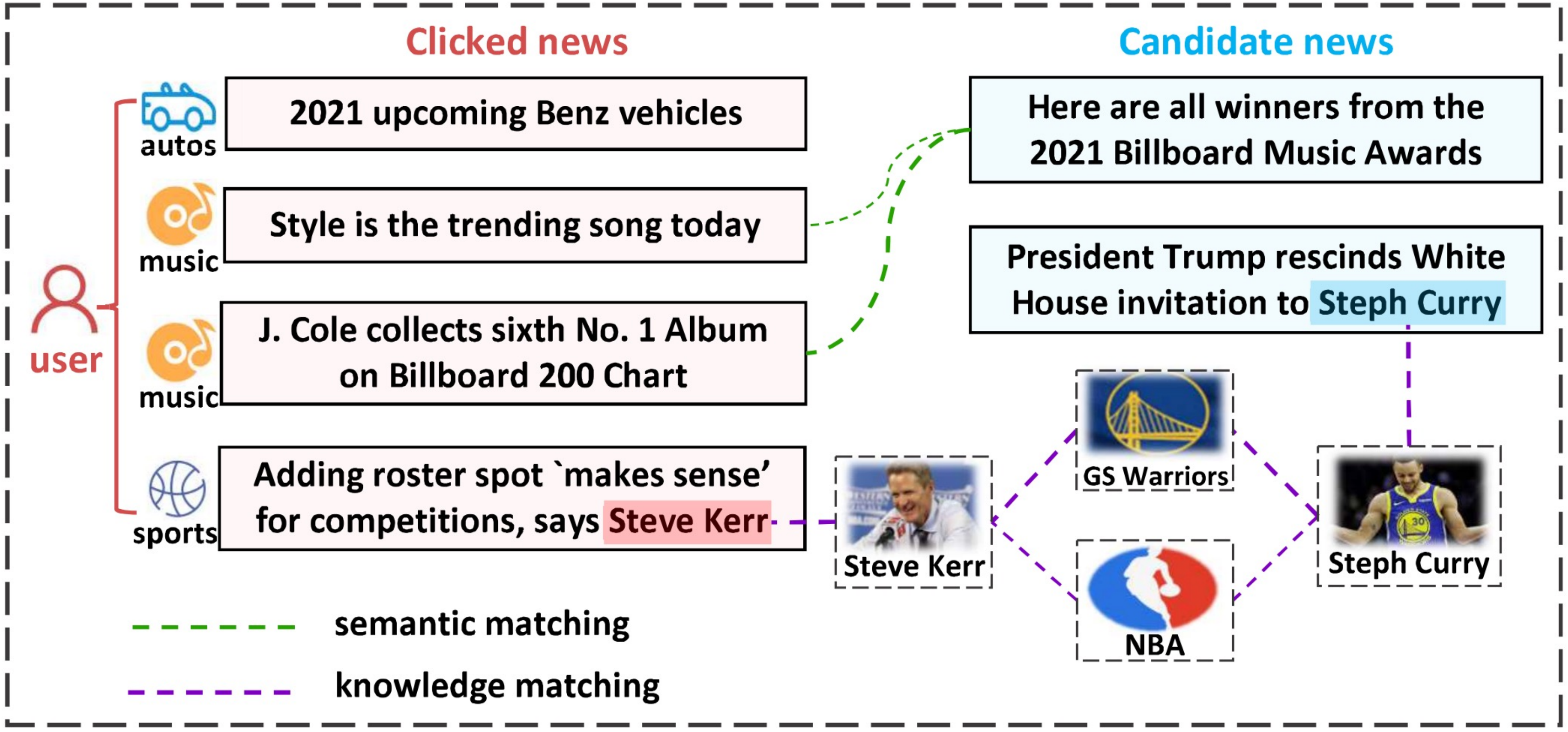
*taoqi.qt@gmail.com

Personalized News Recommendation

- Important for improving user experience on online news platforms
- Accurate matching of user interest and candidate news is critical for personalized news recommendation

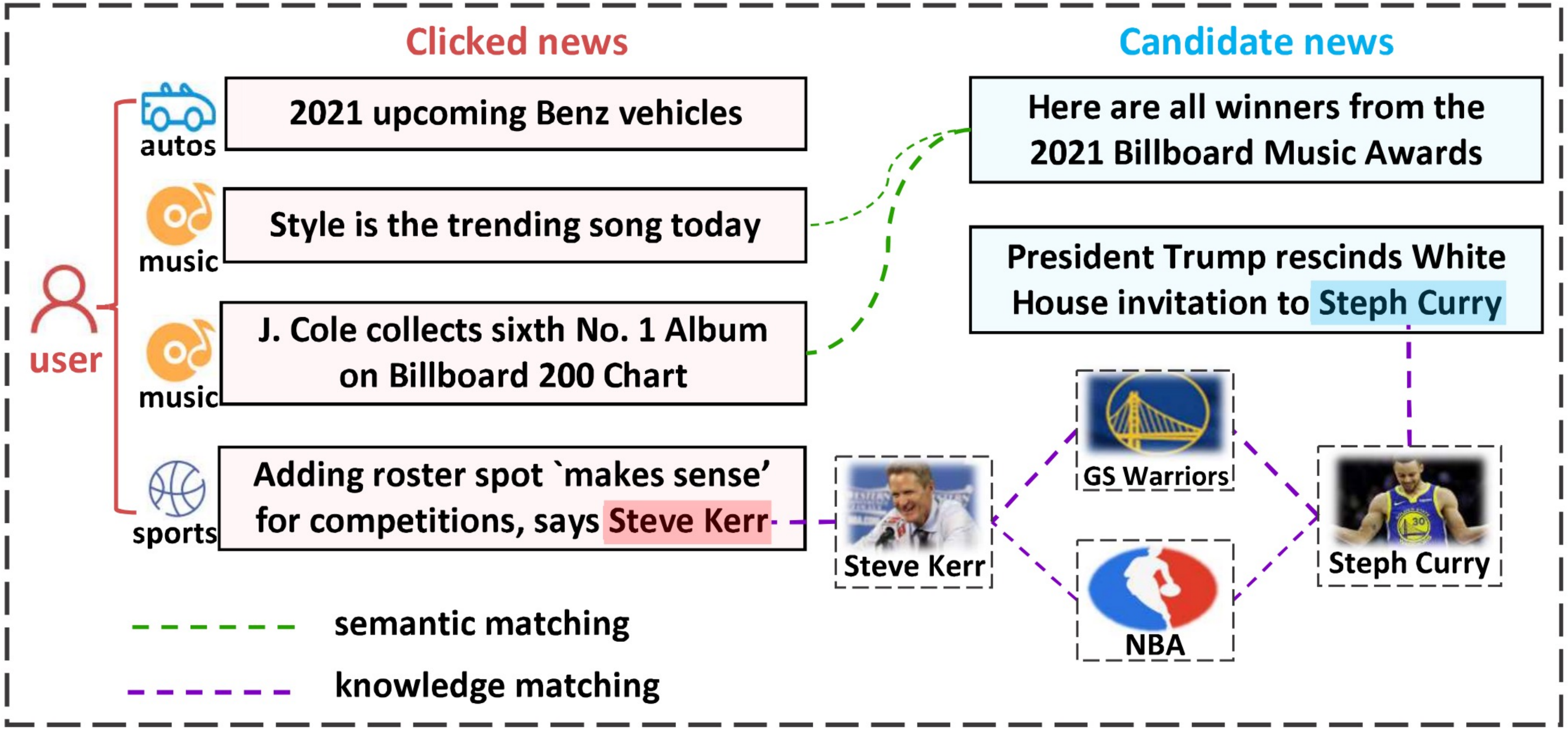
Motivation

- User interests are usually diverse
- Candidate news may cover multiple entities and aspects
- Independent modeling of them is sub-optimal for interest matching



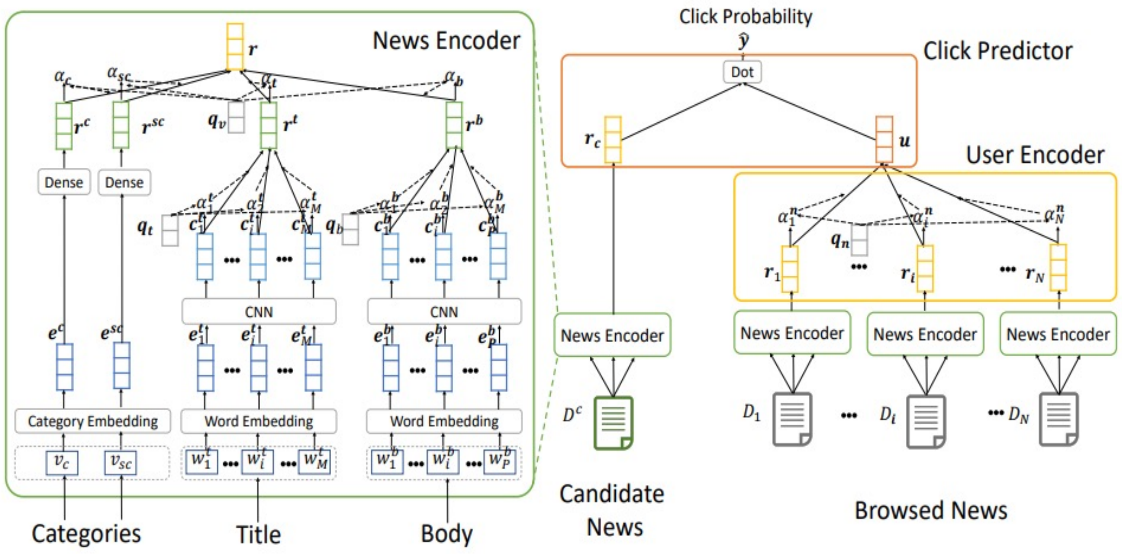
Motivation

- Matching between clicked news and candidate news in both knowledge and semantic level is useful for interest matching

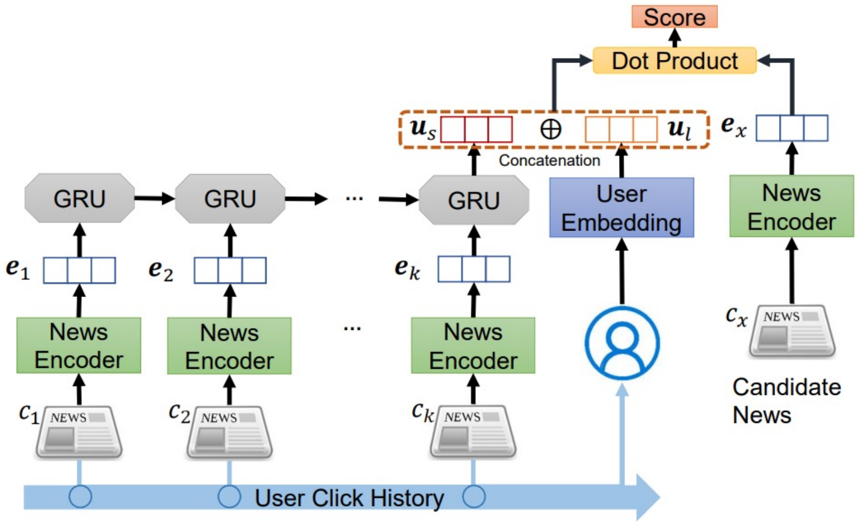


Related Work

- Personalized news recommendation
  - Model candidate news and user interests in an independent way
  - E.g., NAML[1], NRMS[2], LSTUR[3], KRED[4]



NAML

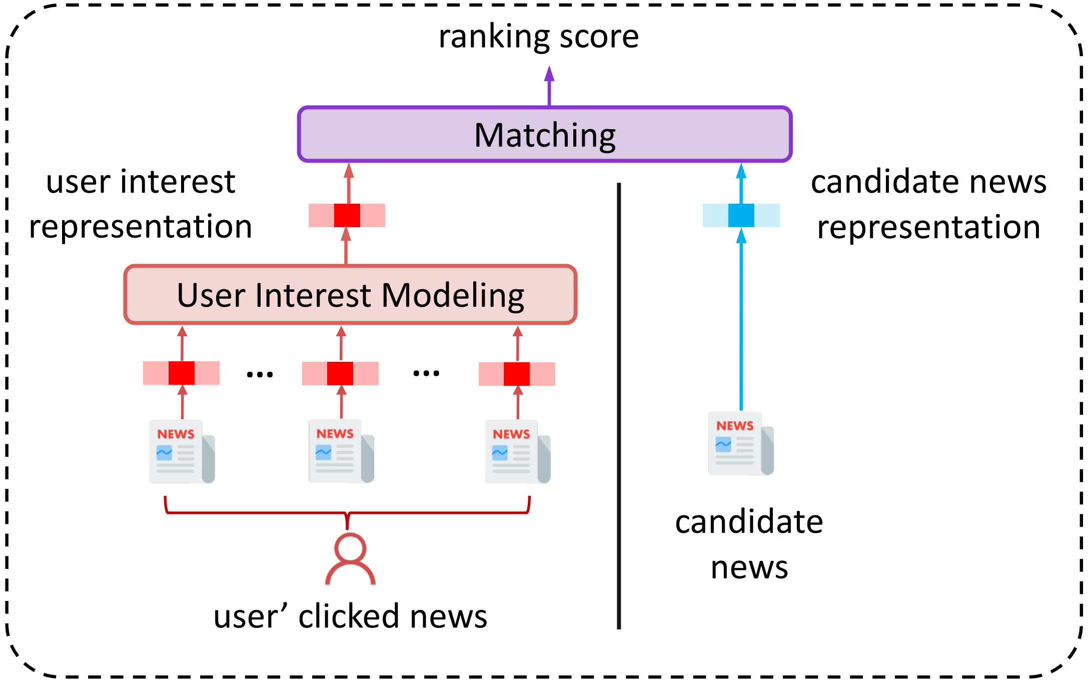


LSTUR

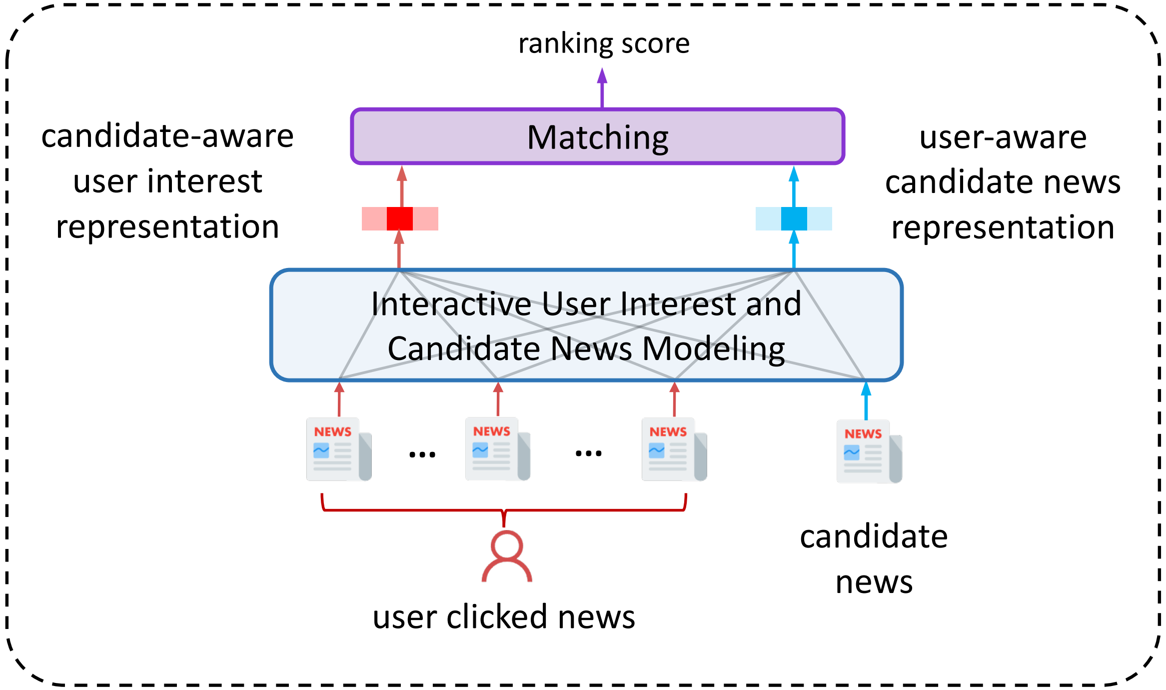
- Challenges:
  - Independent modeling of user interests and candidate news is sub-optimal for interest matching

Knowledge-aware Interactive Matching

- News recommendation with knowledge-aware interactive matching (KIM)



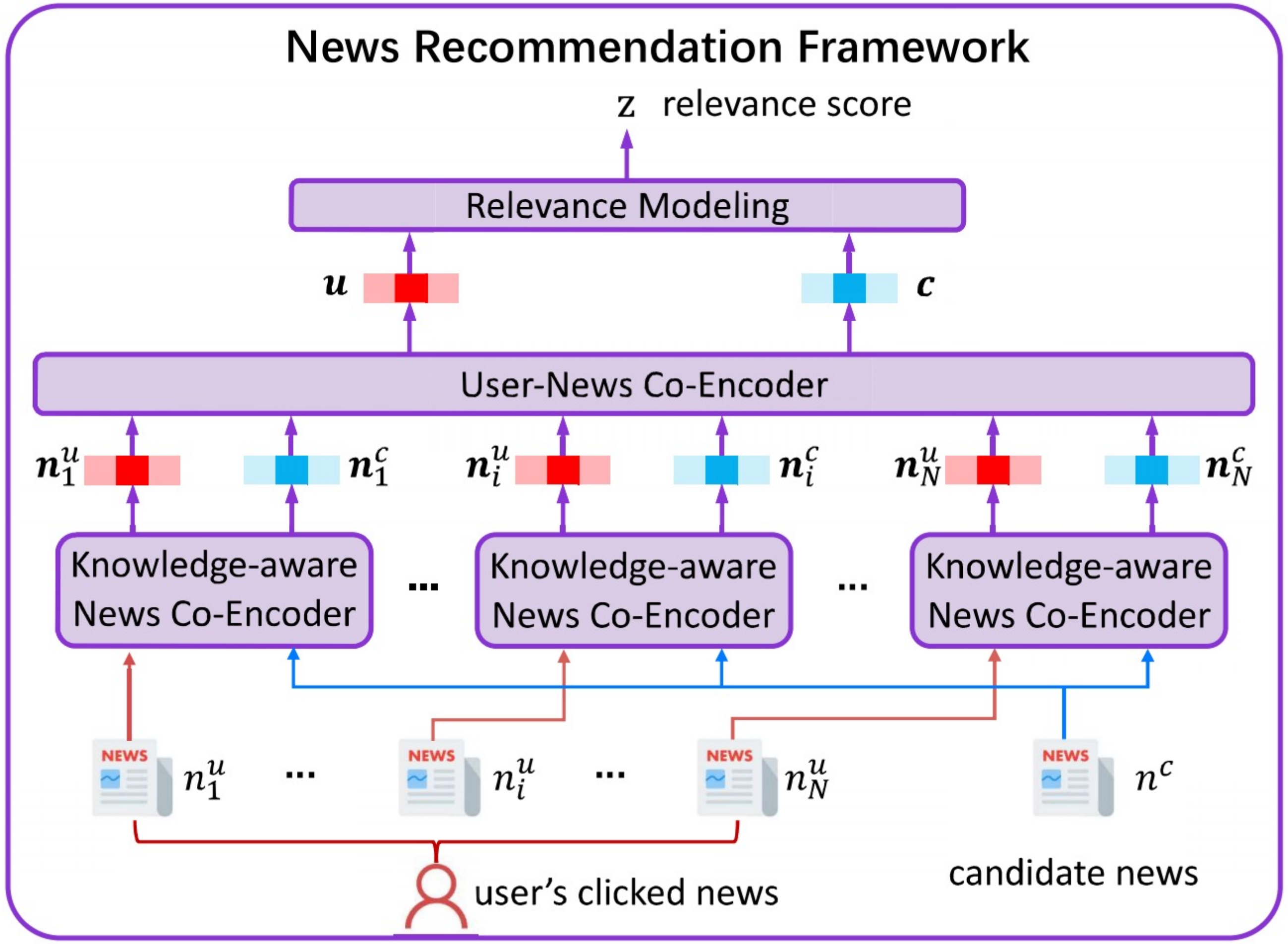
Mainstream Matching Framework



Our Work: Interactive Interest Matching

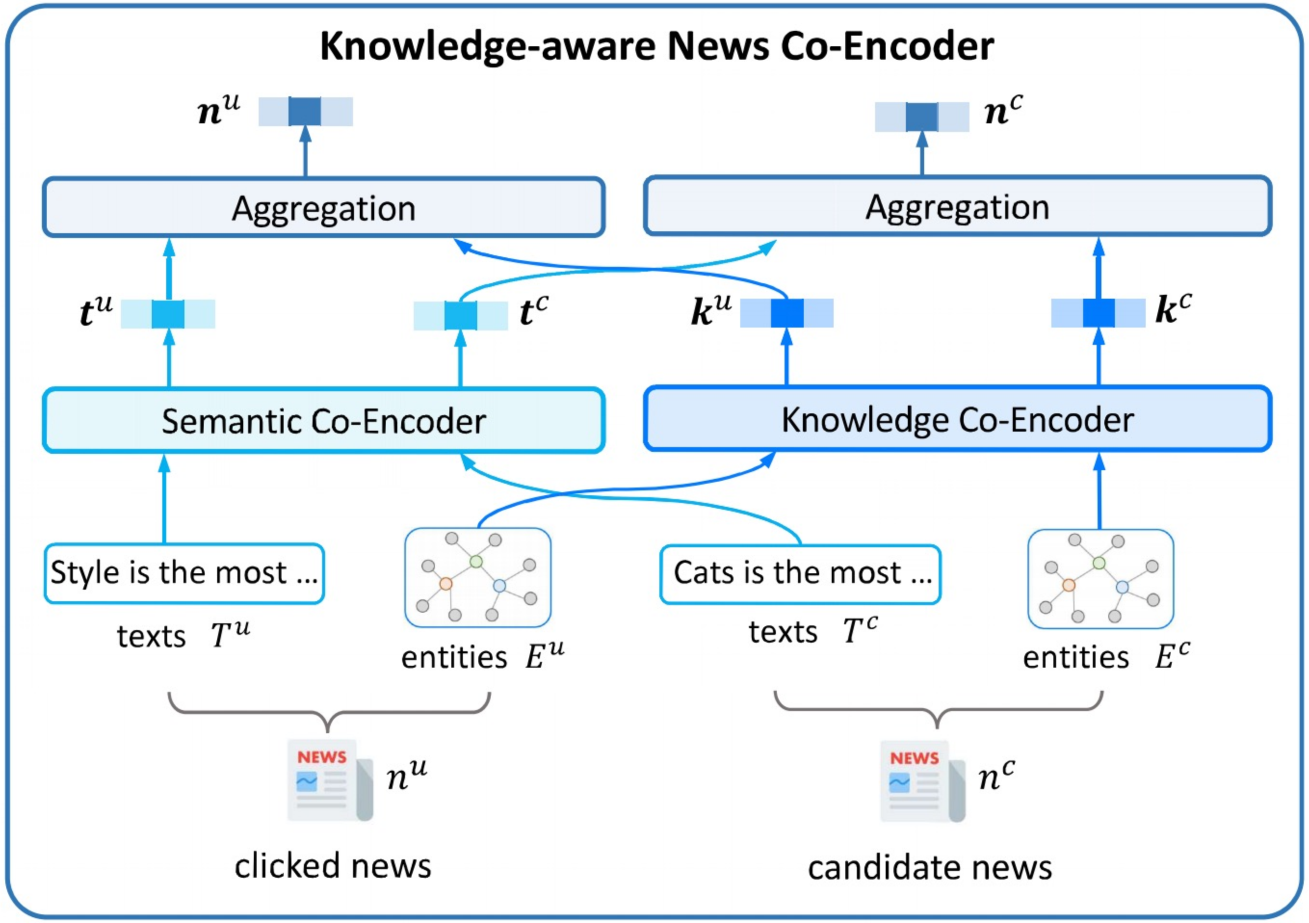
KIM

- Framework Overview
  - Knowledge-aware news co-encoder
  - User-news co-encoder
  - Relevance modeling



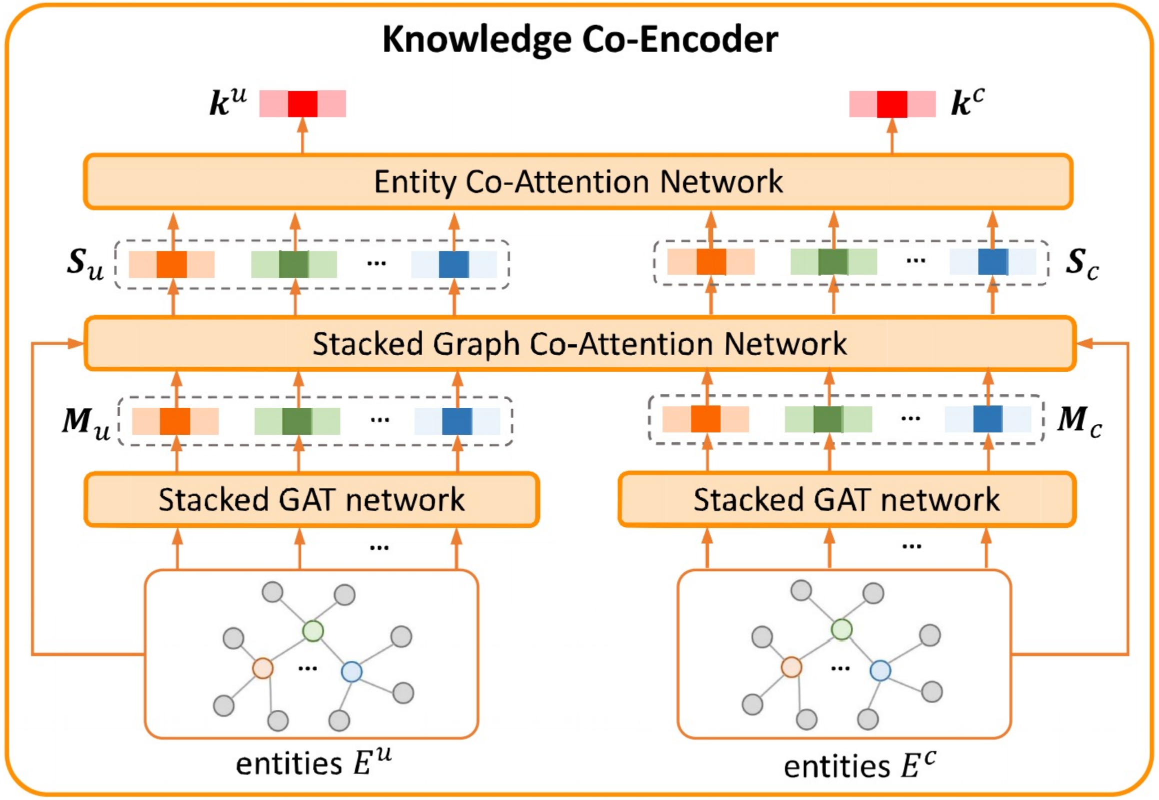
KIM

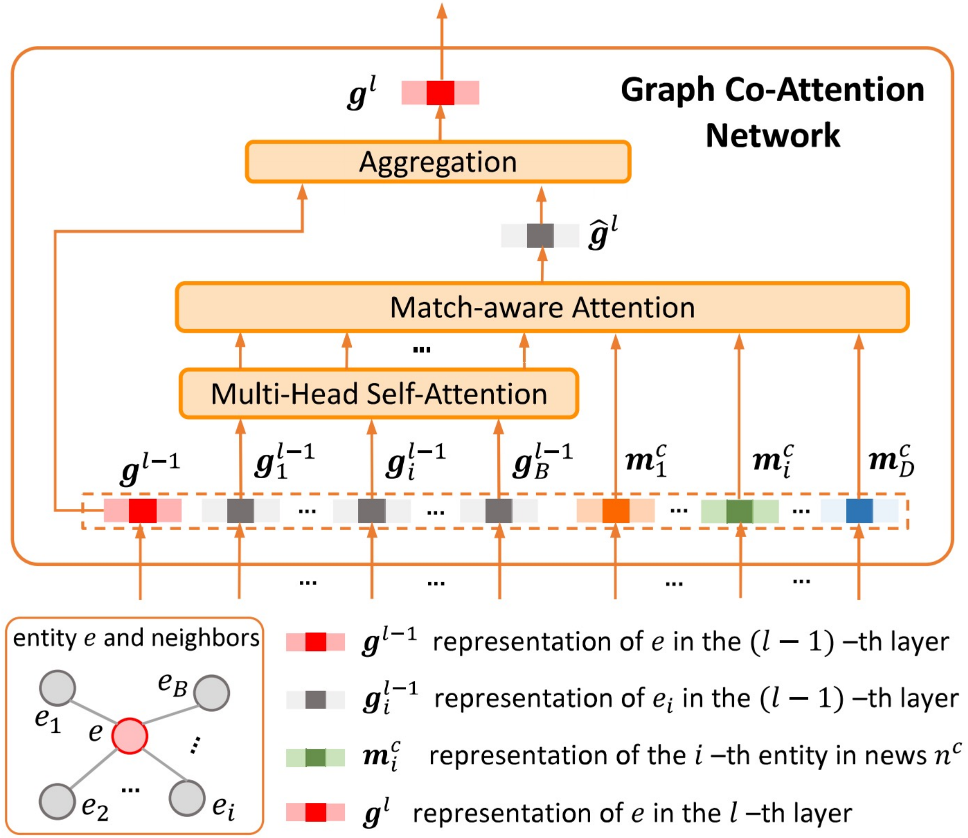
- Knowledge-aware news co-encoder
  - Knowledge co-encoder
  - Semantic co-Encoder
  - Aggregation



KIM

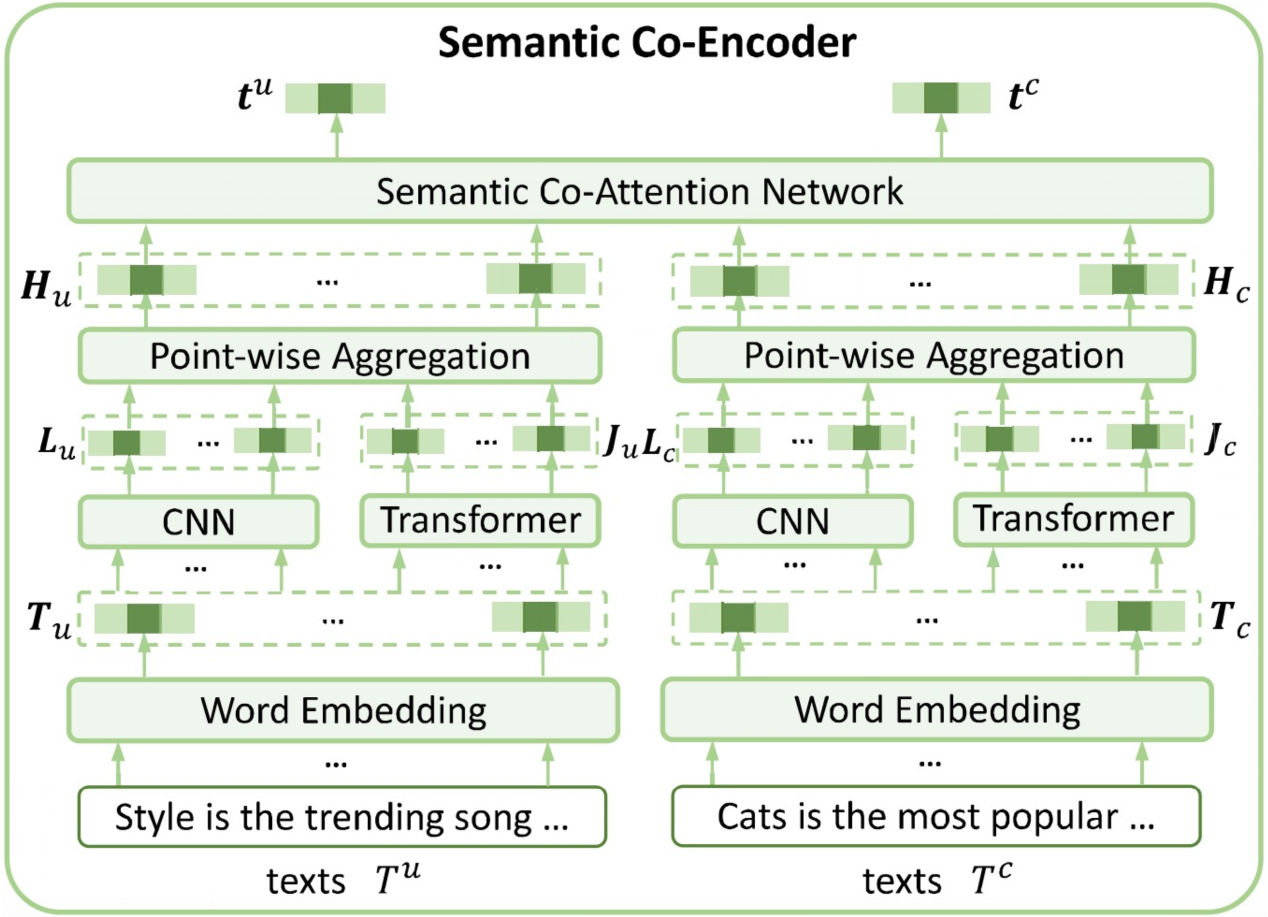
- Knowledge co-encoder
  - Stacked graph attention network (GAT)
  - Stacked graph co-attention network (GCAT)
  - Entity co-attention network





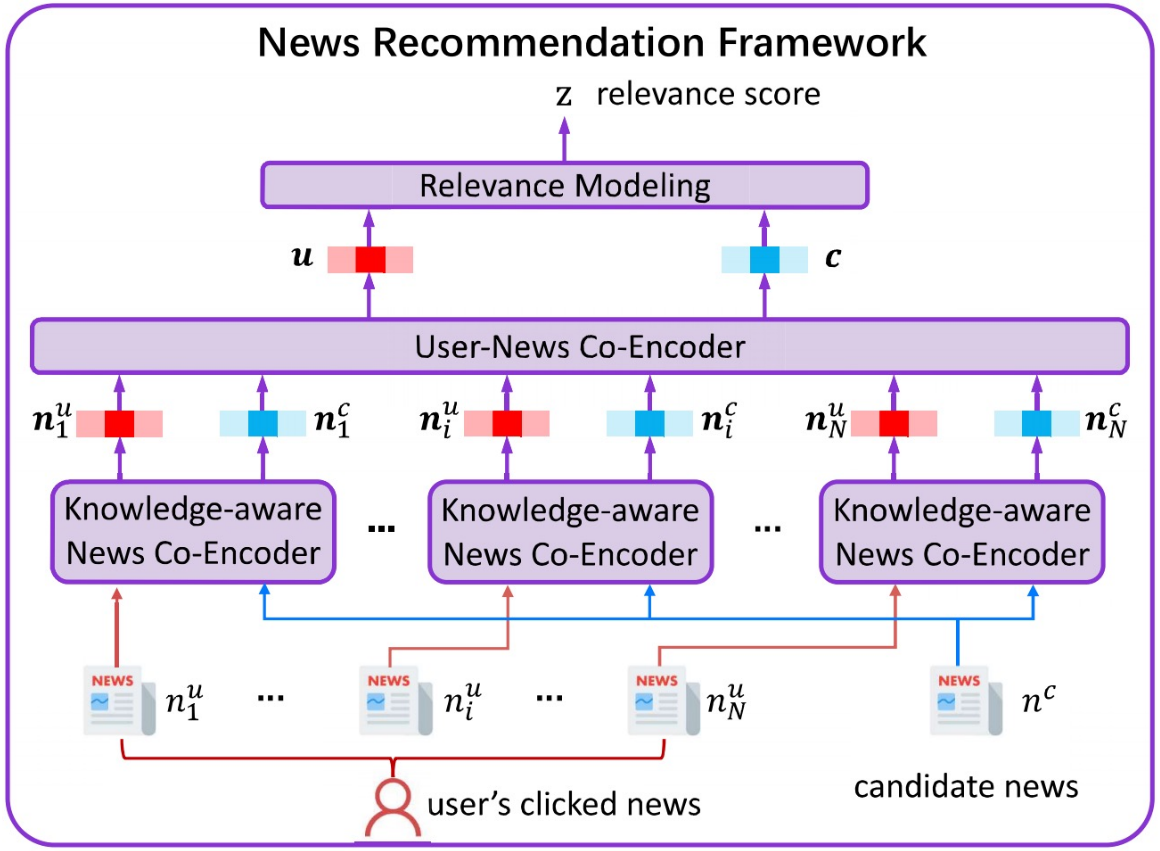
KIM

- Semantic co-encoder
  - Word embedding layer
  - Contexts modeling
  - Semantic co-attention network



KIM

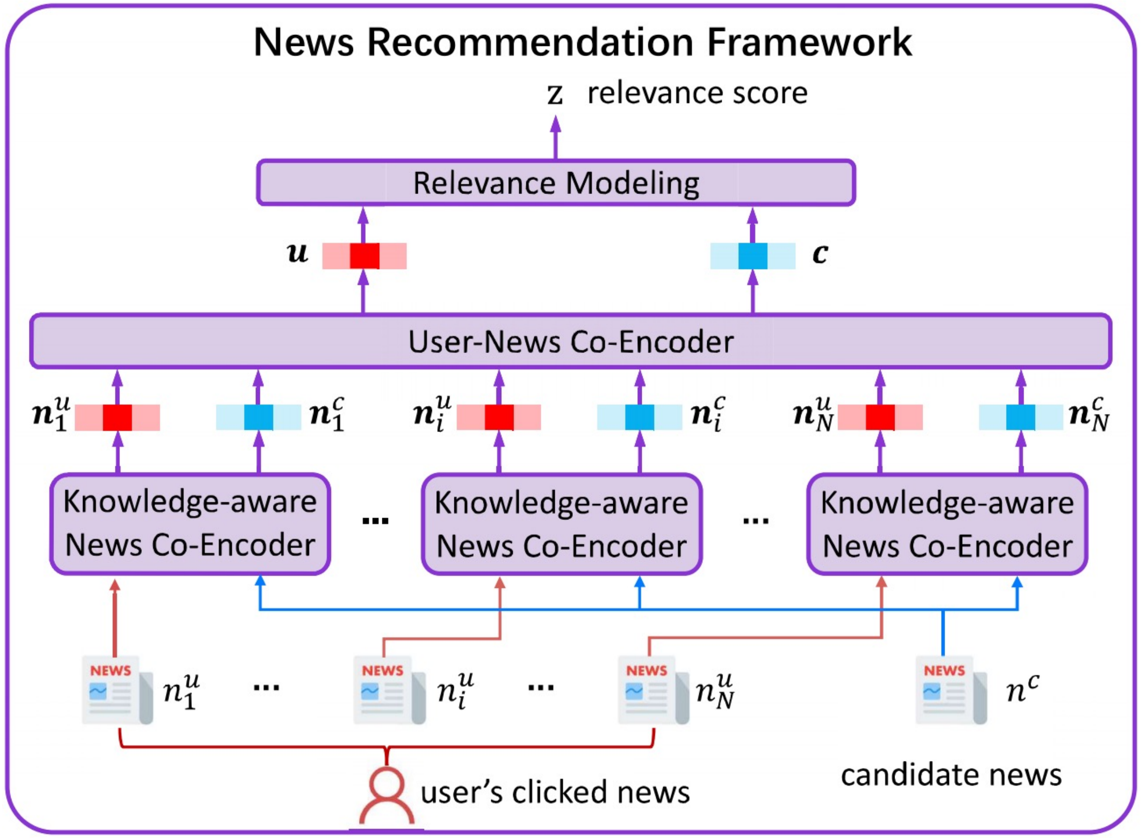
- User-news co-encoder
  - Model candidate news-aware user interest
  - Model user-aware candidate news
  - News co-attention network



KIM

- Interest matching
  - Measuring user interest in candidate news
  - $z = \boldsymbol{u} \cdot \boldsymbol{c}$

- Model Training
  - NCE loss
  - $\mathcal{L} = -\frac{1}{|S|}\sum_{i=0}^{|S|} \log(\frac{\exp(z_+^i)}{\exp(z_+^i)+\sum_{j=1}^{K}\exp(z_j^i)})$



Datasets

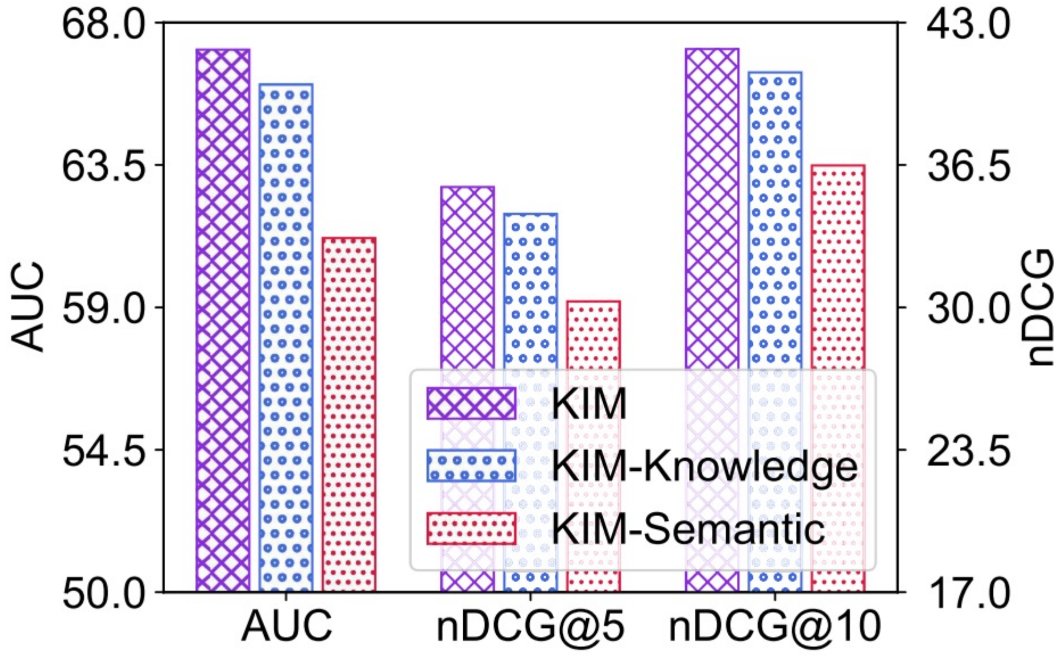
- MIND:
  - Based on user logs on Microsoft news
  - Collect user logs from 10.19 to 11.15, 2019
  - Using user logs in the last week for evaluation
  - Entities in news are extracted and linked to WikiData
- Feeds:
  - Based on user logs on a commercial news feeds in Microsoft
  - Collect user logs from 1.23 to 4.23, 2020
  - Using user logs in the last three weeks for evaluation
  - Entities in news are extracted and linked to WikiData

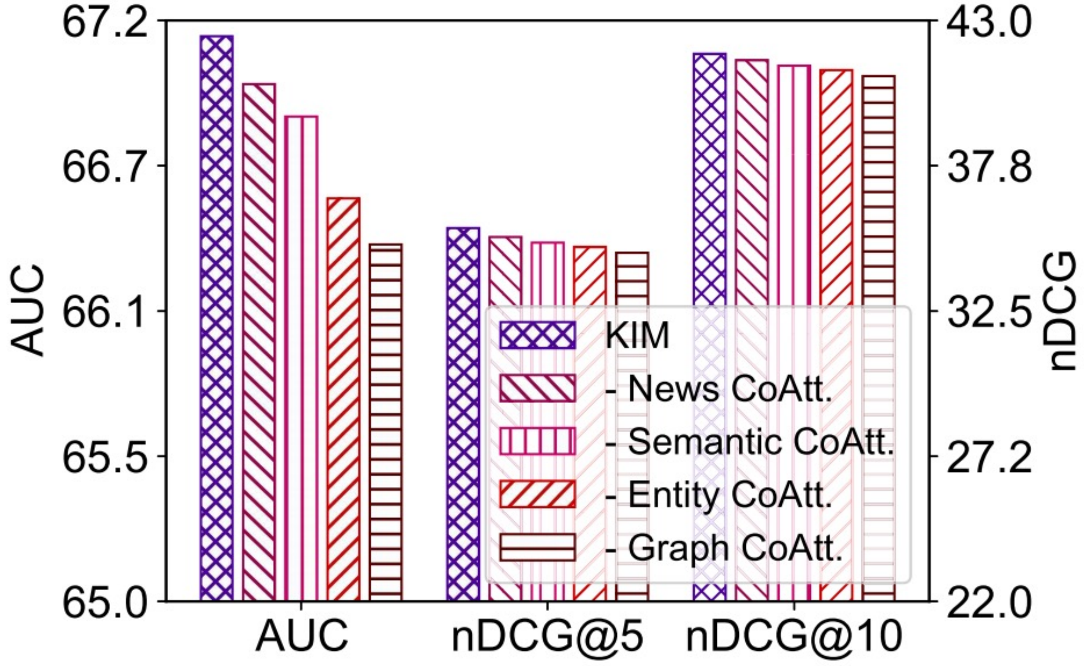
Performance Evaluation

	MIND				*Feeds*			
	AUC	MRR	nDCG@5	nDCG@10	AUC	MRR	nDCG@5	nDCG@10
EBNR	61.28±0.27	27.77±0.21	30.10±0.28	36.75±0.24	63.44±0.39	27.97±0.25	32.01±0.32	37.57±0.35
DKN	64.08±0.12	29.06±0.16	31.82±0.11	38.52±0.14	62.91±0.26	28.08±0.20	32.20±0.24	37.75±0.22
DAN	65.14±0.16	30.04±0.20	32.98±0.22	39.52±0.19	62.65±0.49	27.79±0.32	31.79±0.40	37.37±0.39
NAML	64.21±0.20	29.71±0.13	32.51±0.20	39.00±0.12	64.24±0.38	28.81±0.21	33.06±0.28	38.52±0.29
NPA	63.71±0.27	29.84±0.12	32.40±0.19	39.02±0.20	63.69±0.75	28.51±0.47	32.74±0.64	38.27±0.62
LSTUR	65.51±0.29	30.22±0.31	33.26±0.38	39.76±0.34	64.66±0.33	29.04±0.26	33.44±0.32	38.82±0.30
NRMS	65.36±0.21	30.02±0.11	33.11±0.15	39.61±0.14	65.15±0.13	29.29±0.12	33.78±0.13	39.24±0.13
FIM	64.46±0.22	29.52±0.26	32.26±0.24	39.08±0.27	65.67±0.20	29.83±0.24	34.51±0.31	39.97±0.25
KRED	65.61±0.35	30.63±0.27	33.80±0.24	40.23±0.27	65.47±0.07	29.59±0.04	34.15±0.05	39.69±0.05
KIM	**67.13**±0.29	**32.08**±0.24	**35.49**±0.34	**41.79**±0.28	**66.45**±0.13	**30.27**±0.09	**35.04**±0.09	**40.43**±0.12

KIM significantly outperforms other baseline methods

Ablation Study

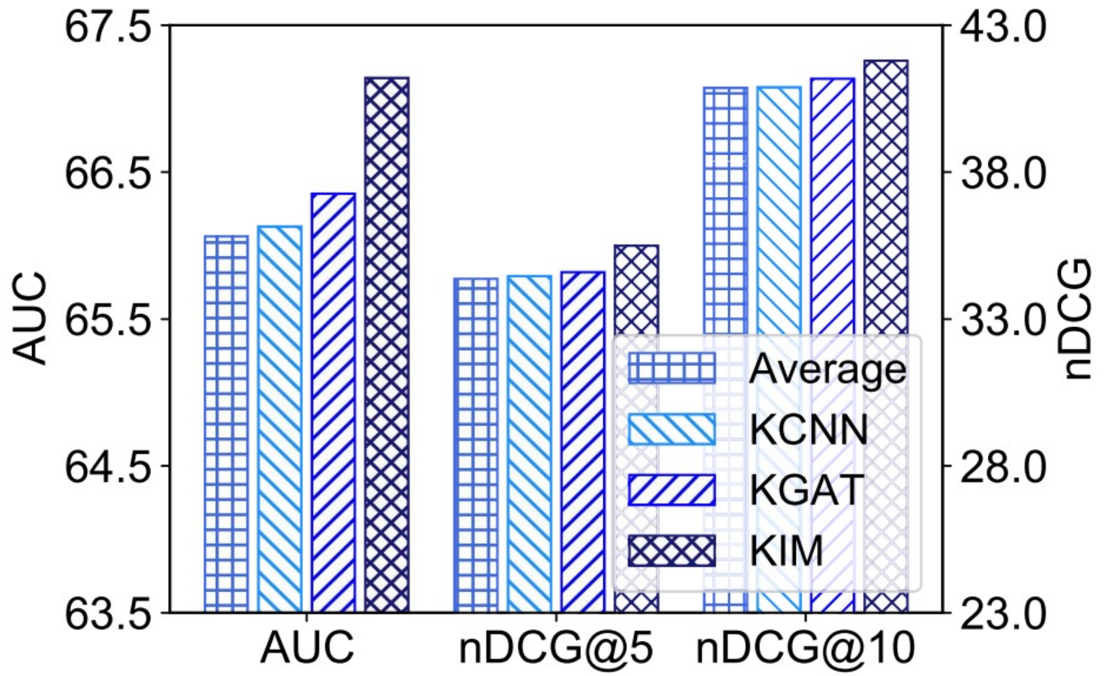




Both knowledge and semantic matching are useful for the interest matching

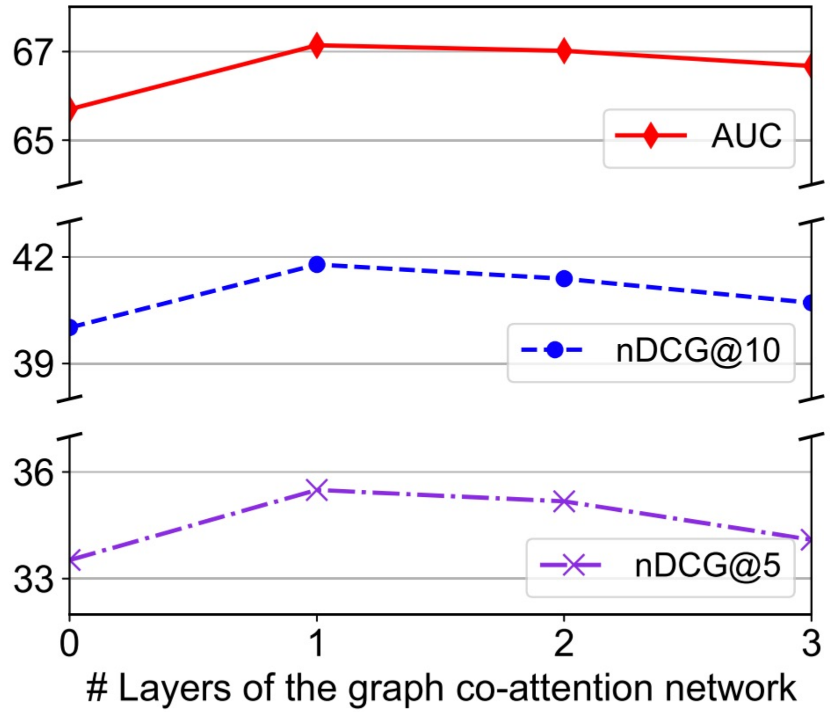
Different modules are important in our method

Knowledge Modeling



Knowledge co-encoder of KIM achieves best performance

Hyper-parameter Influence



A moderate value of # layers of GCAT is best for our method

Conclusion

- Propose a knowledge-aware interactive matching framework for personalized news recommendation
- Propose a knowledge co-encoder to learn knowledge-based representations of clicked news and candidate news from their knowledge relatedness
- Propose a semantic co-encoder to learn semantic-based representations of clicked news and candidate news from their semantic relatedness
- Propose a user-news co-encoder to learn candidate news-aware user interest representation and user-aware candidate news representation

Reference

[1] Wu et al. Neural News Recommendation with Attentive Multi-View Learning. IJCAI2019

[2] Wu et al. Neural News Recommendation with Multi-Head Self-Attention Network. EMNLP2019

[3] An et al. Neural News Recommendation with Long- and Short-term User Representations. ACL 2019

[4] Liu et al. KRED: Knowledge-Aware Document Representation for News Recommendations. RecSys. 2020

Tao Qi

taoqi.qt@gmail.com