Uni-FedRec: A Unified Privacy-Preserving News Recommendation Framework for Model Training and Online Serving

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

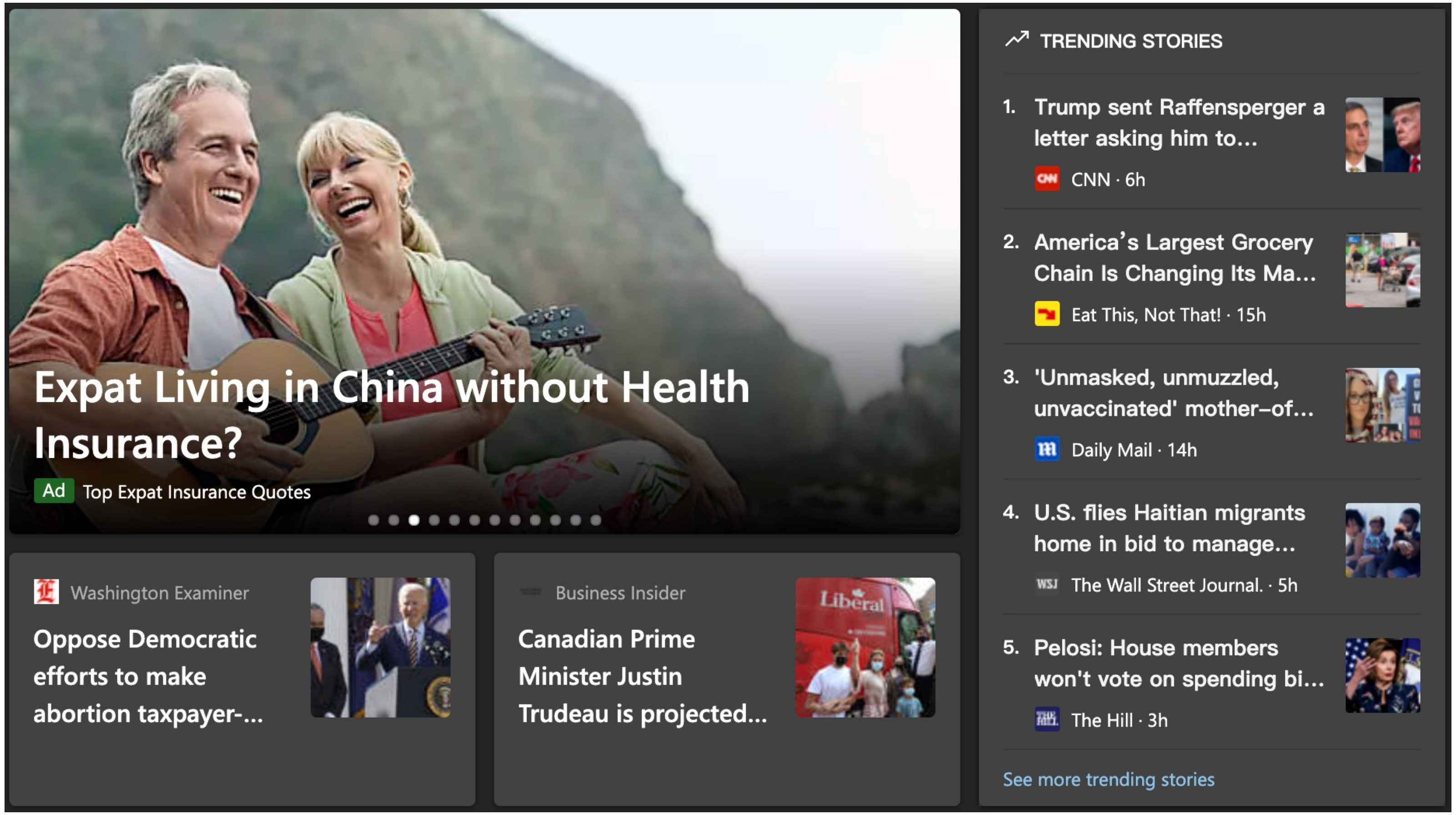
[1]Department of Electronic Engineering & BNRist, Tsinghua University, Beijing 100084, China

[2]Microsoft Research Asia, Beijing 100080, China

taoqi.qt@gmail.com

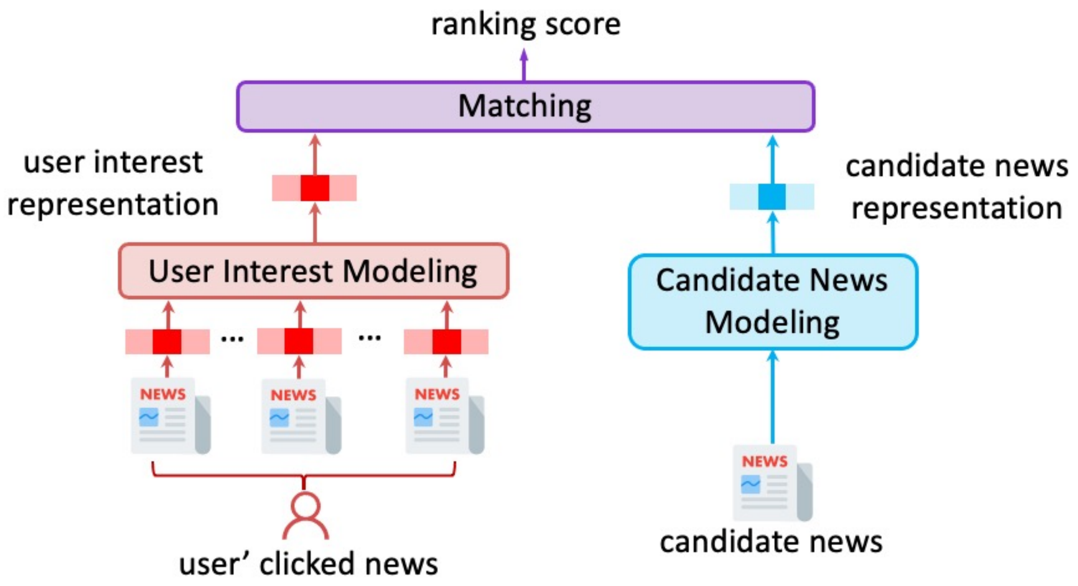
News Recommendation

- Online news platforms become popular for people to read news
- News recommendation is critical for improving user experience

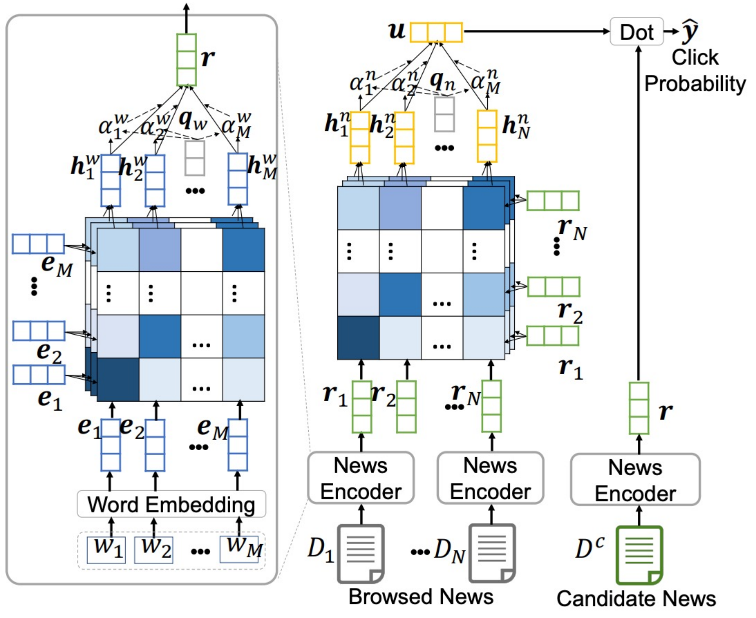


News Recommendation Methods

- Most existing methods rely on centralized storage of user data to train models and serve users



Mainstream framework

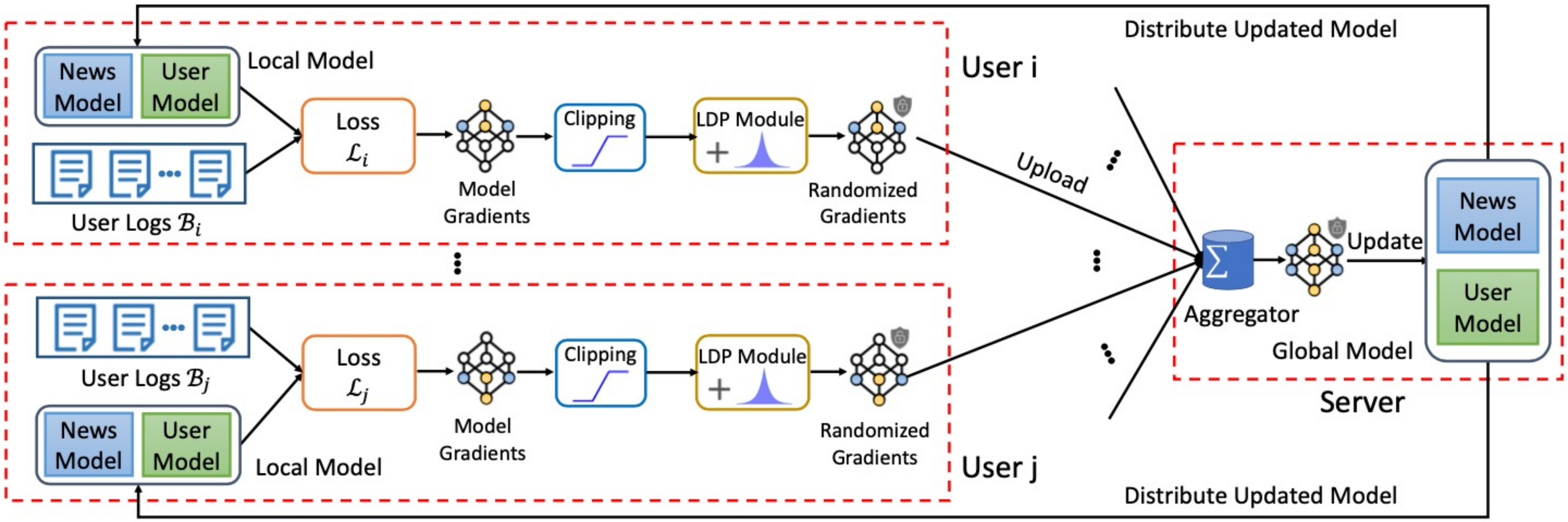


NRMS

- Challenges
  - Centralized storage of user data may arouse privacy concerns and risks
  - Application of these methods may violate some privacy regulations

Privacy-Preserving Recommender Systems

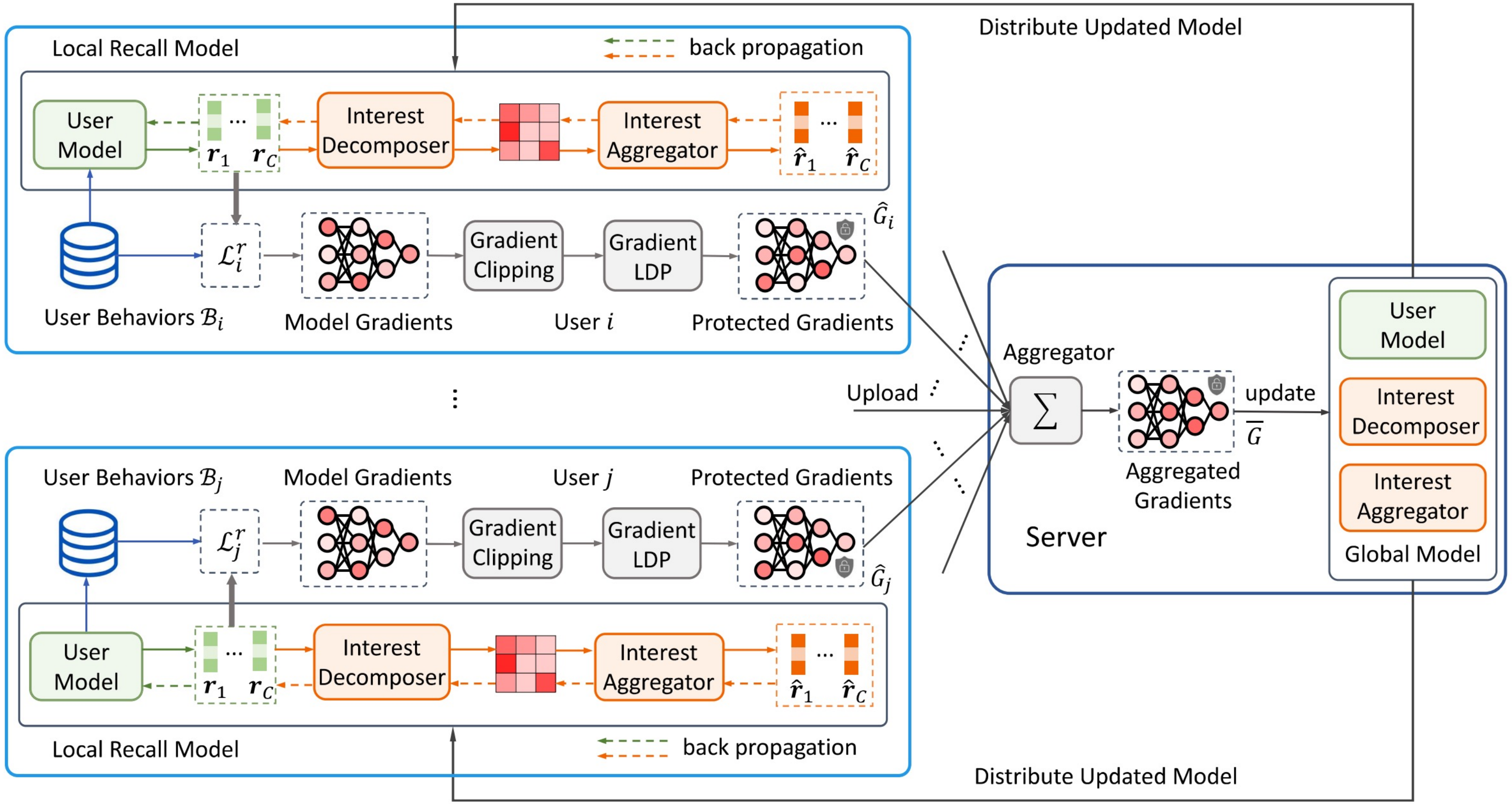
- Most existing methods focus on training a recommend model for ranking candidate items (e.g. news) in a privacy-preserving way



- Challenges
  - How to generate candidate news and serve users with decentralized user data in a privacy-preserving way remains an open problem

Uni-FedRec

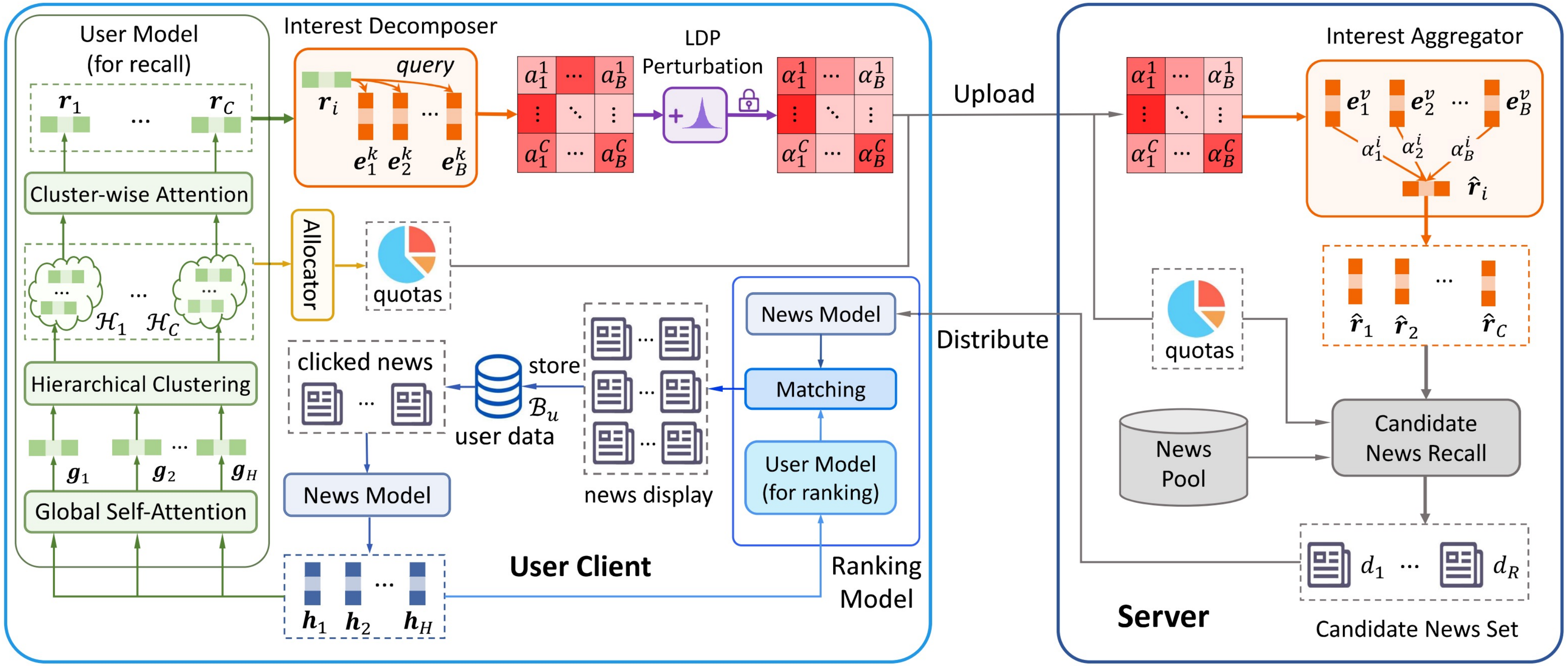
- A unified privacy-preserving news recommendation framework for both model training and online serving



The framework of Uni-FedRec for privacy-preserving model training

Uni-FedRec

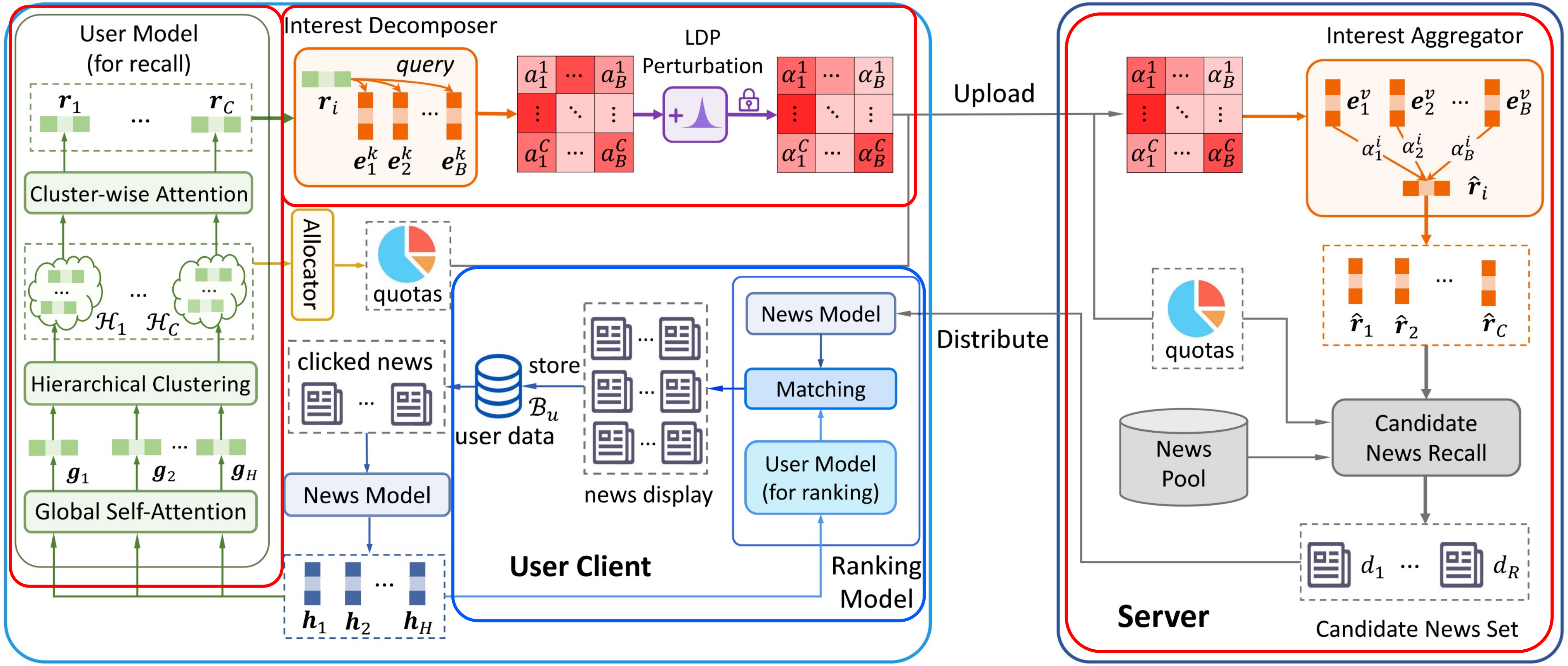
- A unified privacy-preserving news recommendation framework for both model training and online serving



The framework of Uni-FedRec for privacy-preserving online serving

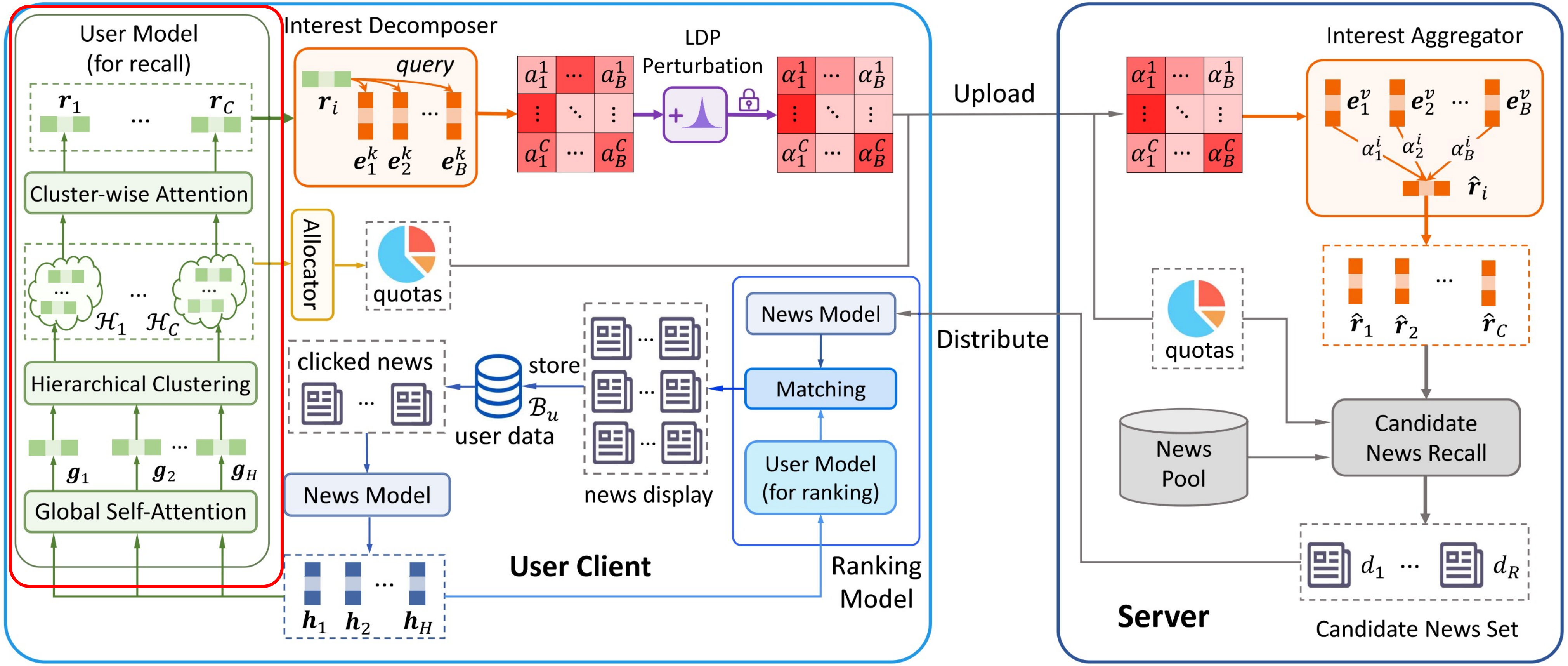
Uni-FedRec: Privacy-Preserving Online Serving

- Recommend news according to user interest with decentralized user data
  - Privacy-preserving news recall framework; Local news ranking framework



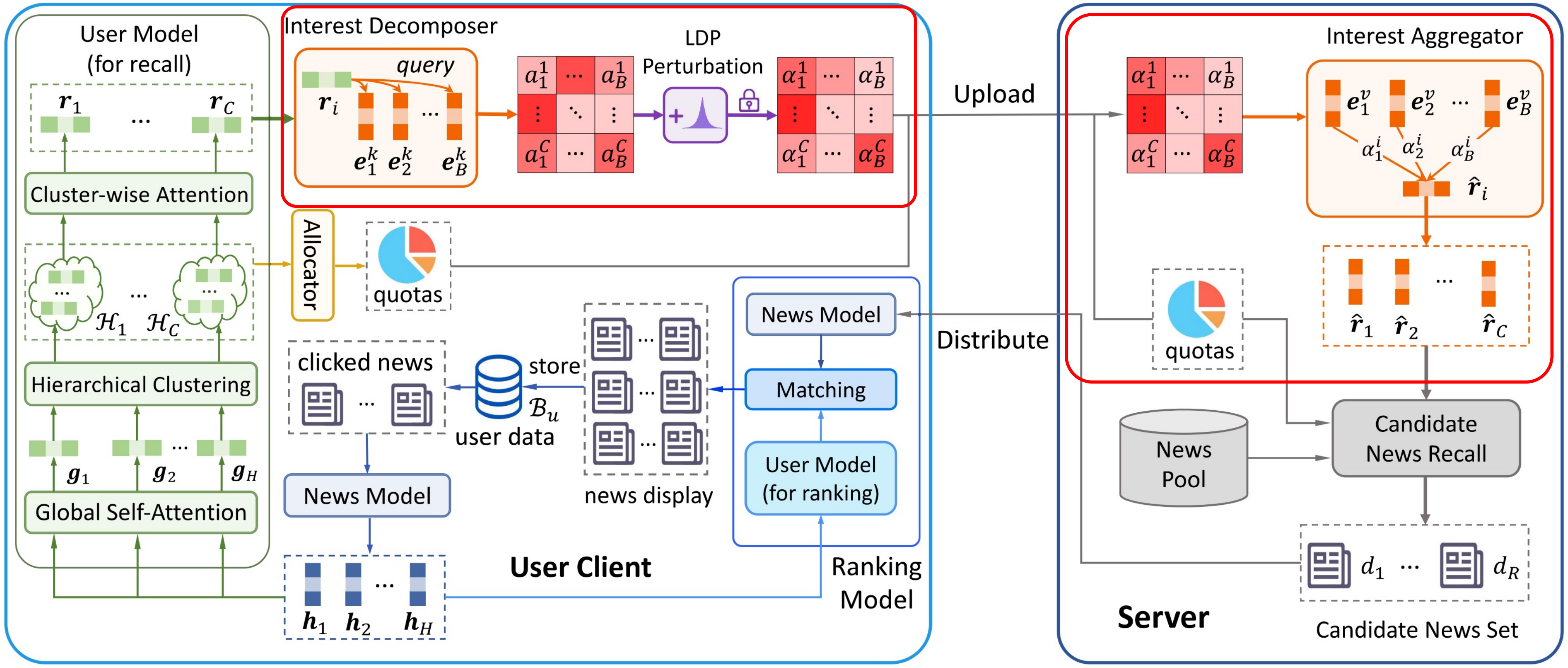
User Model for News Recall

- Users usually have multiple interest
- Learn multiple representation to model diverse user interest



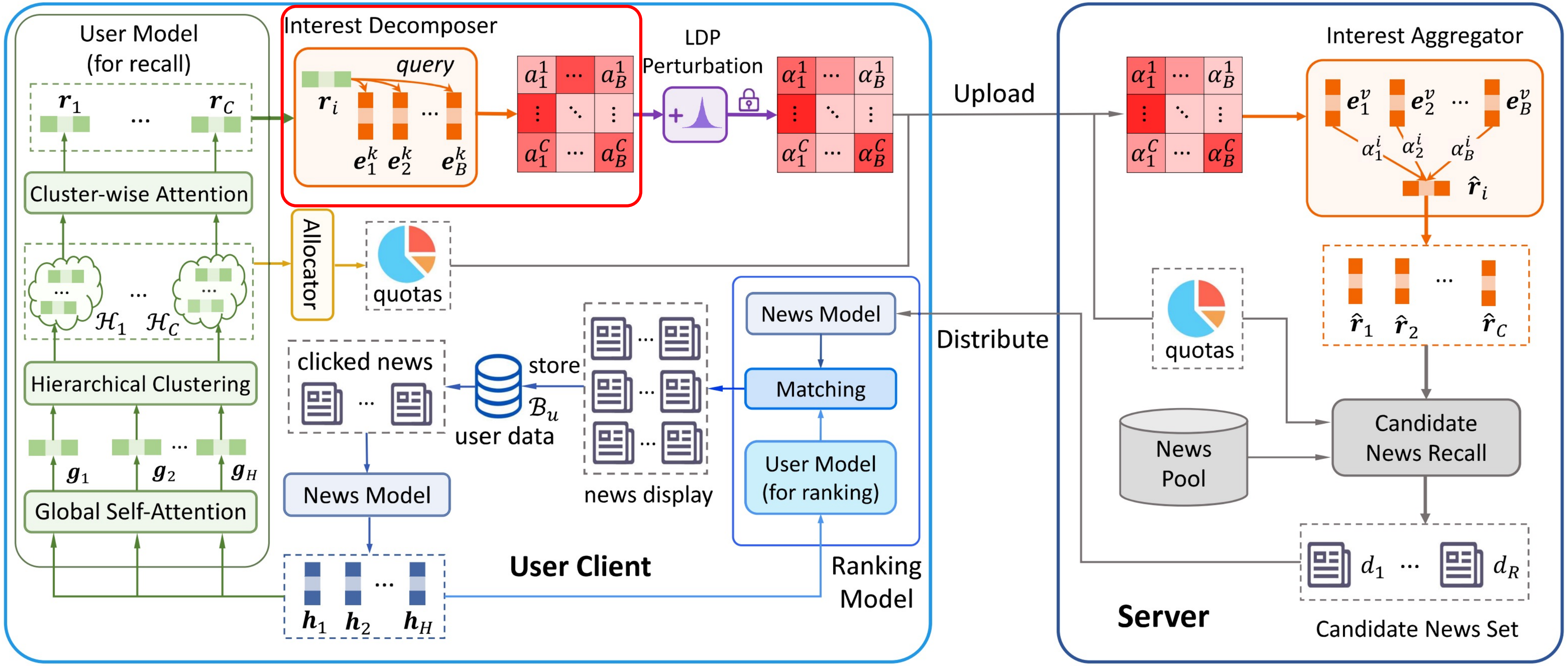
Interest Decomposer-Aggregator Framework

- Synthesize interest representations via basic interest embeddings



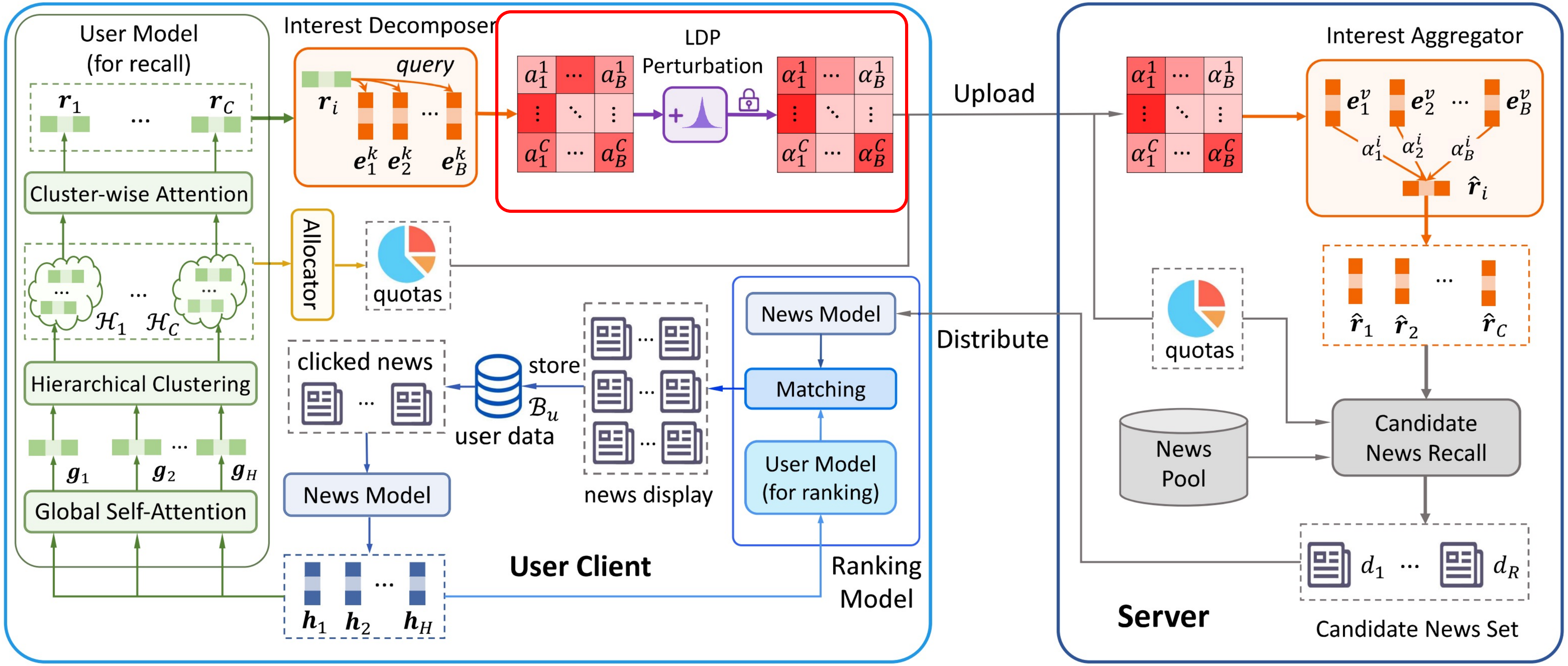
Interest Decomposer-Aggregator Framework

- Synthesize interest representations via basic interest embeddings
- Interest decomposer: $a_j^i = \boldsymbol{r}_i \cdot \boldsymbol{e}_j^k,\ j = 1,2,\dots,B$



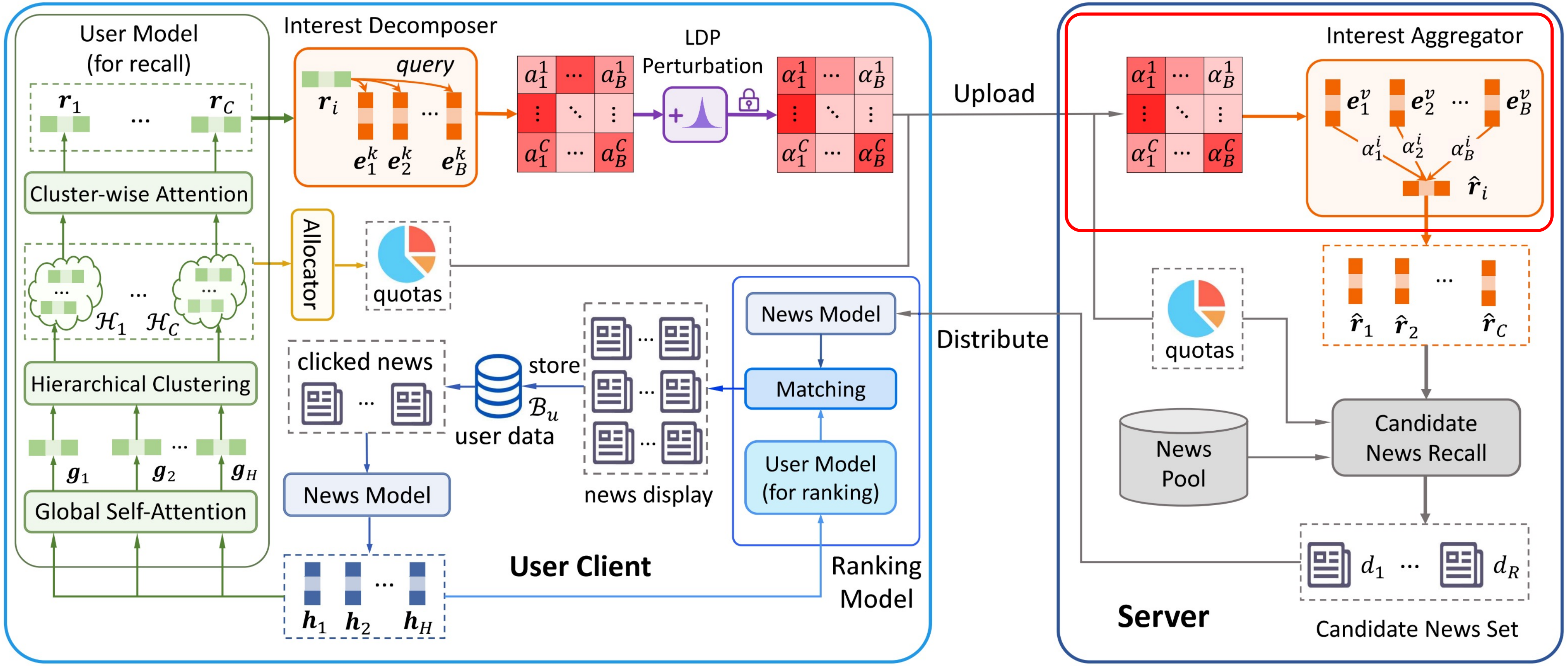
Interest Decomposer-Aggregator Framework

- Synthesize interest representations via basic interest embeddings with noise
- Perturbation noise: $\hat{a}_j^i = f_\delta(a_j^i) + n_I,\ n_I \sim La(0, \lambda_I)$



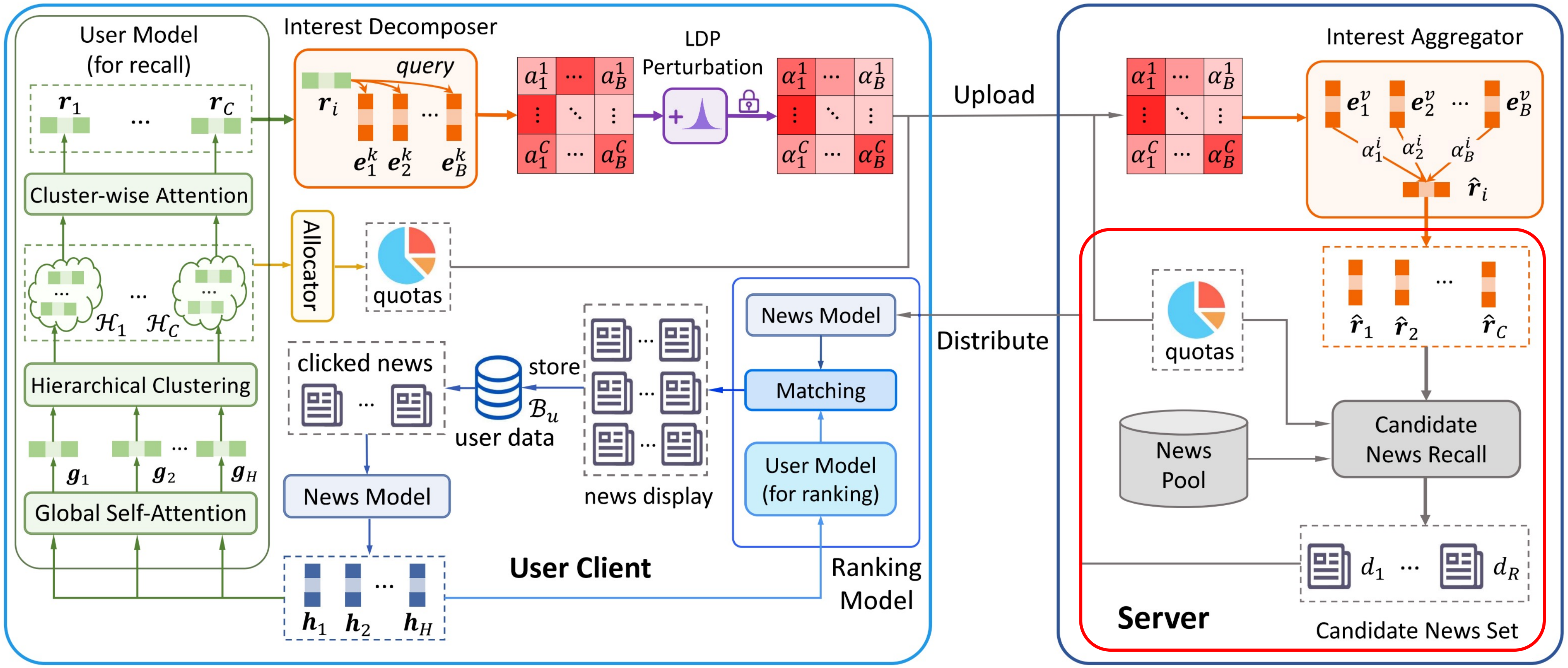
Interest Decomposer-Aggregator Framework

- Synthesize interest representations via basic interest embeddings with noise
- Interest aggregator: $\hat{r}_i = \sum_{j=1}^{B} \alpha_j^i \boldsymbol{e}_j^v$



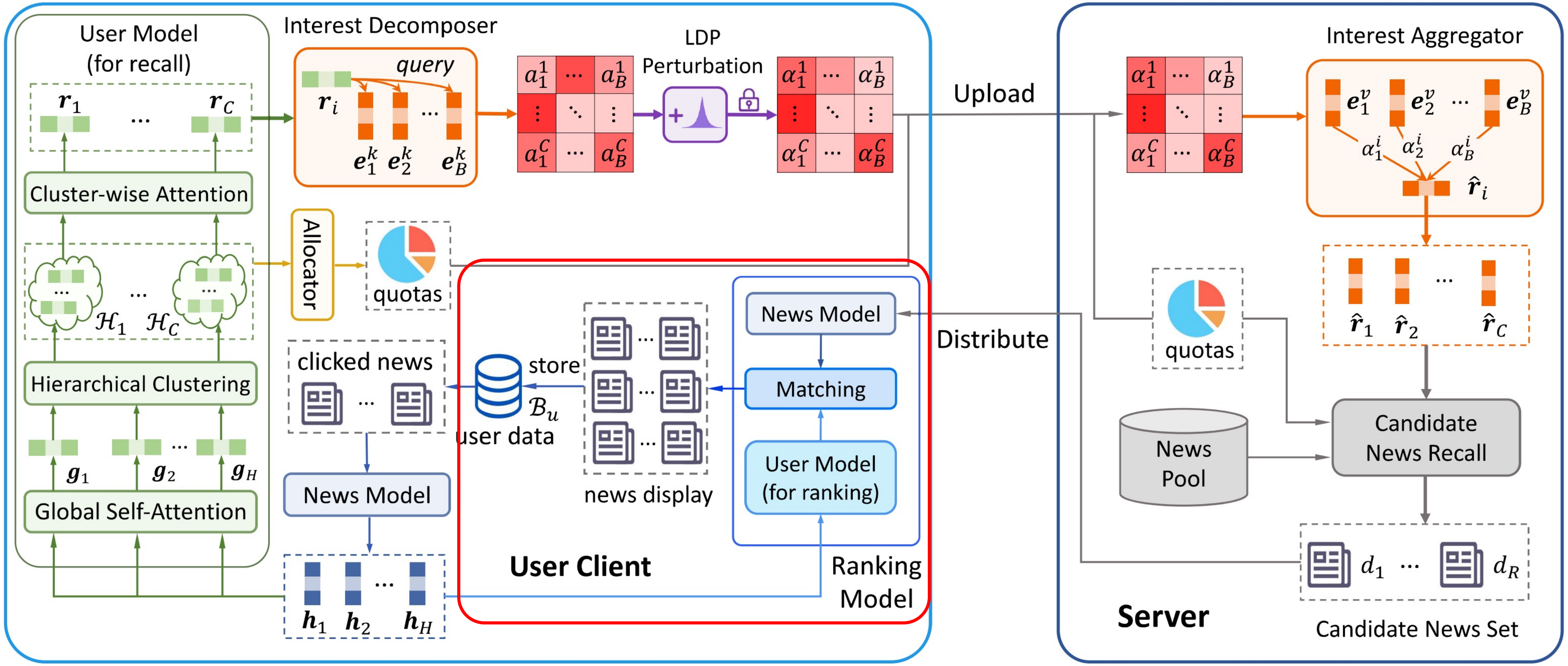
Multi-Channel News Recall

- Recall news according to different user interest representations
- Integrate candidate news generated by different channels



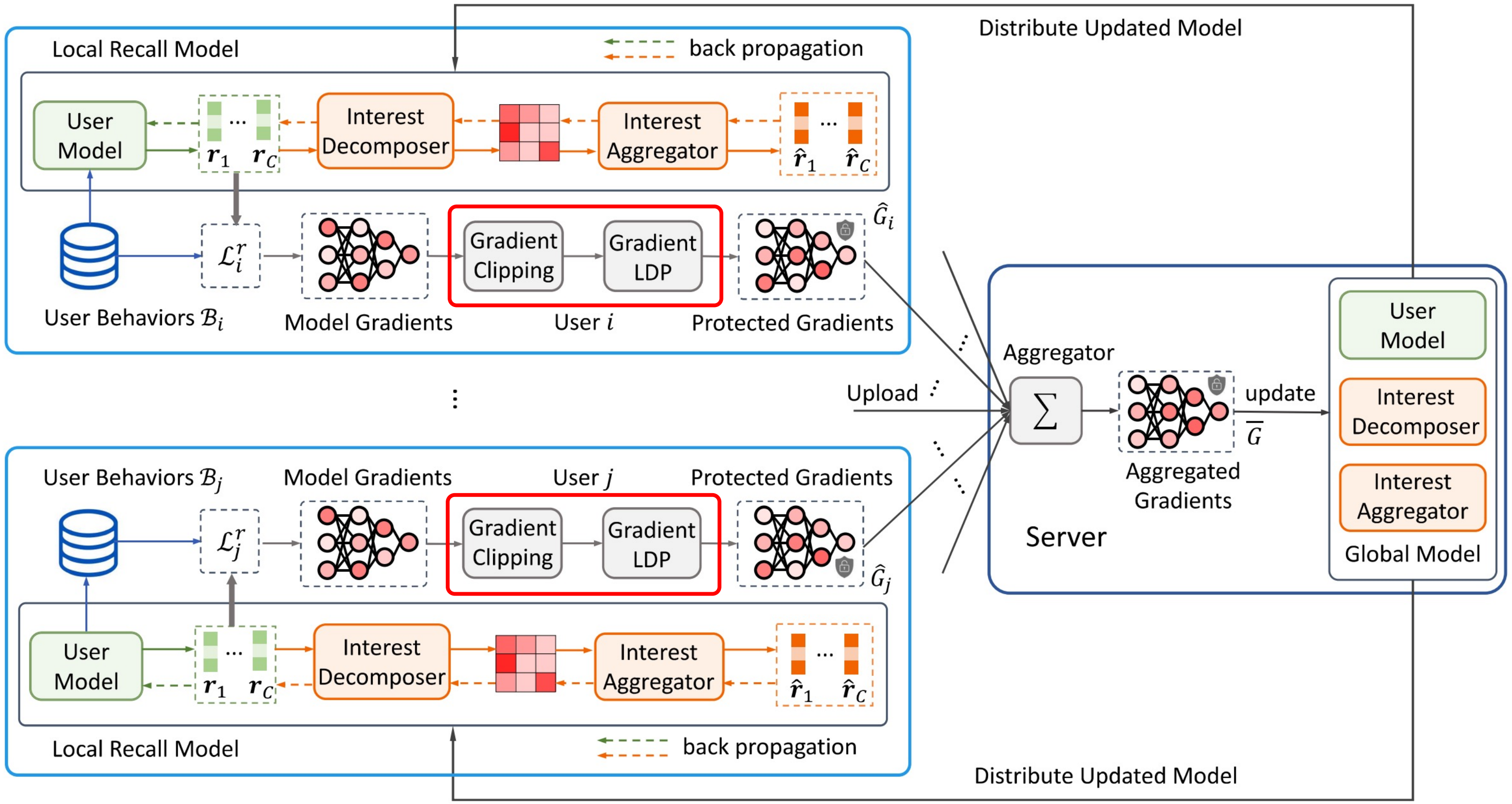
Local News Ranking

- Locally rank candidate news in the user client via existing personalized news ranking methods



Privacy-Preserving Model Training

- Privacy-preserving model training with federated learning
- Protect uploaded gradients with LDP: $\widehat{\boldsymbol{G}}_u = f_\theta(\boldsymbol{G}_u) + n_g, n_g \sim La(0, \lambda_g)$



Datasets

- MIND:
  - A public news recommendation dataset based on Microsoft News
  - Constructed by user logs from 2019.10.19 to 2019.11.15 (6 weeks)
- NewsFeeds:
  - Constructed by user logs on a news feeds app in Microsoft
  - Constructed by user logs from 2020.01.23 to 2020.04.23 (13 weeks)

	# News	# Users	# Clicks	#Impressions
MIND	161,013	1,000,000	24,155,470	15,777,377
NewsFeeds	120,219	20,000	112,927	48,923

News Recall Performance Comparison

	MIND				NewsFeeds			
	R@100	R@200	R@300	R@400	R@100	R@200	R@300	R@400
YoutubeNet	1.50±0.03	2.43±0.08	3.34±0.08	3.96±0.13	0.60±0.02	0.92±0.01	1.17±0.01	1.45±0.02
HUITA	1.69±0.06	2.67±0.04	3.37±0.06	3.97±0.06	0.60±0.01	0.91±0.01	1.18±0.03	1.45±0.01
EBNR	2.31±0.17	3.72±0.13	4.69±0.17	5.61±0.17	0.64±0.03	0.96±0.05	1.28±0.06	1.55±0.06
SASRec	2.22±0.05	3.51±0.07	4.54±0.07	5.38±0.07	0.62±0.06	0.96±0.01	1.20±0.06	1.49±0.05
PinnerSage	1.22±0.14	1.85±0.28	2.69±0.23	3.53±0.20	0.59±0.01	0.93±0.01	1.15±0.01	1.45±0.02
Octopus	1.26±0.03	1.93±0.07	2.74±0.06	3.55±0.06	0.60±0.02	0.92±0.02	1.17±0.02	1.44±0.03
Uni-FedRec	**2.95**±0.11	**4.13**±0.12	**5.13**±0.12	**5.99**±0.11	**0.80**±0.08	**1.14**±0.10	**1.60**±0.12	**2.03**±0.12

News recall performance of different methods.
Higher recall rates means **better** performance.

Privacy Protection Performance Comparison

	MIND				NewsFeeds			
	R@100	R@200	R@300	R@400	R@100	R@200	R@300	R@400
YoutubeNet	12.29	15.91	18.48	20.64	29.43	31.22	32.46	33.47
HUITA	13.44	16.11	17.98	19.49	29.51	31.24	32.44	33.39
EBNR	5.49	8.27	10.30	12.05	11.35	13.08	14.14	14.86
SASRec	6.00	8.71	10.81	12.52	7.78	9.18	10.16	11.03
PinnerSage	16.91	21.35	24.48	27.18	29.43	31.10	32.32	33.38
Octopus	17.04	21.62	24.72	27.31	29.45	31.15	32.36	33.38
Uni-FedRec	**0.55**	**1.14**	**1.69**	**2.22**	**0.23**	**0.54**	**0.83**	**1.08**

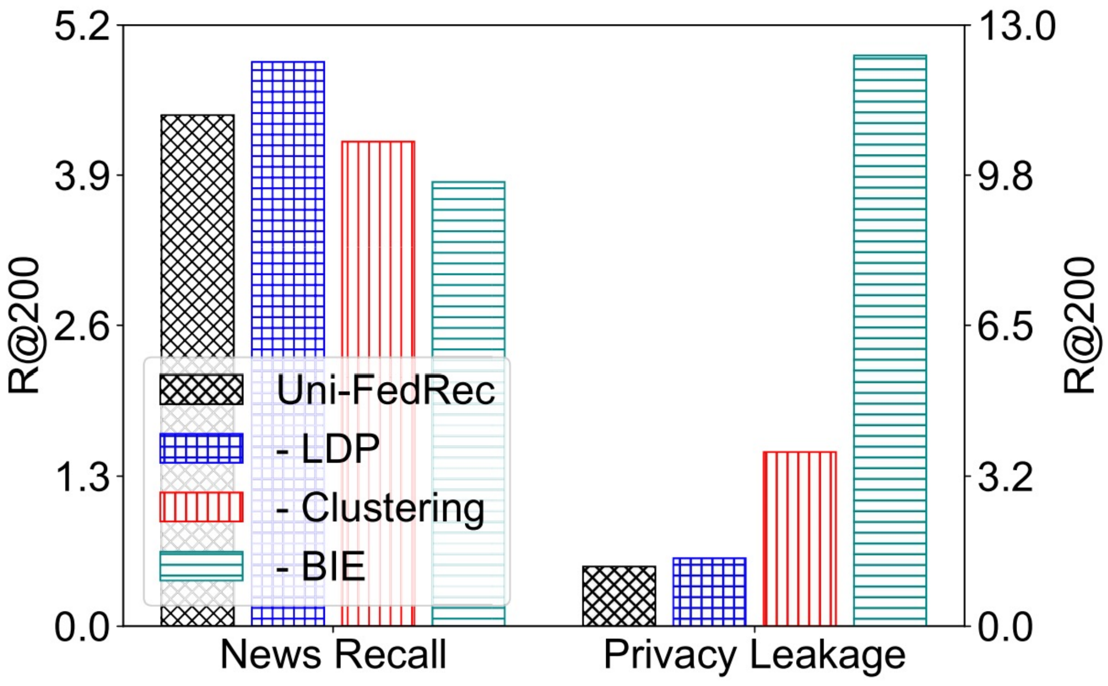
Privacy protection ability is measured by rates of user’s historical clicked news recalled from the news pool. **Lower** recall rates means **better** privacy protection performance.

Recommendation Performance

	FedRec	LSTUR	NRMS	NAML
YoutubeNet	70.65	68.53	68.79	65.93
HUITA	70.48	68.76	70.48	68.76
EBNR	75.56	73.82	75.01	70.89
SASRec	75.07	72.51	73.35	70.51
PinnerSage	69.26	68.96	67.28	66.09
Octopus	69.76	69.12	67.11	65.75
Uni-FedRec	**79.26**	**77.31**	**78.91**	**75.40**

Recommendation performance (AUC) of different methods, where rows and columns are different recall and ranking methods, respectively.

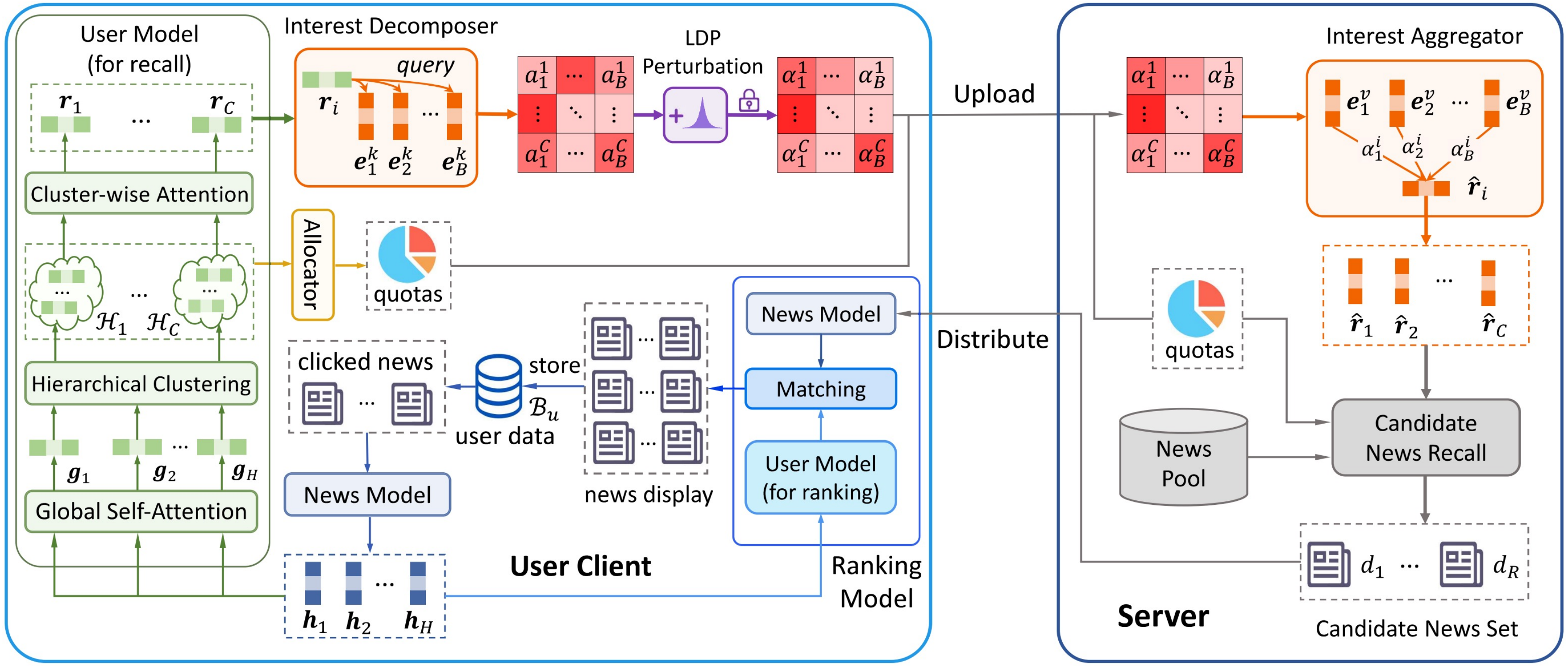
Ablation Study on Uni-FedRec



Effectiveness of different modules in recall performance and privacy protection

Conclusion

- Propose a unified privacy-preserving news recommendation framework for both online serving and model training
- Propose a privacy-preserving recall model which can compressively model user interest and protect user privacy



Tao Qi

taoqi.qt@gmail.com