上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

Positive, Negative and Neutral: Modeling Implicit Feedback in Session-based News Recommendation

Shanghai Jiao Tong University
Shanghai, China
gongshansan@sjtu.edu.cn

Shansan Gong

Kenny Q. Zhu

Contents.

01 Background and intuition

02 Modeling and methods

03 Experiments and Analysis

04 Conclusion and future work

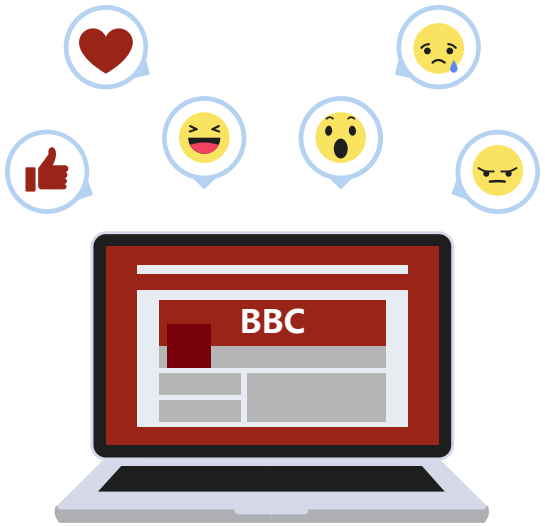
01

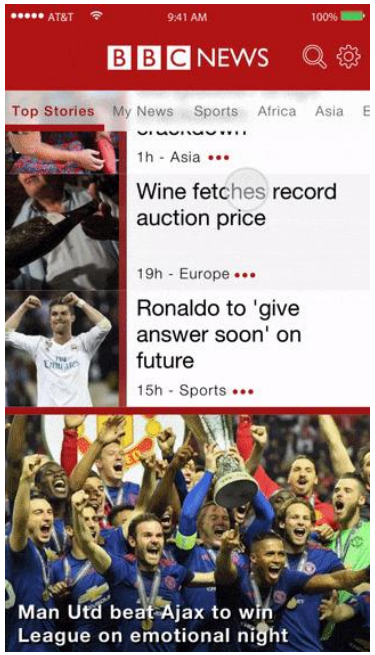
Background and intuition

Introduction of session-based news recommendation

Background

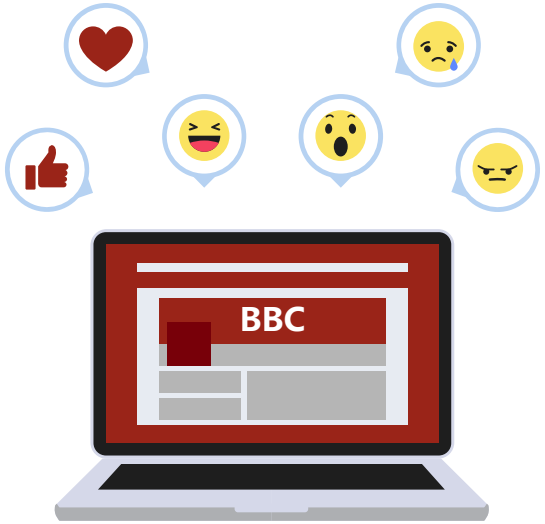
Sparse users | Dynamic news | Unstable interests | Diversity requirement





Background

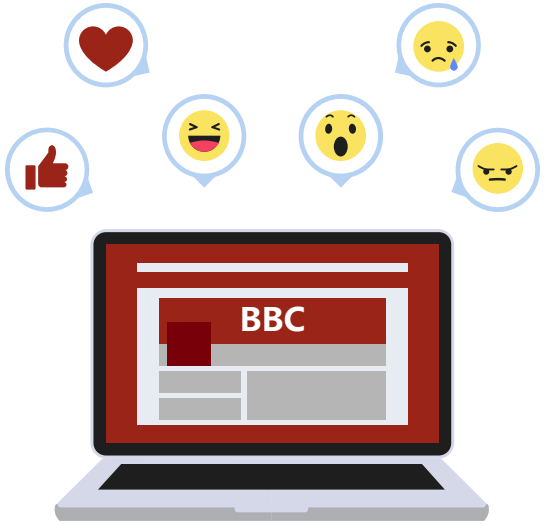
Sparse users | Dynamic news | Unstable interests | Diversity requirement



- Rely on long term user id in UI matrix
- Rely on pure text information

Background

Sparse users | Dynamic news | Unstable interests | Diversity requirement



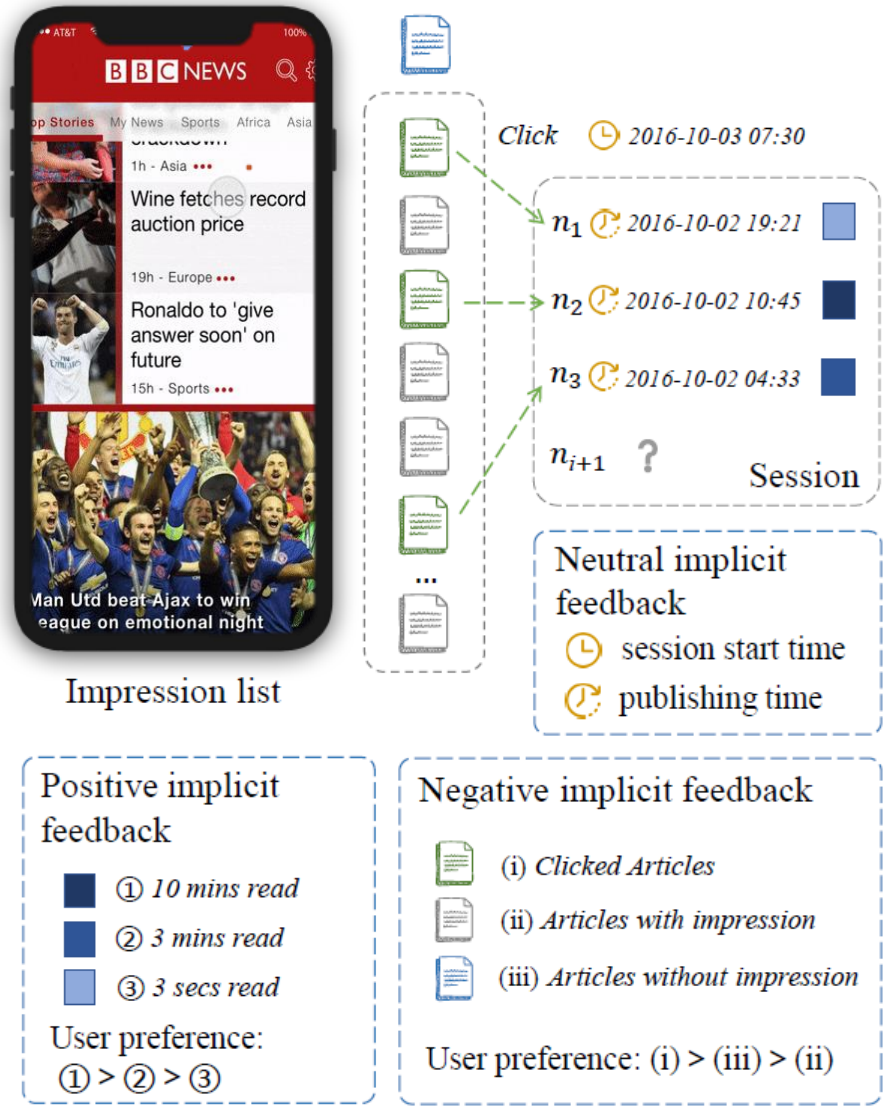
- Rely on long term user id in UI matrix
- Click means “like”
- Rely on pure text information

Intuition

"click"means "like"

Rely on textual information

Implicit feedback



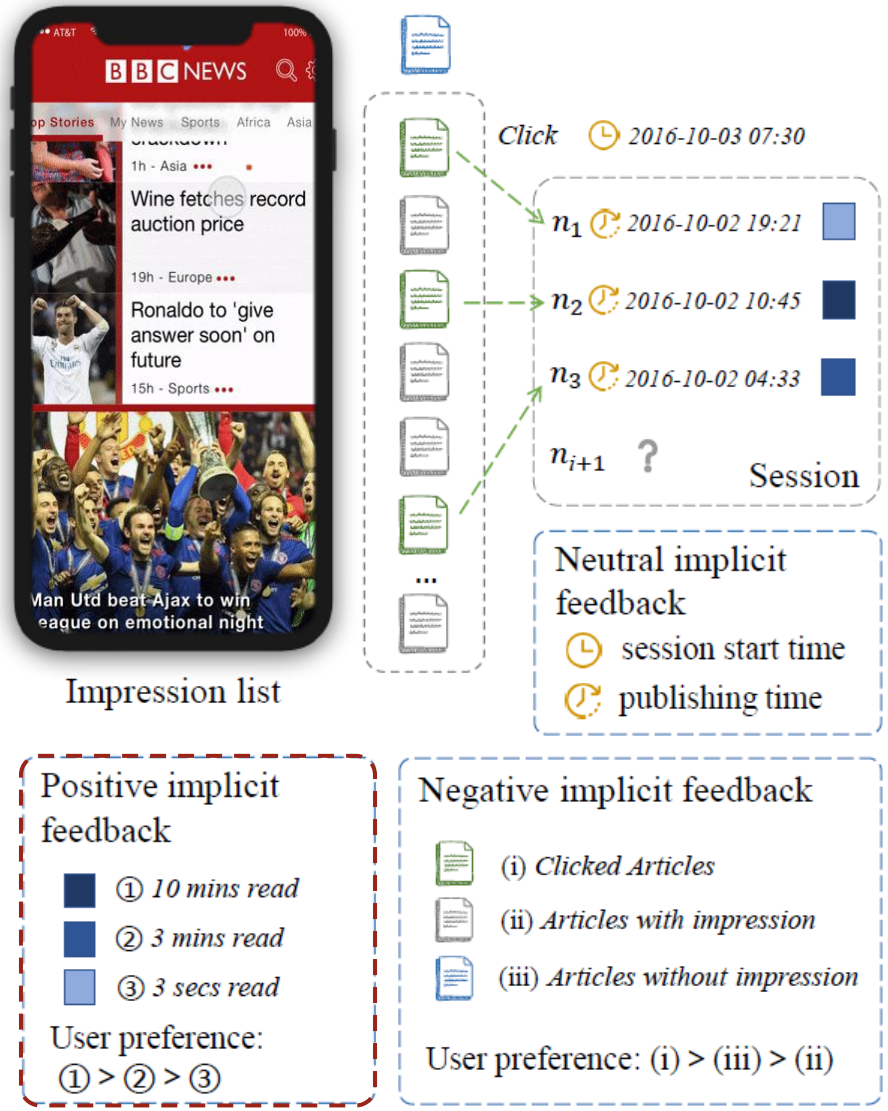
Intuition

"click"means "like"

Rely on textual information

Implicit feedback

Positive Implicit feedback (active time)



Intuition

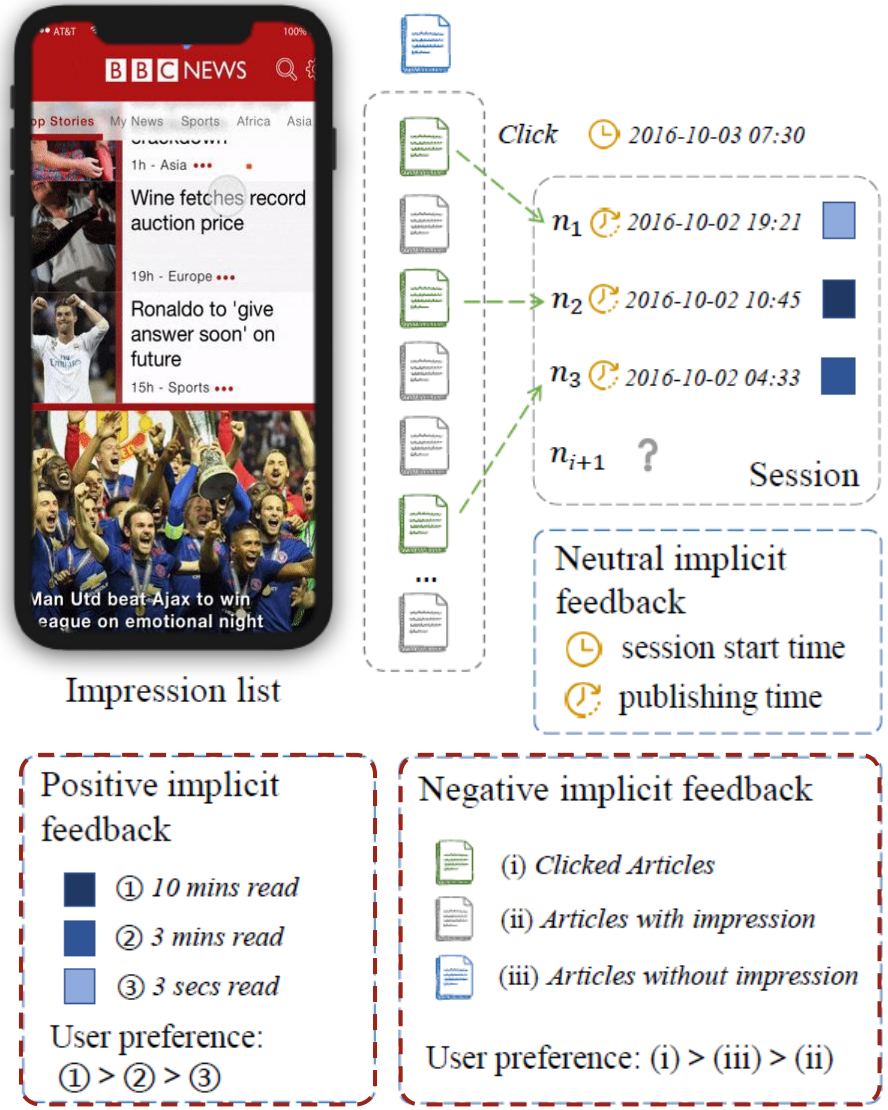
"click"means "like"

Rely on textual information

Implicit feedback

Positive Implicit feedback (active time)

Negative Implicit feedback (not clicked in impression)



Intuition

"click"means "like"

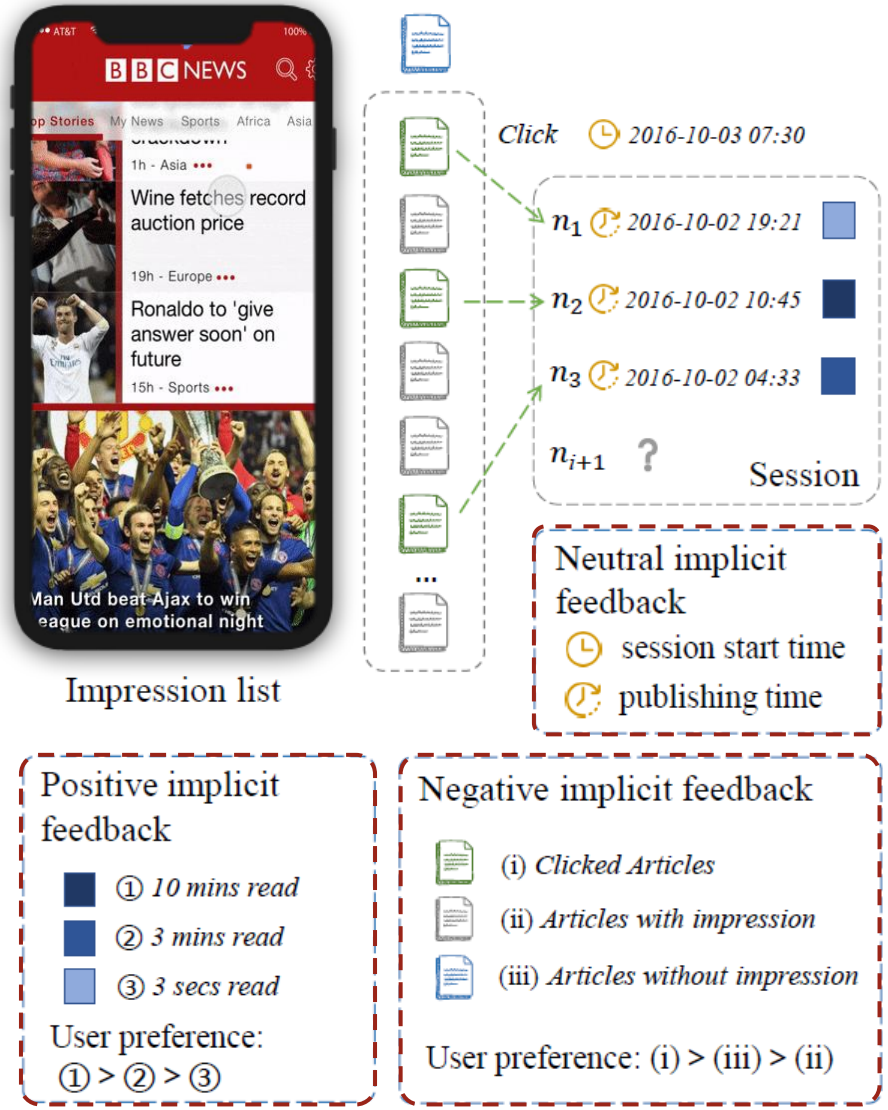
Rely on textual information

Implicit feedback

Positive Implicit feedback (active time)

Negative Implicit feedback (not clicked in impression)

Neutral Implicit feedback (start time & publishing time)



Session-based Recommendation

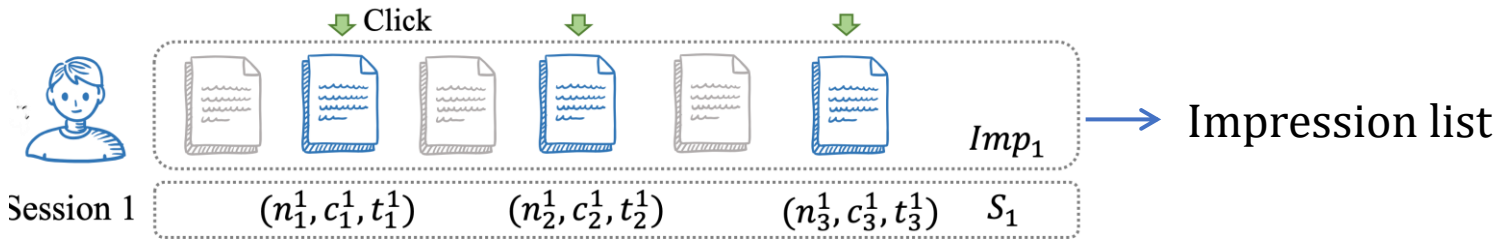
Rely on textual information

Rely on user id → Session-based recommender

Session-based Recommendation

Rely on textual information

Rely on user id → Session-based recommender

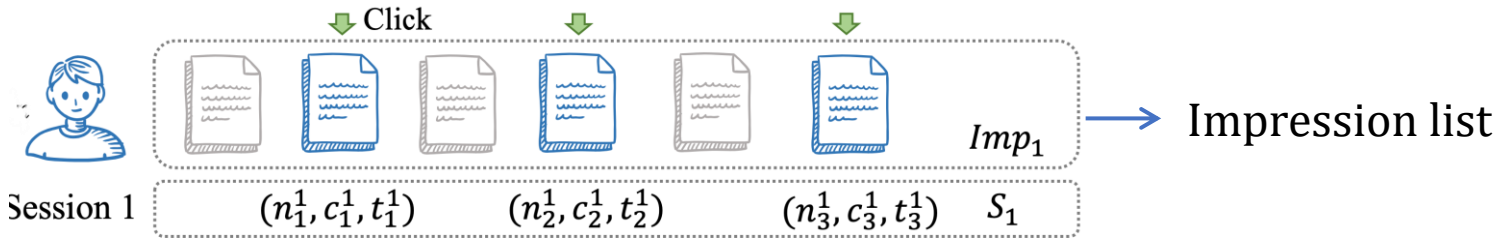


→ Impression list

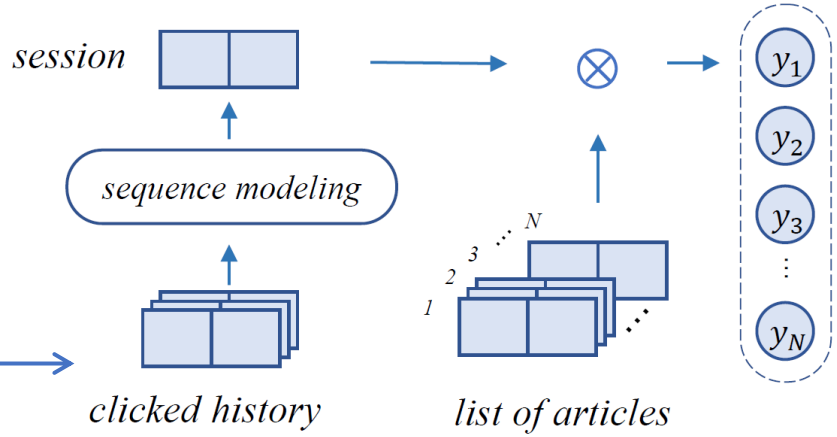
Session-based Recommendation

Rely on textual information

Rely on user id → Session-based recommender



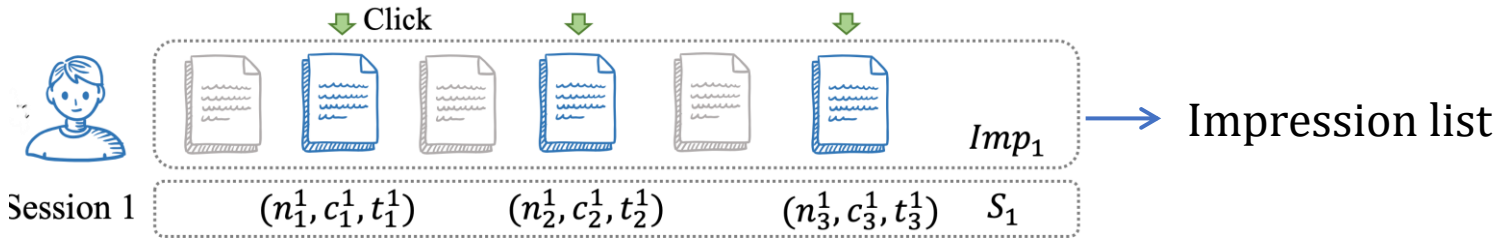
$$S_u : (n_1^u, c_1^u, t_1^u) \rightarrow (n_2^u, c_2^u, t_2^u) \rightarrow \ldots \rightarrow (n_i^u, c_i^u, t_i^u),$$



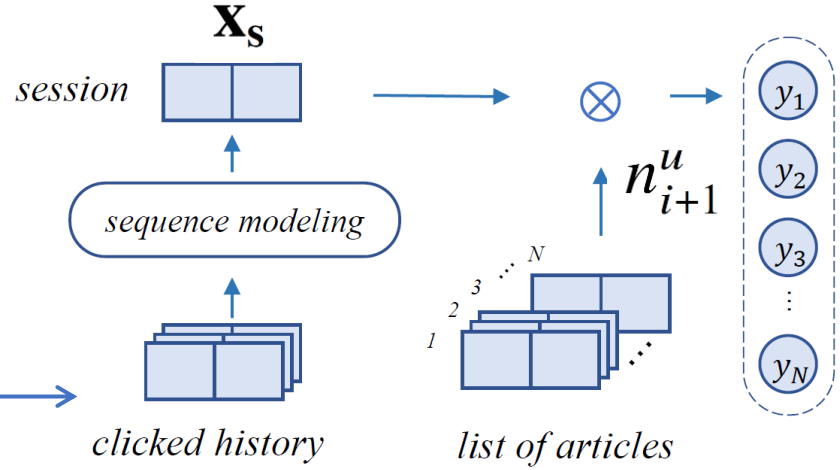
Session-based Recommendation

Rely on textual information

Rely on user id → Session-based recommender



$$S_u : (n_1^u, c_1^u, t_1^u) \rightarrow (n_2^u, c_2^u, t_2^u) \rightarrow \ldots \rightarrow (n_i^u, c_i^u, t_i^u),$$



02

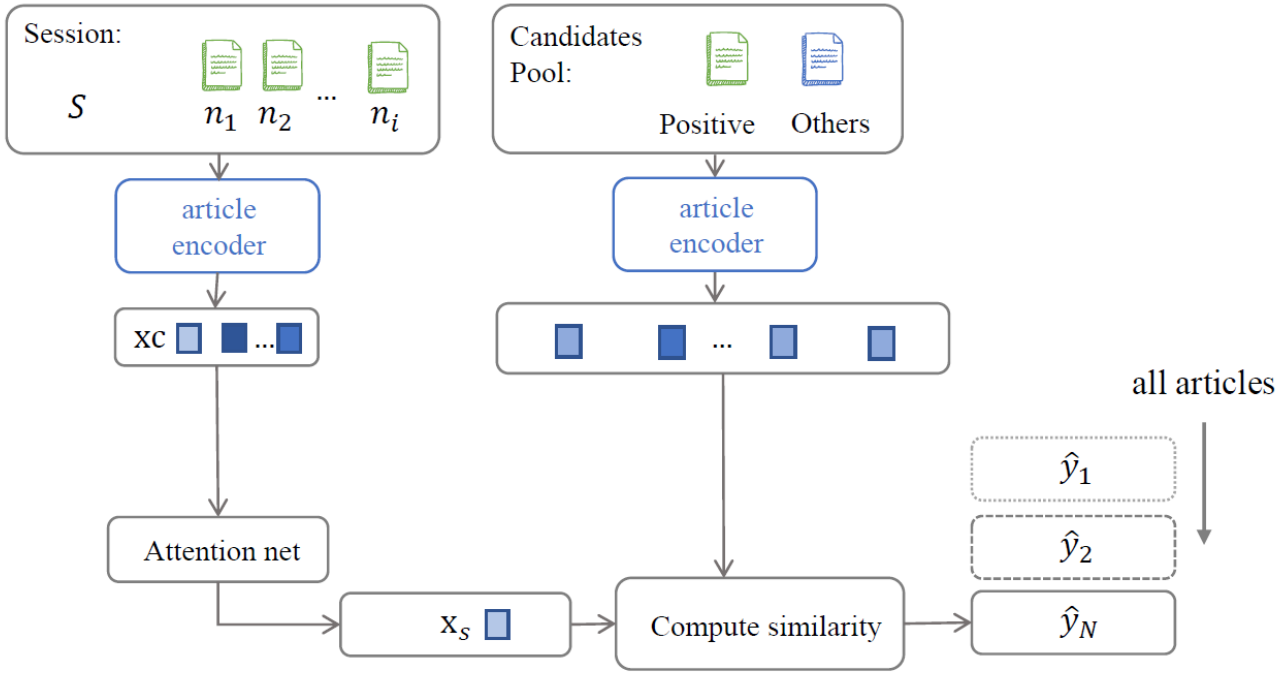
Modeling and methods

Our architecture and representation

Base model

article encoder

Content representative vector ⊕ **co-occurrence representative vector**

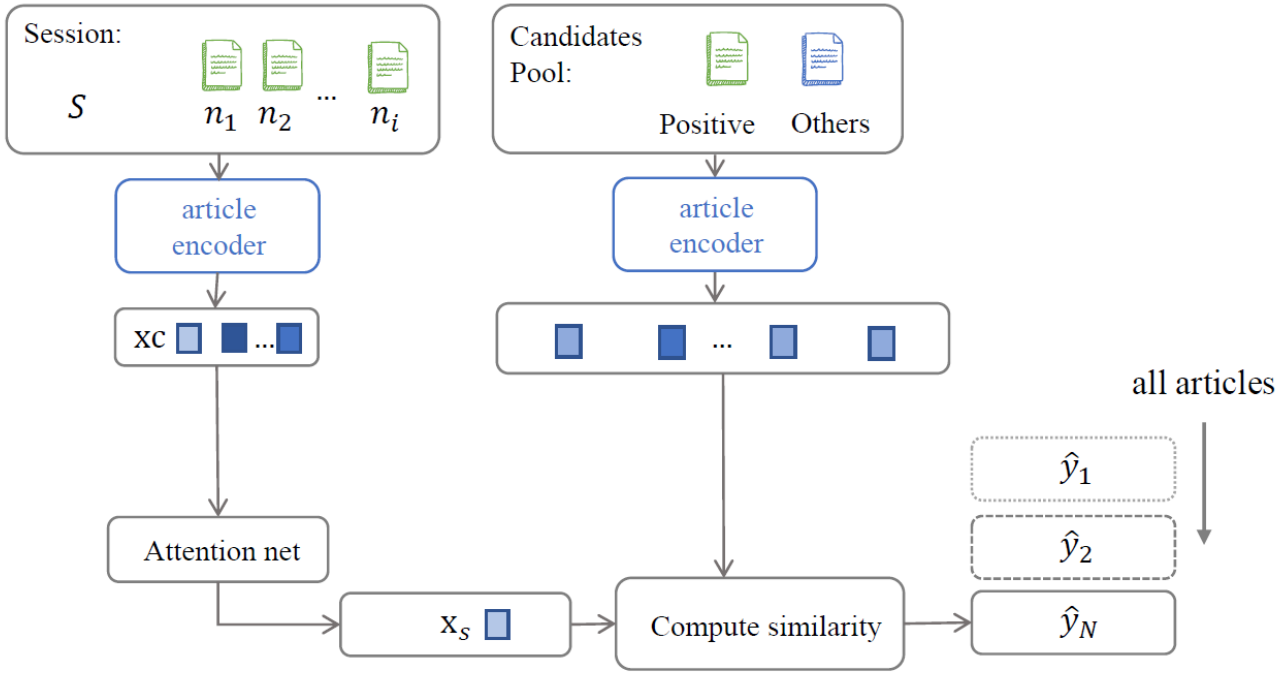


Base model

article encoder

Content representative vector $\oplus$ **co-occurrence representative vector**

Attention network: $\alpha_i = W_0 \times \sigma(W_1 \times \mathbf{xc}_i + b_0)$



Base model

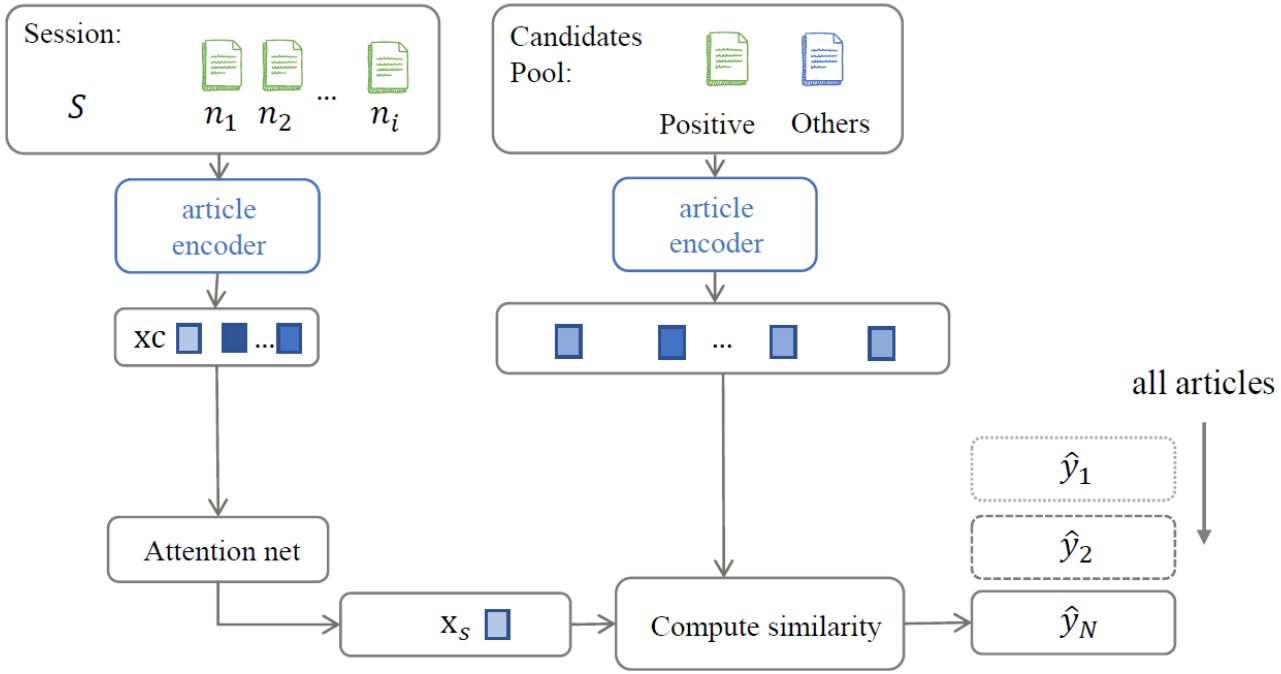
article encoder

Content representative vector $\oplus$ **co-occurrence representative vector**

Attention network:

$$\alpha_i = W_0 \times \sigma(W_1 \times \mathbf{xc}_i + b_0)$$

$$\mathcal{L}_1 = -\frac{1}{|S|}\sum_{S_u \in S}\sum_{j=1}^{N}(y_j^u \log(\hat{y}_j^u) + (1 - y_j^u)\log(1 - \hat{y}_j^u))$$

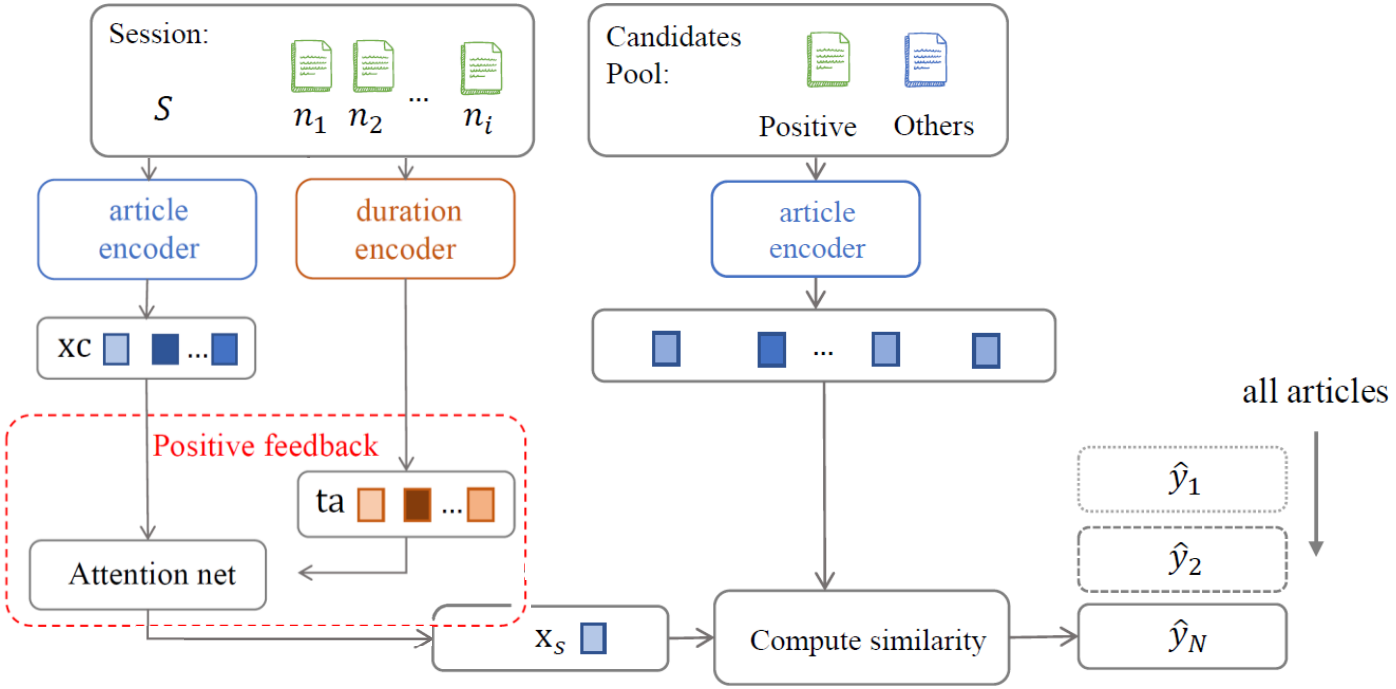


Positive implicit feedback

duration encoder

Bucketize continuous active time $ta'_i = \lfloor log_2 ta_i \rfloor$

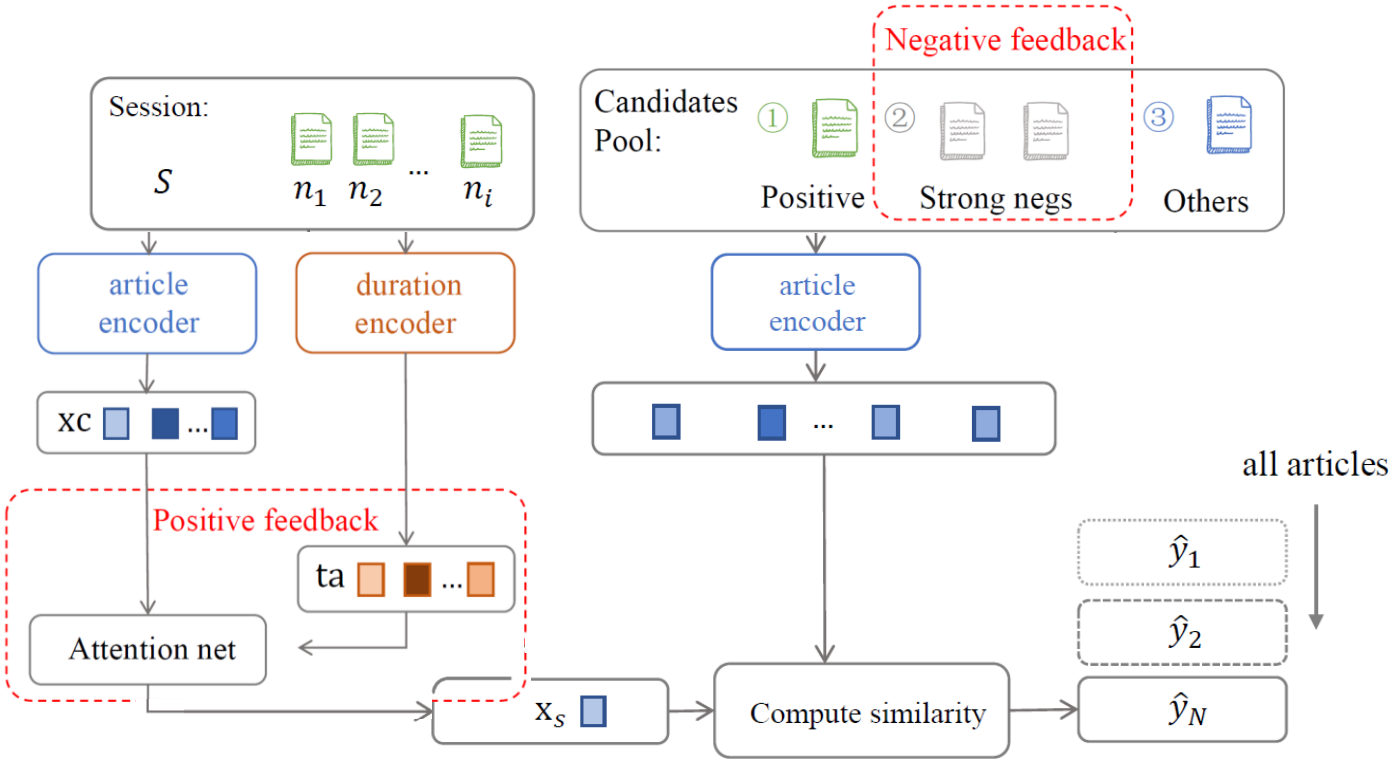
Add to attention network: $\alpha_i = W_0 \times \sigma(W_1 \times \mathbf{xc}_i + W_2 \times \mathbf{ta}_i + b_0)$



Negative implicit feedback

Modify loss function:

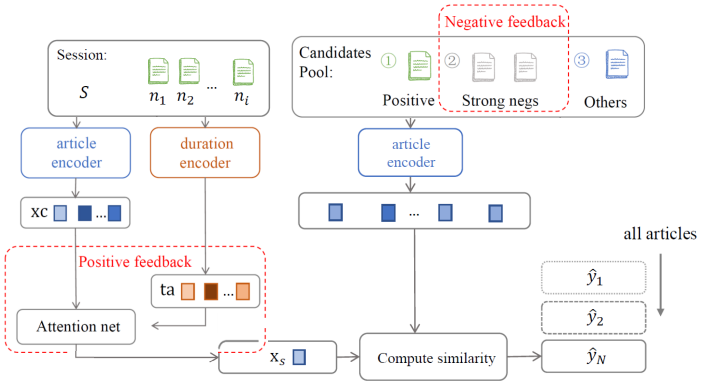
$$\mathcal{L}_2 = -\frac{1}{|S|}\sum_{S_u \in S}\sum_{j=1}^{N}(y_j^u \log(\hat{y_j^u}) + (1 - y_j^u)\log(1 - \hat{y_j^u})$$



Negative implicit feedback

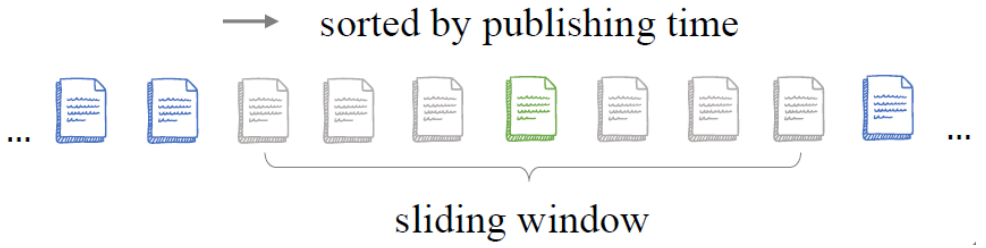
Modify loss function:

$$\mathcal{L}_2 = -\frac{1}{|S|}\sum_{S_u \in S}\sum_{j=1}^{N}(y_j^u \log(\hat{y_j^u}) + (1-y_j^u)\log(1-\hat{y_j^u}) + \lambda \mathbb{1}(j \in Ne_u)\log(\sigma(1-\mathbf{xc}_j^T\mathbf{xc}_s))),$$



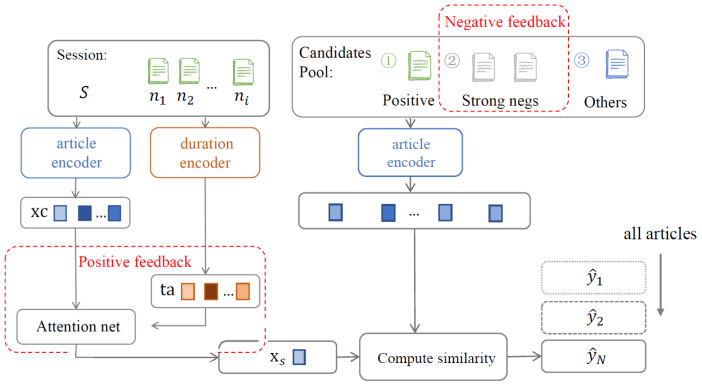
Negative implicit feedback

Negative sampling: Reconstruct impression list by publishing time



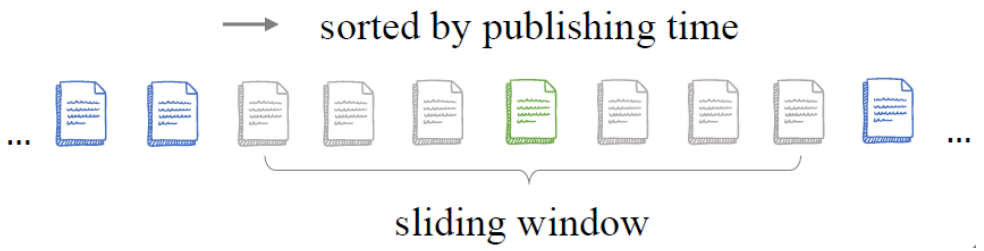
Modify loss function:

$$\mathcal{L}_2 = -\frac{1}{|S|}\sum_{S_u \in S}\sum_{j=1}^{N}(y_j^u \log(\hat{y_j^u}) + (1-y_j^u)\log(1-\hat{y_j^u}) + \lambda \mathbb{1}(j \in Ne_u)\log(\sigma(1-\mathbf{xc}_j^T\mathbf{xc}_s))),$$



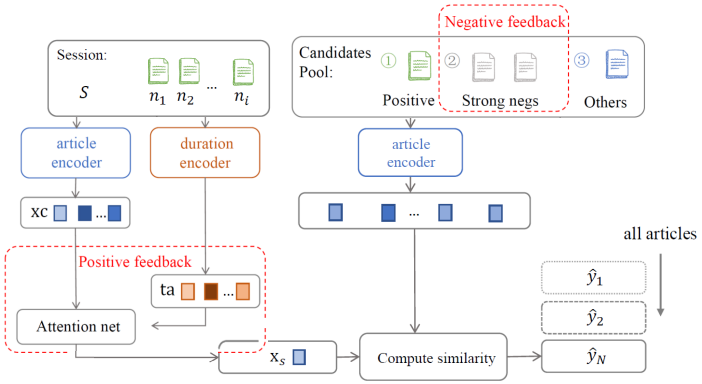
Negative implicit feedback

Negative sampling: Reconstruct impression list by publishing time



Modify loss function:

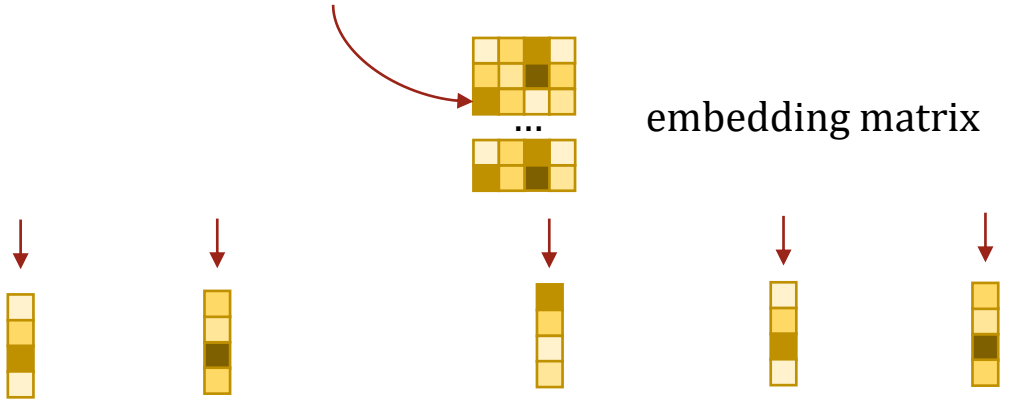
$$\mathcal{L}_2 = -\frac{1}{|S|}\sum_{S_u \in S}\sum_{j=1}^{N}(y_j^u \log(\hat{y_j^u}) + (1-y_j^u)\log(1-\hat{y_j^u}) + \lambda \mathbb{1}(j \in Ne_u)\log(\sigma(1-\mathbf{xc}_j^T\mathbf{xc}_s))),$$



Neutral feedback

temporal encoder

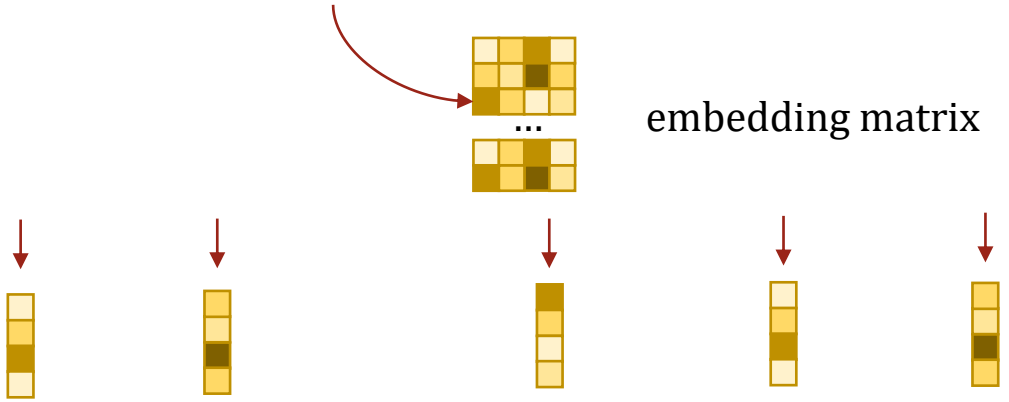
Month, day of a week, week of a month, hours, minute



Neutral feedback

temporal encoder

Month, day of a week, week of a month, hours, minute

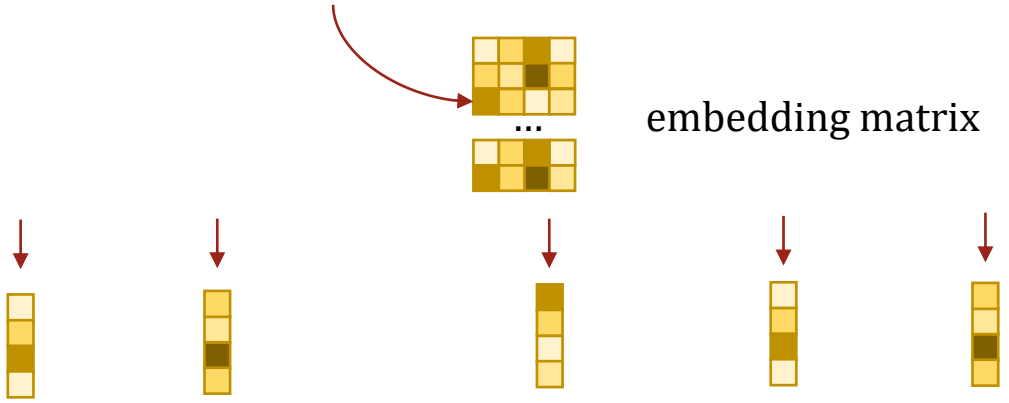


Session start time → **query to attention**

Neutral feedback

temporal encoder

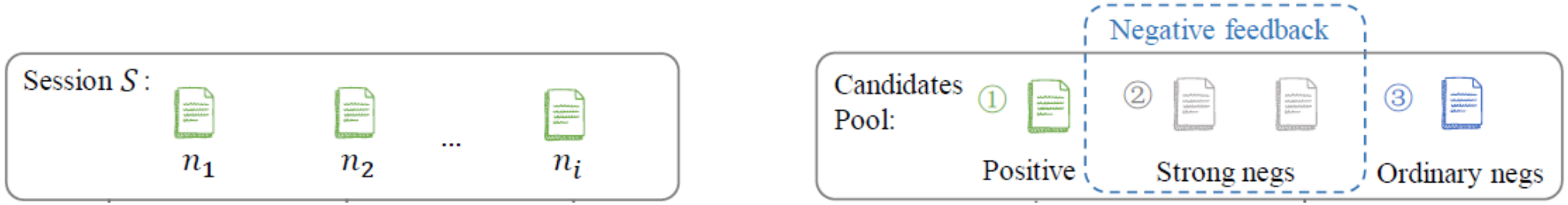
Month, day of a week, week of a month, hours, minute



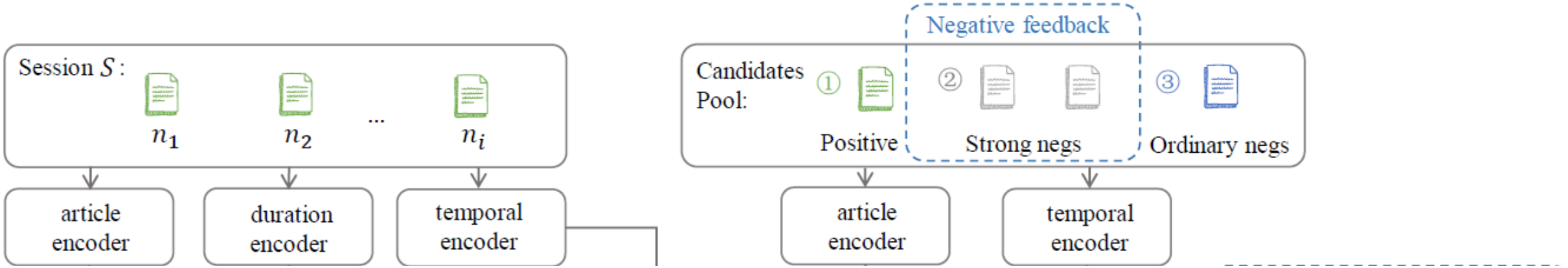
Session start time → **query to attention**

The publishing time $\mathbf{tp_1, tp_2, \dots, tp_n \rightarrow xt_S}$ ⊕ $\mathbf{xc_S \rightarrow x_s}$

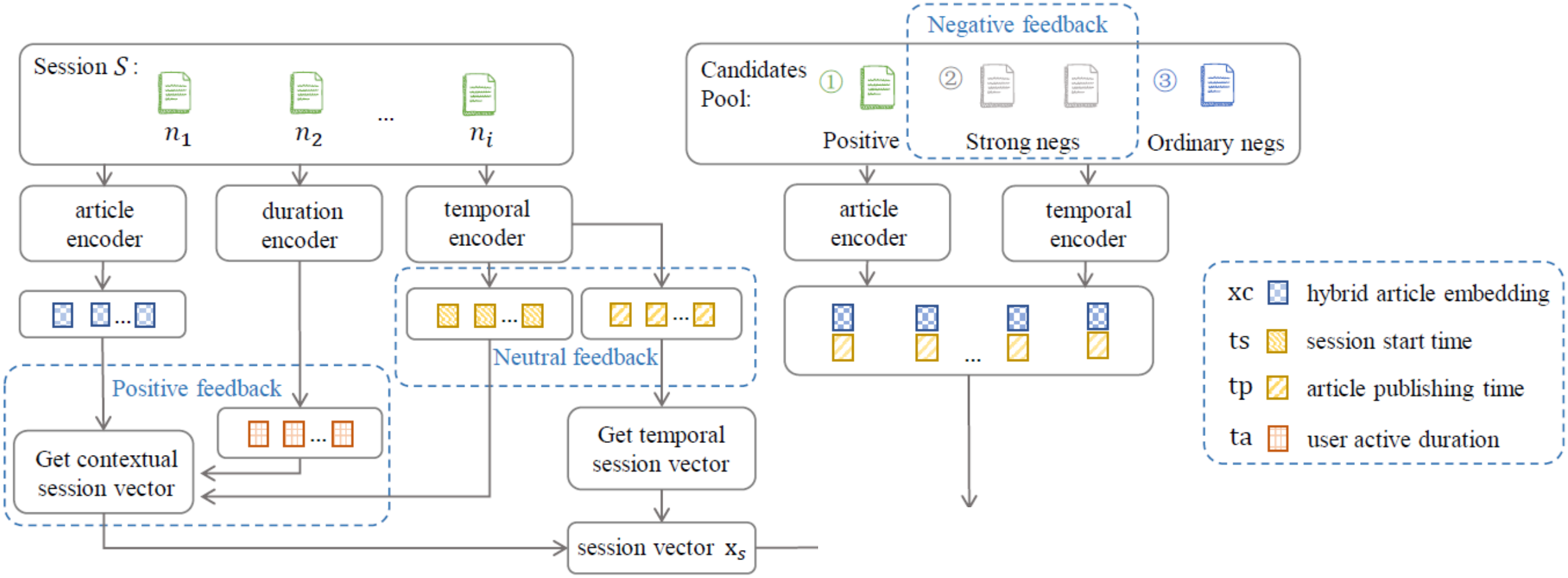
Architecture



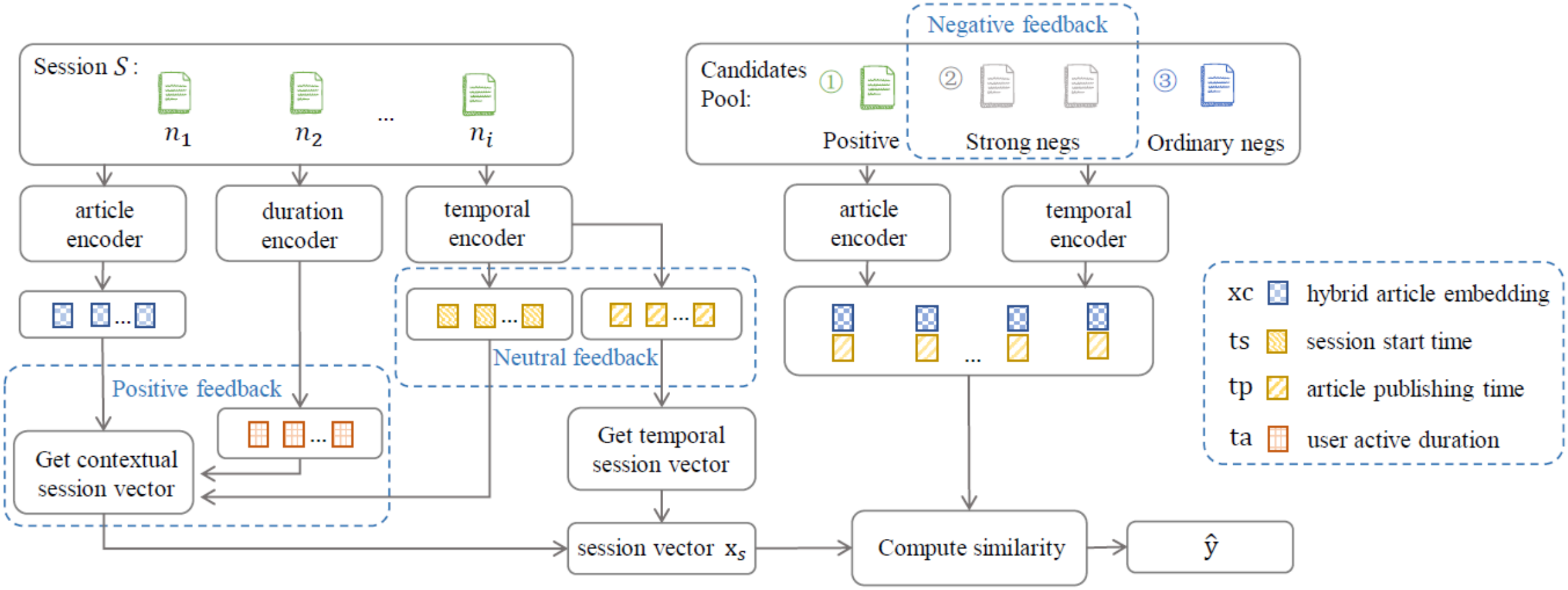
Architecture



Architecture



Architecture



03

Experiments and Analysis

Discussion about the results

Dataset

Globo.com | Adressa | MIND

Dataset	# sessions	# articles	# topics	clicks per session	clicks per article
Globo	1M	45k	461	2.69	64
Adressa	0.5M	12k	23	2.79	117
MIND	0.2M	7k	16	2.38	59

Dataset

Globo.com | Adressa | MIND

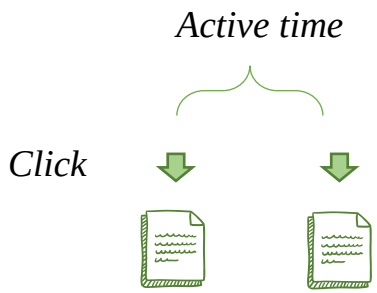
Dataset	# sessions	# articles	# topics	clicks per session	clicks per article
Globo	1M	45k	461	2.69	64
Adressa	0.5M	12k	23	2.79	117
MIND	0.2M	7k	16	2.38	59

Dataset

Globo.com | **Adressa** | **MIND**

Dataset	# sessions	# articles	# topics	clicks per session	clicks per article
Globo	1M	45k	461	2.69	64
Adressa	0.5M	12k	23	2.79	117
MIND	0.2M	7k	16	2.38	59

Dataset



Globo.com | **Adressa** | **MIND**

Dataset	# sessions	# articles	# topics	clicks per session	clicks per article
Globo	1M	45k	461	2.69	64
Adressa	0.5M	12k	23	2.79	117
MIND	0.2M	7k	16	2.38	59

Metrics & Setup

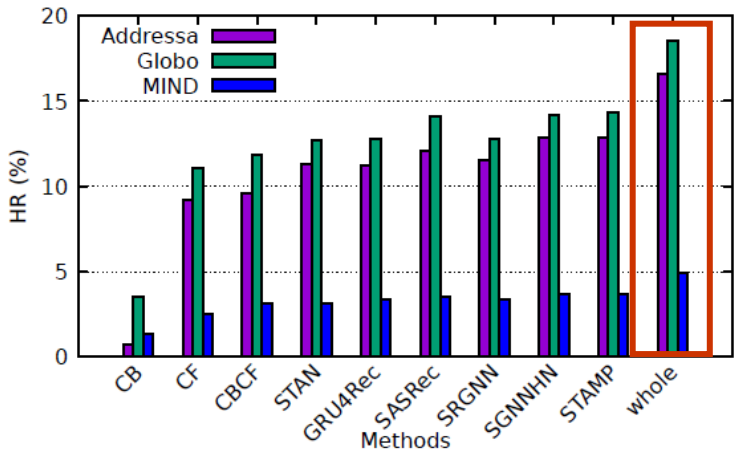
HR | **NDCG** | **ILD** | **unEXP**

Intra-list diversity | Unexpectedness

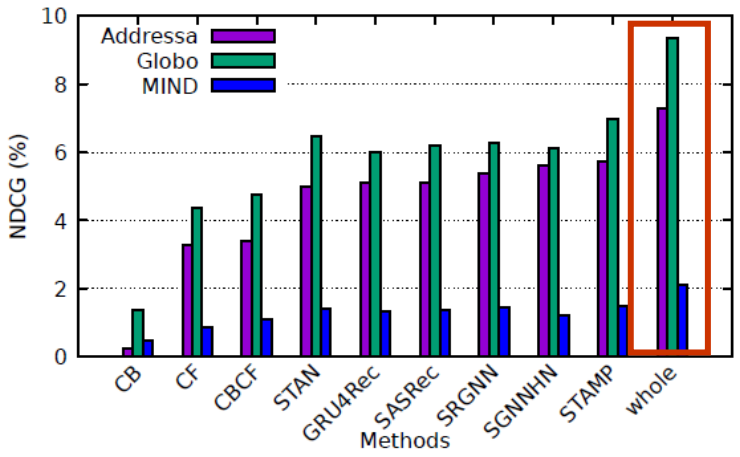
Baseline

CBCF	GRU4Rec	STAMP	CPRS
STAN	SASRec	SR-GNN	
		SGNN-HN	

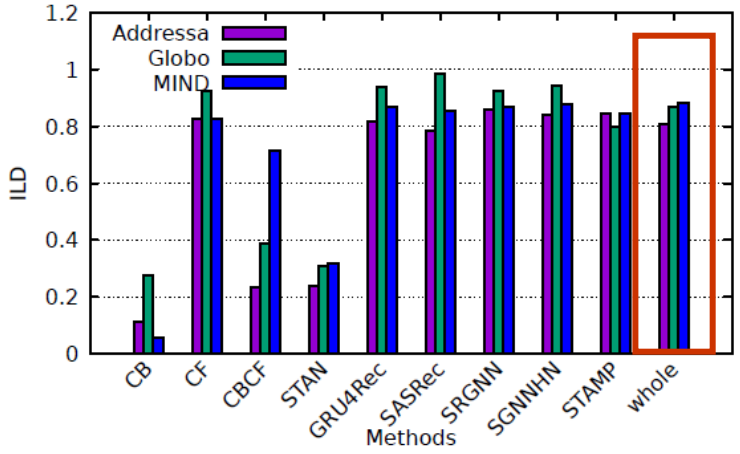
Main result



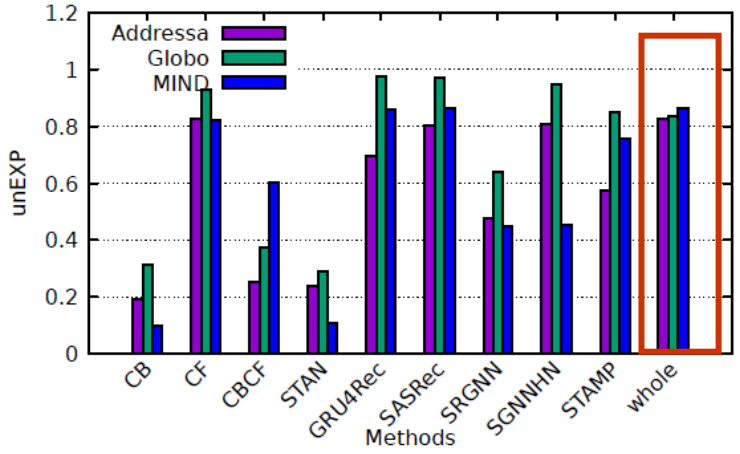
a) HR@20



b) NDCG@20

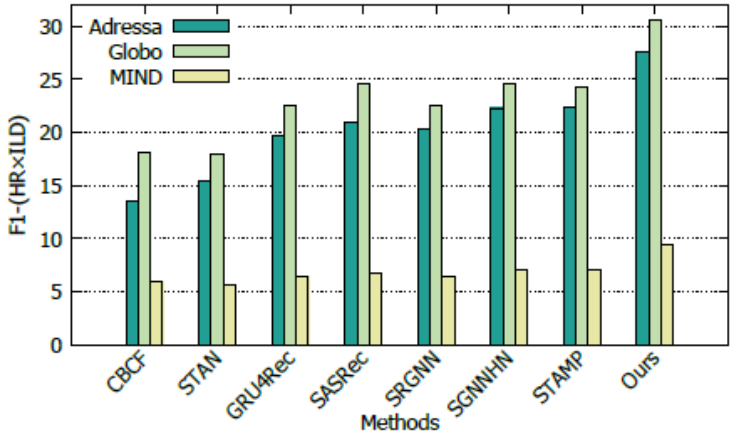


c) ILD@20

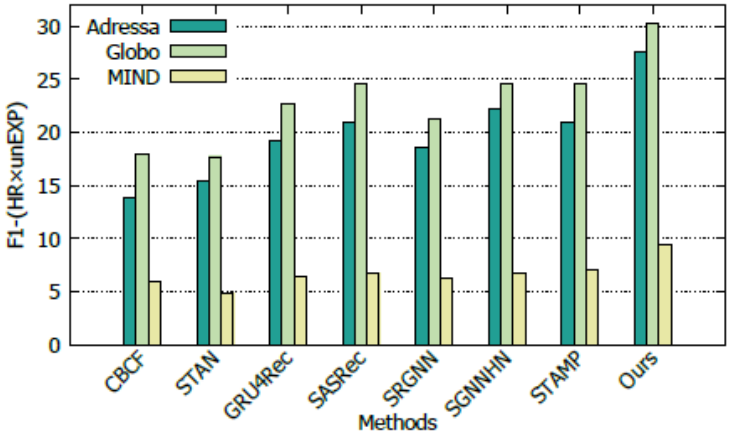


d) unEXP@20

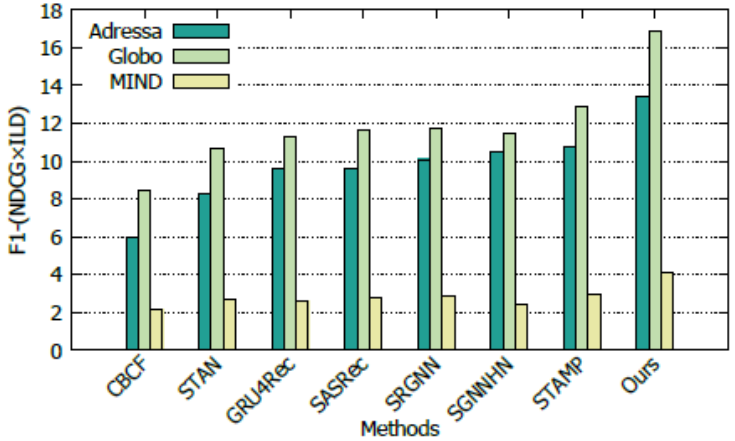
Main result



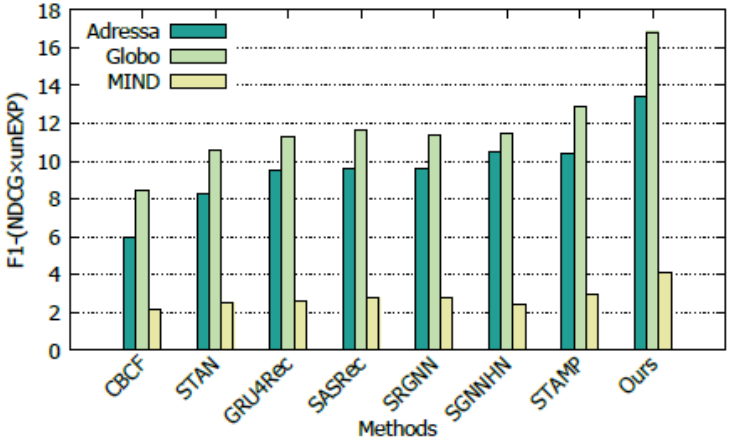
(a) F1-(HR×ILD)



(b) F1-(HR×unEXP)



(c) F1-(NDCG×ILD)



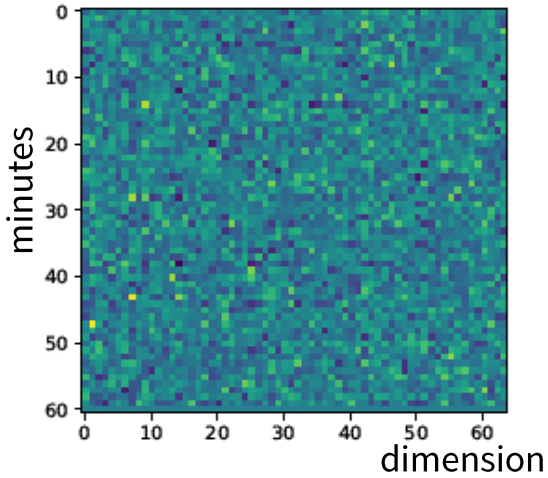
(d) F1-(NDCG×unEXP)

Embedding sharing

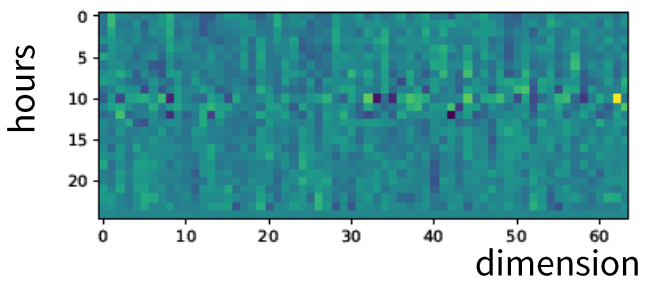
Methods	HR@20	NDCG@20	ILD@20	unEXP@20
whole (shared)	0.1658	0.0730	0.8085	0.8279
whole (no share)	0.1620 ↓	0.0727 ↓	0.8310	0.8215 ↓

Embedding sharing

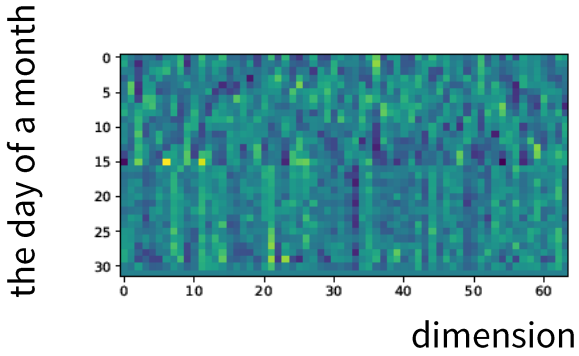
Methods	HR@20	NDCG@20	ILD@20	unEXP@20
whole (shared)	0.1658	0.0730	0.8085	0.8279
whole (no share)	0.1620 ↓	0.0727 ↓	0.8310	0.8215 ↓



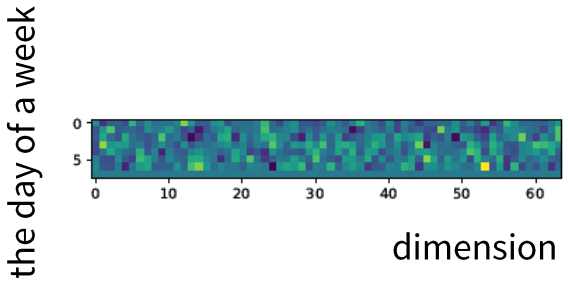
(a) Embedding of minutes



(b) Embedding of hours



(c) Embedding of days



(d) Embedding of weeks

Article cold-start

COV: article coverage,
recommend different articles to different users

Cold: all articles in a test session have not appeared in the training set

Article cold-start

COV: article coverage,
recommend different articles to different users

Cold: all articles in a test session have not appeared in the training set

Methods	Cold(80.3%)		non-Cold(19.7%)		Total	
	HR@20	COV@20	HR@20	COV@20	HR@20	COV@20
CBCF	0.0369	5.06	0.2488	5.54	0.0787	3.95
STAN	-	-	0.2652	1.63	0.0522	0.93
GRU4Rec	0.0151	0.03	0.2093	0.88	0.0533	0.50
SASRec	0.0080	0.01	0.2335	1.28	0.0525	0.73
SR-GNN	0.0100	0.01	0.2365	0.99	0.0546	0.57
STAMP	0.0172	0.01	0.2184	1.04	0.0568	0.59
SGNNHN	0.0089	0.01	0.2486	0.05	0.0561	0.04
Ours	0.0496	0.74	0.2527	1.87	0.0896	1.20
Only publishing time	0.0211	1.49	0.0803	1.55	0.0328	1.34
Ours w/o publishing time	0.0120	0.67	0.2534	1.69	0.0596	1.11

Article cold-start

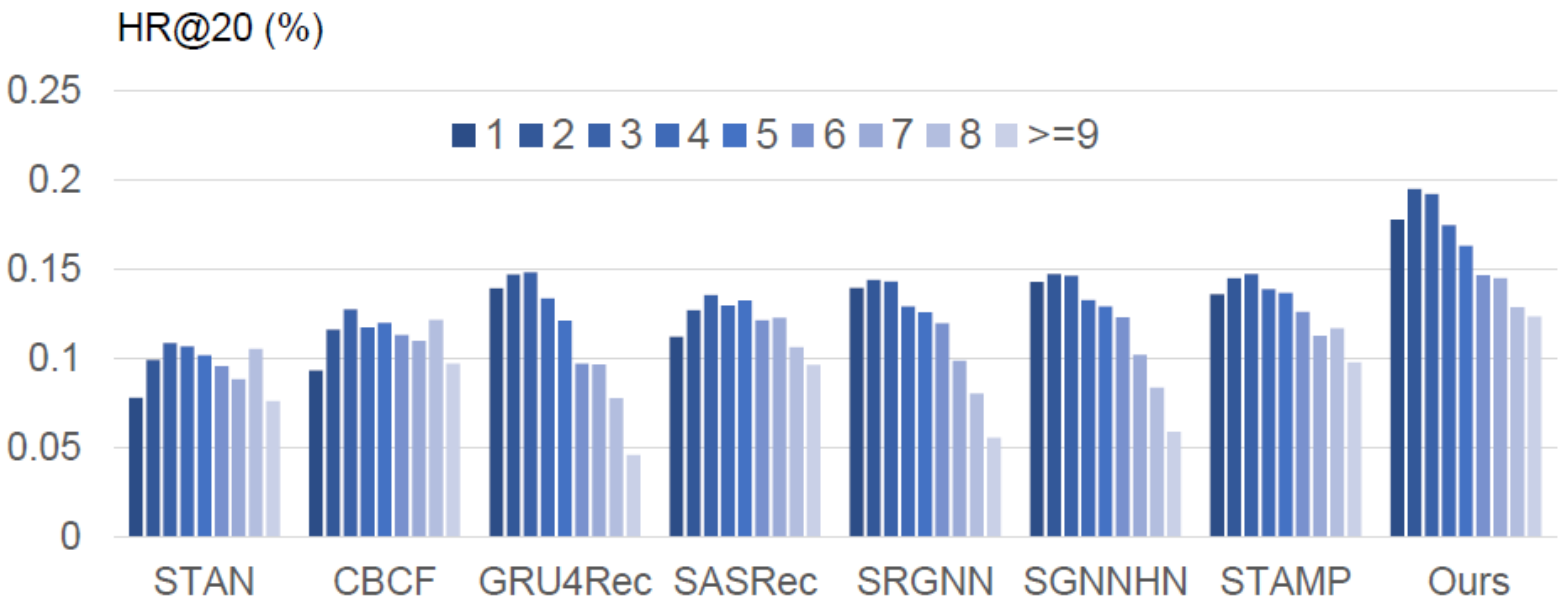
COV: article coverage,
recommend different articles to different users

Cold: all articles in a test session have not appeared in the training set

Methods	Cold(80.3%)		non-Cold(19.7%)		Total	
	HR@20	COV@20	HR@20	COV@20	HR@20	COV@20
CBCF	0.0369	5.06	0.2488	5.54	0.0787	3.95
STAN	-	-	0.2652	1.63	0.0522	0.93
GRU4Rec	0.0151	0.03	0.2093	0.88	0.0533	0.50
SASRec	0.0080	0.01	0.2335	1.28	0.0525	0.73
SR-GNN	0.0100	0.01	0.2365	0.99	0.0546	0.57
STAMP	0.0172	0.01	0.2184	1.04	0.0568	0.59
SGNNHN	0.0089	0.01	0.2486	0.05	0.0561	0.04
Ours	0.0496	0.74	0.2527	1.87	0.0896	1.20
Only publishing time	0.0211	1.49	0.0803	1.55	0.0328	1.34
Ours w/o publishing time	0.0120	0.67	0.2534	1.69	0.0596	1.11

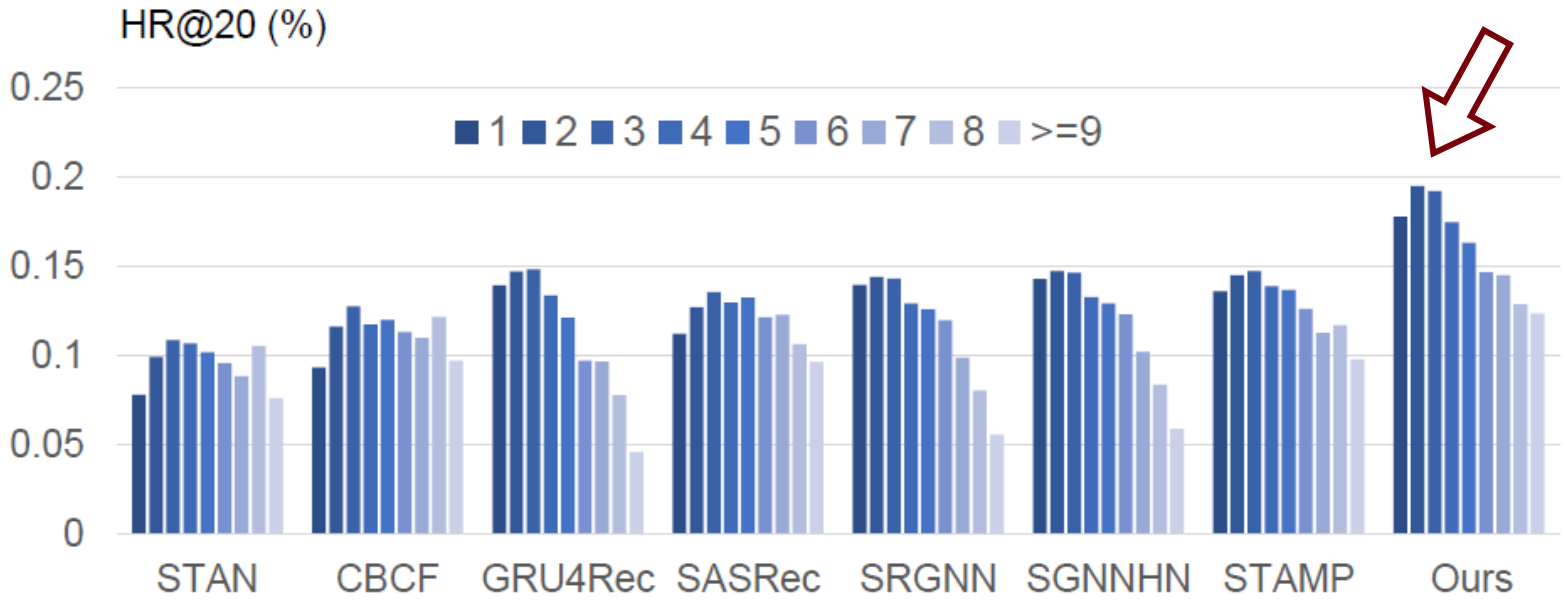
User cold-start

For different session lengths (taking Globo as an example)



User cold-start

For different session lengths (taking Globo as an example)



04

Conclusion and future work

Summary

Sparse users | Dynamic news | Unstable interests | Diversity requirement

Article transition	Find they not like	Temporal information
Session-based	Trade-off	Cold-start

Future work

Applied in Recall stage

Long term interests

Real-time application

Sophisticated architecture

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

THANKS FOR ALL

Shansan Gong

SHANGHAI JIAO TONG UNIVERSITY

Thank you for watching