International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 2, Number 3, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v2n3.04
Journal homepage: https://wepub.org/index.php/IJCSIT/index

A News Recommendation Method for User Privacy Protection

Yiqing Chen

School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454003, China

ABSTRACT

News recommendations play an important role in daily life, but there is a problem of privacy leakage. To address this problem, a news recommendation method for user privacy protection is proposed. It combines news recommendation with a federated learning framework to ensure accurate recommendations while protecting user privacy. To be specific, the method contains three key components: 1) It combines federated learning to ensure that user data is processed locally and not leaked to third parties, thus achieving privacy protection. 2) It fully considers the long-term and short-term interests of users and learns their short-term interests using the GRU model based on the item features of interaction, which provides a more comprehensive perspective on user modelling. 3) We design a score prediction method that fuses linear and nonlinear features to better capture the interaction between users and news and improve the accuracy of recommendation. The experimental validation on two public news datasets, Adressa and MIND, demonstrates that the method can still achieve efficient news recommendation without compromising user privacy.

KEYWORDS

News Recommendation; Privacy protection; Federated Learning

1. INTRODUCTION

With the widespread popularization of the Internet, the dissemination speed of online news is changing day by day. A large number of news articles are released on the Internet every day, which leads to a serious information overload problem for users of online news services [1]. Facing such a huge amount of news information, users cannot browse one by one to find the news content they are interested in, and different users have different preferences for news information. Therefore, effectively filtering and presenting personalized news information to solve the information overload problem of users is an important problem to be solved in the field of online news service. In recent years, there are numerous methods for news recommendation. When constructing news recommendation models, they are usually divided into two categories: user models and news models [2, 3]. User models are dedicated to learning and extracting users' interest representations from their historical behaviors, while news models focus on learning the intrinsic representations of news based on news content. By matching these two representations, accurate recommendation of news can be achieved. Okura et al [2] uses an autoencoder approach to learn and extract the intrinsic representation of news from news content. Wang et al [3] proposes a candidate-aware attention network to provide an effective way for the user model, which helps to capture the user's interest in the candidate news. Wu et al [4] uses a multi-head self-attention network for in-depth modeling and analysis of news headlines. It is worth mentioning that Wu et al [5] innovatively introduces a pre-trained language model into the news modeling domain, which provides strong support for improving the accuracy and effectiveness of news recommendation.

In the news recommendation domain, one of the core challenges is accurately capturing users' interests and provide them with personalized news content. Traditional news recommendation methods usually store and train user behavior data centrally. However, this approach carries the risk of data leakage, which in turn may expose users' private information

Federated learning is a distributed training framework that allows multiple clients to jointly train learning models without sending raw data to a central server, enabling multiple clients to collaboratively learn global models without sharing local data. Combining news recommendation and federated learning to train accurate news recommendation models using useful information from massive user behaviors while eliminating the need for centralized storage of them is one of the research hotspots in news recommendation. Wu et al [6] proposes a personalized news recommendation method that protects users' privacy, where a local copy of the news recommendation model is retained on each user's device and, based on the users' behaviors in that device behavior, the gradient of the local model is calculated and uploaded to the server to update the global model. However, this method requires a huge computational and communication cost from the client, which is unacceptable to the user. Wu et al [7] separates the user model and the news model, and trains only highly private and small data volume in the client, and privacy-insensitive and large data volume in the server, which also protects the user's privacy as well as reduces the computational cost of the client.

When modelling users, we need to consider that online users have diverse interests in news. Some users may have a sustained interest in particular people or events for a long time. For example, if a user is passionate about table tennis, that user may read many various news related to table tennis over the next period of time. This user preference is called long-term interest. In addition, many user interests may change over time and may be triggered by specific contextual or temporal demands. For example, reading the news about the ‘2023 Nobel Prize in Literature’ will cause the user to read frequently related news such as ‘The Collected Works of Jon Fosse’ because John Fosse is the winner of the Nobel Prize in Literature for that year, although the user may have never read the news before. This user interest is called short-term interest. Thus, users usually have both long-term interests and short-term interests, and learning users' long-term and short-term interests can better recommend satisfactory news for them. However, most of the existing news recommendation methods learn a single long-term representation of the user, and less often take into account the short-term interests of user interactions.

The main task of news recommendation is to provide users with a set of personalized news rankings to satisfy their news reading needs. Traditional methods mostly rank candidate news articles using a direct dot product of news and user interest representations, which may ignore some potential relevance. For example, Okura et al [2] uses the dot product between news and user representations to predict relevance scores. Iana et al [8] uses multiple dot product aggregations to model clicked news and candidate news.

To address the privacy leakage problem and enhance news recommendation, we propose a news recommendation method for user privacy protection that combines the user's long-term and short-term interest representations and utilizes linear and non-linear fusion for news recommendation. Specifically, we focus on the user's recent interaction history and learn the user's long- and short-term interest representations from the user's historical browsing history via a news encoder. When ranking candidate news items, we utilize a multilayer perceptron (MLP) nonlinear model to learn the interaction function between users and news. Experimental results on the publicly available datasets Adressa and MIND show that our method achieves significant performance improvements in all four evaluation metrics, AUC, MRR, nDCG@5 and nDCG@10, compared to the current optimal baseline method.

In summary, the contributions of this paper are as follows:

When modelling users, we consider the problem that users' interests change over time, use GRU networks to capture users' short-term interests, and combine users' long-term interests with short-term interests to obtain more accurate user representations.

When personalizing rankings, we use a combination of linear and non-linear methods to get predictive scores for ranking.

Experiments are conducted on two real and publicly available datasets, and the results show that the method proposed in this paper can effectively improve the recommendation metrics.

2. RELATED WORK

2.1. News Recommendation

The news recommendation model mainly consists of a news model and a user model. News modelling is a key step in personalized news recommendation methods to capture the characteristics of news articles and understand their content. News modelling mostly uses news ID to represent news articles [9]. However, simple ID cannot model news representations well, and many approaches incorporate news content for news modelling. Gershman et al [10] considers Term Frequency Inverse Document Frequency (TF-IDF) features extracted from news texts. Capelle et al [11] proposes to use Synset Frequency based on Synset Frequency in WordNet -Inverse Document Frequency (SF-IDF) to model news. Some scholars have also incorporated topic categories and labels of news into news modelling. For example, Zhang et al [12] learns news representations from news ID, categories, keywords and entities via character-level CNNs. TANR et al [13] incorporates an assisted news topic prediction task to learn topic-aware news representations. Other scholars combine news recommendation with graph neural networks to enhance candidate news representations. Yang et al [14] enhances candidate news representations by constructing a global click-through news encoder using global entity graph and candidate news fusion. User modelling is a key step in personalized news recommender systems to infer a user's personal interest in news. User modelling is user-click based news modelling which introduces additional user functionality for better personalized user understanding. Wu et al [15] and Liu et al [16] use news-level attention networks to learn user representations from clicked news representations. Zhang et al [17] also uses attention networks to aggregate different information from clicked news and candidate news for user modelling. Wu et al [18] uses cross-modal candidate-aware attention network which selects click news based on cross-modal correlation between click news and candidate news. Wu et al [19] proposes a temporal diversity-aware news recommendation method to predict clicks more accurately.

2.2. Federated Learning

Federated learning [20] is a privacy-preserving machine learning technique. There are several approaches combining federated learning and recommendation models. Ammad et al [21] proposes a Federated Collaborative Filtering (FCF) approach, which stores the user matrix locally on the client side and the item matrix on the server side. Wu et al [22] proposes the FedGNN model, which combines a federated learning framework with a graph neural network, where each user stores a local user-item graph locally that finds neighboring users with common interactions and exchanges their embeddings to extend the local user-item graph. Chen et al [23] uses a federated learning framework to improve the recommendation model by considering the knowledge migration of the data. Li et al [24] proposes a federated sequential recommendation method using a low-rank tensor projection to model the user's dynamic preferences. Liu et al [25] decomposes the user interest vector into a low-dimensional attention vectors and designs a random anonymous behavior vector filling method to reduce noise and improve usability. Combining news recommendation and federated learning is also one of the research hotspots in news recommendation. Wu et al [6] proposes a personalized news recommendation method that protects users' privacy by keeping the news recommendation model's locally and calculating the gradient of the local model based on the user's behavior in the device and uploading it to the server for updating the global model. Wu et al [7] trains the data with high privacy and small data volume on the client side, and the data with privacy-insensitive and large data volume on the client side. privacy-insensitive and high data volume data is trained in the server, which reduces the computational cost of the client while protecting the user's privacy. Yi et al [26] proposes an untargeted attack on federated news recommendation, which effectively reduces the model performance of a small group of malicious clients by exploiting the a priori knowledge of news recommendation and federated learning

3. OUR METHODOLOGY

3.1. News recommendation model

The news recommendation model consists of three main parts: news model, user model and prediction score. The model graph is shown in Fig. 1.

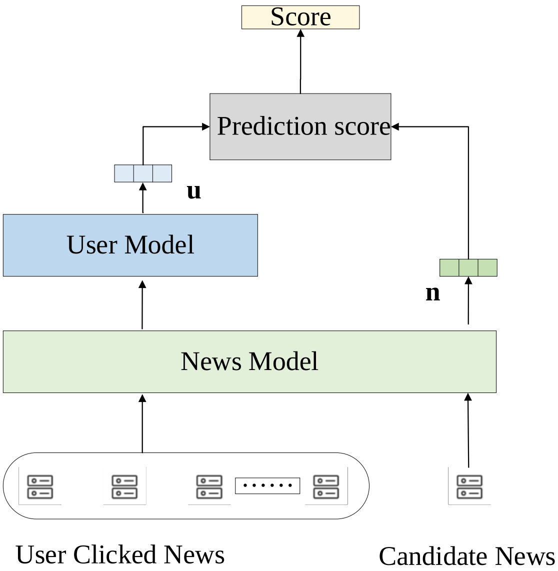

Fig. 1 News recommendation model

3.1.1. News Models

For a given news item containing information such as headline, summary, category, text, etc., the news model takes one or more of these elements as inputs and learns news representations from them through an encoder. We use the PLM-NR news model [5] to learn news representations from news headlines. The model contains four layers. Firstly, there is a word embedding layer, which converts the words in the news headlines into semantic word embedding vectors. The second layer learns word representations by capturing the local context through a CNN network layer. The third layer learns contextual word representations by modelling the correlation between different words with a multi-head self-attention network. The last layer is the attention network layer, which constructs the news representation vector n from the output of the multi-head self-attention network by selecting informative words.

3.1.2. User Models

The user's long-term inherent interests are shown in the user's historical clicking behavior. The user's historical click sequence is Lu=[n1, n2,…nm] and m is the length of the user's click sequence. We use a multi-head self-attention network to capture the relationship between historical click news, and then weight it using an additive attention network to obtain the long-term user representation ul.

User interest is dynamically changing, in selecting the news of recent interactions of user interactions to obtain the short-term user interest, the short-term click sequence of the user is Su=[nm-k+1,…nm-1,nm], where k is the length of the sequence.

The representation of these clicked news is obtained as [em-k+1,…em-1,em] using news encoder. We use GRU for feature extraction and use the hidden state representation of the last output as the short-term user representation us=ht. The generic formula for GRU is:

$$\boldsymbol{r}_t = \sigma\left(\boldsymbol{W}_r\left[\boldsymbol{h}_{t-1}, \boldsymbol{e}_t\right]\right) \tag{1}$$

$$\boldsymbol{z}_t = \sigma\left(\boldsymbol{W}_z\left[\boldsymbol{h}_{t-1}, \boldsymbol{e}_t\right]\right) \tag{2}$$

$$\tilde{\boldsymbol{h}}_t = \tanh\left(\boldsymbol{W}_{\tilde{h}}\left[\boldsymbol{r}_t \odot \boldsymbol{h}_{t-1}, \boldsymbol{e}_t\right]\right) \tag{3}$$

$$\boldsymbol{h}_t = \boldsymbol{z}_t \odot \boldsymbol{h}_t + \left(\mathbf{1} - \boldsymbol{z}_t\right) \odot \tilde{\boldsymbol{h}}_t \tag{4}$$

Finally, the user representation u is obtained by combining the long-term user interest ul and the short-term user interest us. We use additive attention network considering that users attach different importance to long-term and short-term interests. The user model graph is shown in Fig. 2.

$$u = u_l \oplus u_s \tag{5}$$

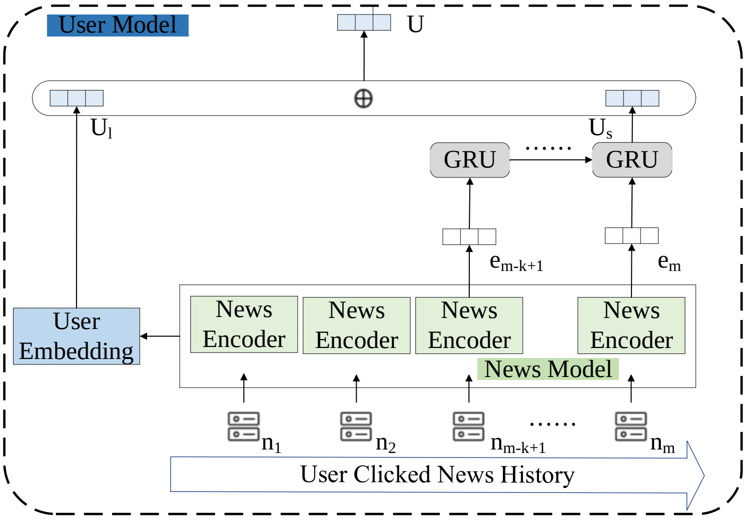

Fig. 2 User model

3.1.3. Prediction Score

The purpose of prediction scoring is to take the candidate news representation and the user representation obtained from the user encoder as inputs in order to obtain the user's score for the candidate news. We use a combination of linear and non-linear methods so that they share the same embedding layer, and then combine their outputs for computation to obtain the final score, as shown in Fig. 3.

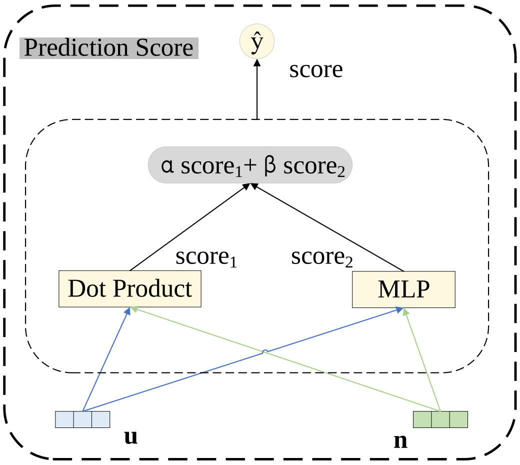

Fig. 3 Prediction score

Based on the matrix decomposition, the user embedding and the candidate news embedding are dot-producted, and then the vectors are projected to the output layer to obtain the prediction score score1.

$$score_1 = a_{out}\left(\mathbf{h}^T\left(\mathbf{p}_u \odot \mathbf{q}_n\right)\right) \tag{6}$$

The prediction score score1 is obtained by projecting the vectors to the output layer. Considering that simple vector dot product is not enough to model the complex interaction between user and news potential features, MLP is used to learn the interaction between them to obtain the prediction score score2. where W, b are trainable parameters, and the activation function is chosen as Relu function.

$$score_2 = a_L\left(\mathbf{W}_L^T\left(a_{L-1}\left(\ldots a_2\left(\mathbf{W}_2^T\begin{bmatrix}\mathbf{p}_u^M\\ \mathbf{q}_i^M\end{bmatrix}+\mathbf{b}_2\right)\ldots\right)\right)+\mathbf{b}_L\right) \tag{7}$$

The final prediction score is:

$$\hat{y} = \alpha score_1 + \beta \mathrm{score}_2 \tag{8}$$

Where, αandβare adjustable parameters.

3.2. Framework of Federal News Recommendation

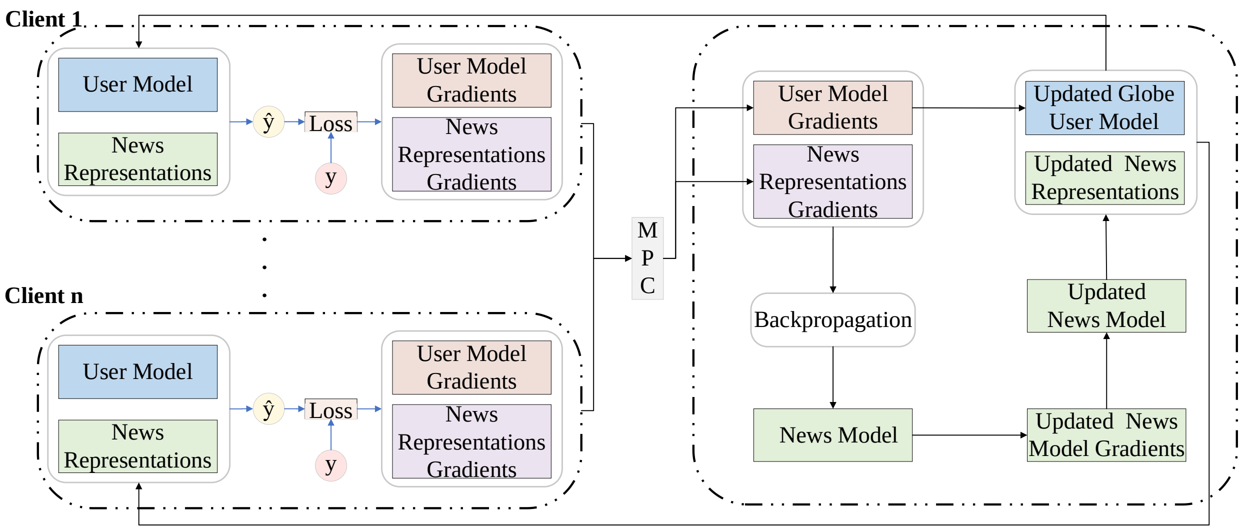

Fig. 4 The framework of federal news recommendation

The general framework of the federated news recommendation framework is shown in Fig. 4. Existing news recommendation methods need to collect users' browsing history to model users, and most of them need to store users' browsing history centrally in the server, which can easily lead to privacy leakage problems. In federated learning, private data is stored locally on different clients. With the introduction of the federated learning framework for news recommendation, user data does not need to be uploaded to the server side, which prevents data leakage and effectively protects user privacy. Users put their browsing history on the local client, train the model locally, and upload the local loss gradient to the server. Then, the server aggregates the local loss gradients, performs global model update, and sends the updated user model and news representation model to each client. Iterate until the model converges.

Since training a news model usually requires a large computational overhead, it would be unacceptable for users if the training is performed entirely on the client side. Considering the communication and computational overhead of clients, the news recommendation model is decomposed into a large news model maintained on the server and a lightweight user model shared between the server and clients. News models that are prone to leaking user privacy are trained in the client, and only news models that do not pose a privacy risk are trained in the server, and users only need to request a small number of news representations from the server instead of requesting the entire news model, which protects user privacy and reduces the client's overhead.

4. EXPERIMENTS

4.1. Datasets

We conducted experiments on two publicly available datasets, Adressa and MIND.MIND is a large news research dataset published by Microsoft containing 6 weeks of data with news, items, impressions, and positive and negative samples. Adressa is a publicly available news dataset from a Norwegian newspaper company containing 7 days of sample data. The first 5 days of news click data is used as the user's history of clicks, the click data from day 6 is used as a training set, and the click data from day 7 is used as a validation set and a test set, where 20% is the validation set and 80% is the test set. Since this dataset is not distinguished from impressions and has no negative samples, for each news item clicked by a user, 20 news items are randomly selected as negative samples, and these 21 news items are considered as impressions for subsequent experiments.

The detailed information is shown in Table 1.

Table. 1 Statistics of datasets.

	Adressa	MIND
News	20428	65238
Users	640503	94057
Impressions	-	230117
Positive samples	3101991	347727
Negative samples	-	8236715

4.2. Experimental setting and evaluation metrics

We used [27] as a pre-trained language model for MIND and [28] as a pre-trained language model for Adressa. The dimensionality of the news representation was set to 400. to reduce overfitting, we set the discard rate in the user model to 0.2. the learning rate to 0.00005. the number of negative samples associated with each positive sample was 4. the user group size was 50. the retained user history browsing records were 50. all hyperparameters were selected based on the results on the validation set. Repeat each experiment independently five times for each experiment and calculate the average result. We measure the recommendation effect by four evaluation metrics: the AUC, MRR, nDCG@5 and nDCG@10.

4.3. Baseline

(1) DFM [29] is a deep fusion model that fuses different depth fully connected layers for news recommendation.

(2) DKN [3] is a knowledge-aware news recommendation method.

(3) NAML [15] uses a multi-view attention mechanism for news representation, which achieves more accurate news representation through multiple features of the news dataset.

(4) NRMS [4] uses a self-attention mechanism to represent news features and user features, and uses a sequence of news words and a sequence of users browsing news as inputs, respectively.

(5) FCF [21] uses federated collaborative filtering approach for news recommendation.

(6) FedRec [6] applies federated learning and local differential privacy for news recommendation.

(7) Efficient-FedRec [7] is an efficient privacy-preserving federated news recommendation method.

4.4. Performance Comparison

This section experimentally compares the performance of our proposed method with other baseline methods. Firstly some centralised storage news recommendation algorithms are compared with some federated news recommendation methods. Among them, DKM, NAML, and NRMS are centralised news recommendation methods, and FCF, FedRec and Efficient-FedRec are federal news recommendation methods. The recommendation performance of different recommendation methods is shown in Table 2, where the bolded data are the optimal results and the data with underline are the sub-optimal results. It can be observed from Table 2:

Centralised news recommendation methods (DKM, NAML, NRMS) have much higher recommendation performance than the traditional deep learning method DFM. This is because they apply more complex models to characterise users and news, and neural networks can capture global and local semantic context in news, which helps to learn more accurate user representations of news and news recommendations.

Federated recommendation methods (FCF, FedRec, and Efficient-FedRec) do not require uploading the user's news browsing history, which can effectively protect user privacy. Among them, FCF has the lowest recommendation performance because there is a cold-start problem using collaborative filtering, and each user and item must participate in the training process. However, FCF cannot handle new items and is not applicable to news recommendation due to the fast news update in news systems.

Compared with traditional methods, our approach provides a significant improvement in recommendation performance while protecting user privacy. We add a GRU sequence network to the user model to capture users' short-term interests and connect between users' short-term and long-term interests to obtain new user vectors. Also, a combination of linear and non-linear approach is used to compute the score of user and news matches for personalized ranking, which is able to model the potential relevance and get better results.

Table. 2 Overall performance

	Adressa				MIND			
	AUC	MRR	nDCG@5	nDCG@10	AUC	MRR	nDCG@5	nDCG@10
DFM	59.90	32.68	29.69	36.43	60.67	28.08	29.93	35.68
DKM	73.73	39.52	40.98	47.48	64.72	30.53	33.01	38.70
NAML	73.09	44.27	43.51	50.02	66.10	31.91	34.52	40.21
NRMS	75.31	42.24	44.66	48.46	66.67	32.25	49.99	40.74
FCF	51.39	18.98	15.42	22.94	50.02	22.37	22.77	29.02
FedRec	71.73	41.37	41.81	47.18	66.54	31.96	34.54	40.30
Efficient-FedRec	79.08	45.09	47.13	53.85	67.25	32.63	35.37	41.11
ours	80.63	47.67	49.88	56.27	67.52	32.90	35.61	41.32

4.5. Ablation Experiment

4.5.1. Effect of short-term user interest

In this section, ablation experiments are conducted to verify the impact of short-term user interest representation on recommendation performance, and the experimental results are shown in Fig. 5. We can see that the performance metrics of the news recommendation model are improved after modelling users' short-term interests through GRU. The reason is that the user's long-term interest can be captured by the self-attention network. At this point, the user representations obtained from news sequences of different order may be the same, because self-attention itself is difficult to capture the sequence order information. However, the GRU sequence network can capture the sequential information of news, and the user's recent browsing history can be used as input to obtain the user's short-term representation. Our combination of long-term and short-term user interests can capture the complex and diverse user interests in news reading.

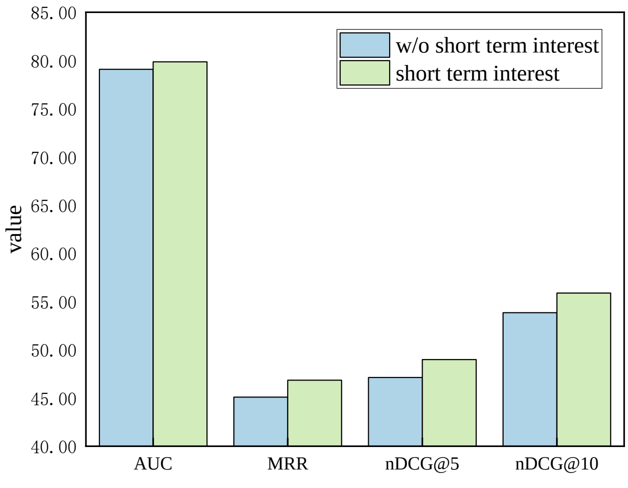

Fig.5 The influence of short-term interest

4.5.2. Impact of combining non-linear scoring

In this section, ablation experiments are conducted to verify the impact of combining linear and non-linear score calculation approaches for personalised ranking on recommendation performance, the results of which are shown in Fig. 6. We can see that fusing linear and non-linear approaches improves the performance of the model. This is because modelling user-news in a linear way using dot product alone is not sufficient to capture their complex relationship. Whereas MLP neural network has strong nonlinear modelling capability and can handle complex data relationships for better recommendation performance.

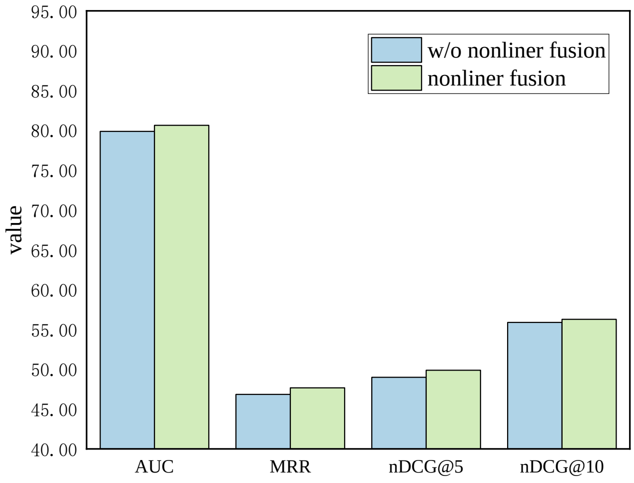

Fig.6 The influence of nonlinear fusion

4.6. Effect of hyperparameters

To verify the effect of the length of the browsed news sequences in the short-term user representation on the news recommendation performance, experiments are conducted in this section and the results are shown in Fig. 7. The length of the sequence is the number of recently browsed news that the user selects when modelling short-term interest in the user, and these news representations capture the user's short-term interest through the GRU sequence network. The total user browsing history in the experiment is 50, and the number of news short-term browsing sequence lengths of the user is set to k. The performance of news recommendation is optimal when k is set to 25. When k is set too small, the number of news samples is insufficient to capture the user's interest and the indicated user embedding is not as accurate. When k is set too large, the user's short-term interest representation tends to coincide with the long-term interest representation and the user's recent interests are not available.

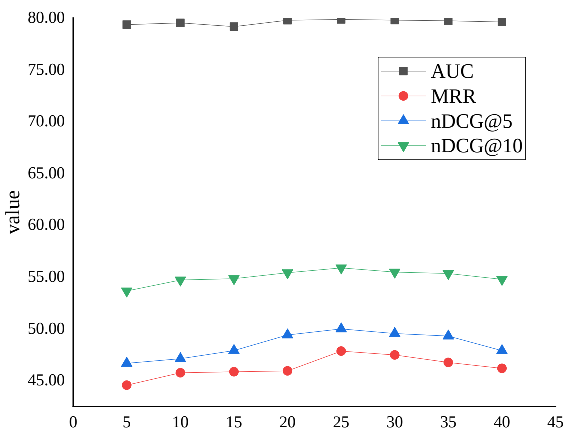

Fig.7 The influence of the number of news items

5. CONCLUSION

In this paper, a news recommendation method for user privacy protection is proposed for the user privacy problem in news recommendation. Firstly, the short and long term interests of users are considered and captured using GRU network. Then for personalized ranking, a combination of linear and non-linear methods are used to obtain predicted scores for ranking. Experimental results on two publicly available real-world datasets, Adressa and MIND, show that the method proposed in this paper can effectively improve the recommendation metrics.

REFERENCES

[1] ZHENG G, ZHANG F, ZHENG Z, et al. DRN: A deep reinforcement learning framework for news recommendation [C]// proceedings of the Proceedings of the 2018 world wide web conference, 2018: 167-176.

[2] OKURA S, TAGAMI Y, ONO S, et al. Embedding-based news recommendation for millions of users [C]// proceedings of the Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, 2017: 1933-1942.

[3] WANG H, ZHANG F, XIE X, et al. DKN: Deep knowledge-aware network for news recommendation [C]// proceedings of the Proceedings of the 2018 world wide web conference, 2018: 1835-1844.

[4] WU C, WU F, GE S, et al. Neural news recommendation with multi-head self-attention [C]// proceedings of the Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP), 2019: 6389-6394.

[5] WU C, WU F, QI T, et al. Empowering news recommendation with pre-trained language models [C]// proceedings of the Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2021: 1652-1656.

[6] QI T, WU F, WU C, et al. Privacy-preserving news recommendation model learning [J]. arXiv preprint arXiv:200309592, 2020.

[7] YI J, WU F, WU C, et al. Efficient-FedRec: Efficient federated learning framework for privacy-preserving news recommendation [J]. arXiv preprint arXiv:210905446, 2021.

[8] IANA A, GLAVAS G, PAULHEIM H. Simplifying content-based neural news recommendation: On user modeling and training objectives [C]// proceedings of the Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2023: 2384-2388.

[9] MOOKIAH L, EBERLE W, MONDAL M. Personalized news recommendation using graph-based approach [J]. Intelligent Data Analysis, 2018, 22(4): 881-909.

[10] GERSHMAN A, WOLFE T, FINK E, et al. News personalization using support vector machines [J]. 2011.

[11] CAPELLE M, FRASINCAR F, MOERLAND M, et al. Semantics-based news recommendation [C]// proceedings of the Proceedings of the 2nd international conference on web intelligence, mining and semantics, 2012: 1-9.

[12] ZHANG L, LIU P, GULLA J A. A deep joint network for session-based news recommendations with contextual augmentation [M]. Proceedings of the 29th on Hypertext and Social Media. 2018: 201-209.

[13] WU C, WU F, AN M, et al. Neural news recommendation with topic-aware news representation [C]// proceedings of the Proceedings of the 57th Annual meeting of the association for computational linguistics, 2019: 1154-1159.

[14] YANG B, LIU D, SUZUMURA T, et al. Going Beyond Local: Global Graph-Enhanced Personalized News Recommendations [C]// proceedings of the Proceedings of the 17th ACM Conference on Recommender Systems, 2023: 24-34.

[15] WU C, WU F, AN M, et al. Neural news recommendation with attentive multi-view learning [J]. arXiv preprint arXiv:190705576, 2019.

[16] LIU D, LIAN J, WANG S, et al. KRED: Knowledge-aware document representation for news recommendations [C]// proceedings of the Proceedings of the 14th ACM Conference on Recommender Systems, 2020: 200-209.

[17] ZHANG Q, JIA Q, WANG C, et al. Amm: Attentive multi-field matching for news recommendation [C]// proceedings of the Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2021: 1588-1592.

[18] WU C, WU F, QI T, et al. Mm-rec: Visiolinguistic model empowered multimodal news recommendation [C]// proceedings of the Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022: 2560-2564.

[19] WU C, WU F, QI T, et al. Is News Recommendation a Sequential Recommendation Task? [C]// proceedings of the Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022: 2382-2386.

[20] MCMAHAN B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data [C]// proceedings of the Artificial intelligence and statistics, 2017: PMLR: 1273-1282.

[21] AMMAD-UD-DIN M, IVANNIKOVA E, KHAN S A, et al. Federated collaborative filtering for privacy-preserving personalized recommendation system [J]. arXiv preprint arXiv:190109888, 2019.

[22] WU C, WU F, CAO Y, et al. Fedgnn: Federated graph neural network for privacy-preserving recommendation [J]. arXiv preprint arXiv:210204925, 2021.

[23] CHEN G, ZHANG X, SU Y, et al. Win-win: a privacy-preserving federated framework for dual-target cross-domain recommendation [C]// proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, 2023: 4149-4156.

[24] LI L, LIN F, XIAHOU J, et al. Federated low-rank tensor projections for sequential recommendation [J]. Knowledge-Based Systems, 2022, 255: 109483.

[25] LIU R, CAO Y, WANG Y, et al. PrivateRec: Differentially Private Model Training and Online Serving for Federated News Recommendation [C]// proceedings of the Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2023: 4539-4548.

[26] YI J, WU F, ZHU B, et al. UA-FedRec: untargeted attack on federated news recommendation [C]// proceedings of the Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2023: 5428-5438.

[27] KENTON J D M-W C, TOUTANOVA L K. Bert: Pre-training of deep bidirectional transformers for language understanding [C]// proceedings of the Proceedings of naacL-HLT, 2019: 2.

[28] KUMMERVOLD P E, DE LA ROSA J, WETJEN F, et al. Operationalizing a national digital library: The case for a Norwegian transformer model [J]. arXiv preprint arXiv:210409617, 2021.

[29] LIAN J, ZHANG F, XIE X, et al. Towards Better Representation Learning for Personalized News Recommendation: a Multi-Channel Deep Fusion Approach [C]// proceedings of the IJCAI, 2018: 3805-3811.