Computers, Materials & Continua CMC, vol.64, no.1, pp.557-570, 2020

Personalized News Recommendation Based on the Text and Image Integration

Kehua Yang[1, *], Shaosong Long[1], Wei Zhang[1], Jiqing Yao[2] and Jing Liu[1]

Abstract: The personalized news recommendation has been very popular in the news recommendation field. In most research, the picture information in the news is ignored, but the information conveyed to the users through pictures is more intuitive and more likely to affect the users' reading interests than the one in the textual form. Therefore, in this paper, a model that combines images and texts in the news is proposed. In this model, the new tags are extracted from the images and texts in the news, and based on these new tags, an adaptive tag (AT) algorithm is proposed. The AT algorithm selects the tags the user is interested in based on the user feedback. In particular, the AT algorithm can predict tags that a user may be interested in with the help of the tag correlation graph without any user feedback. The proposed AT algorithm is verified by experiments. The experimental results verified the AT algorithm regarding three evaluation indexes F1-score (F1), area under curve (AUC) and mean reciprocal rank (MRR). The recommended effect of the proposed algorithm is found to be better than those of the various baseline algorithms on real-world datasets.

Keywords: News recommendation, F1, AUC, MRR, feedback, correlation graph.

1 Introduction

Today, society has entered the era of visual communication. One of the obvious features of this era is that the news media no longer uses only text as the main expression, but also image and video representations. We counted the news released by Toutiao News, Sina News, Tencent News and other platforms, the ratio of images and plain text in the news is about 9:1. But whether it is traditional or current mainstream personalized news recommendation algorithms, they rarely pay attention to pictures in the news. This may cause the user's interest description to be inaccurate, especially his/her short-term interest. For example, users may read this news because of the pictures in the news instead of the content. The news itself is highly time-effective and can be widely disseminated through social media. For personalized news recommendations that it's not as influential as the movies [Deldjoo, Dacrema, Constantin et al. (2019)] and music recommendations [Zheng,

---

[1] College of Computer Science and Electronic Engineering and Key Laboratory for Embedded and Network Computing of Hunan Province, Hunan University, Changsha, 410082, China.

[2] Oath Verizon Company, New York, 10007, USA.

[*] Corresponding Author: Kehua Yang. Email: khyang@hnu.edu.cn.

Received: 26 January 2020; Accepted: 31 March 2020.

CMC. doi:10.32604/cmc.2020.09907 www.techscience.com/journal/cmc

Kondo, Zilora et al. (2018)] over a period of time.

In this paper, the dataset is visualized as shown in Fig. 1. The average number of images in each news type exceeds 3.07, and the news categories, including five or more images, account for 66.67%, especially in tourism and design fields. There are more news photos and photography, and the number of views is relatively large. As presented in Fig. 2. The numbers of images and news views are positively correlated, which can explain why the relevant images are embedded in the news.

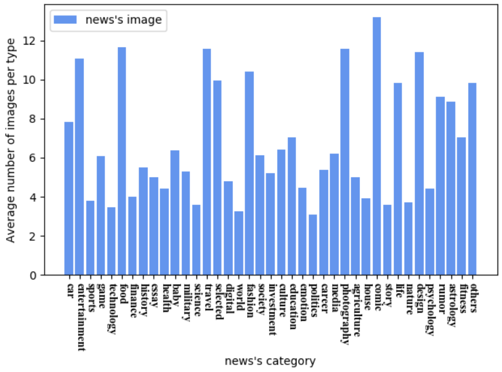

Figure 1: Average pictures of each news category

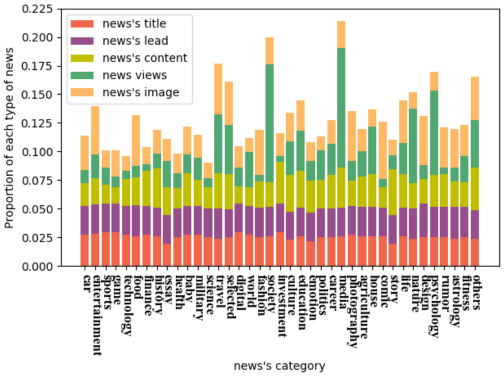

Figure 2: The proportion of the average value of different news contents

We proposed an algorithm based on news content recommendations. Different from the other related algorithms, the proposed algorithm introduces the information of news images, merges images and news texts to reconstruct the news tags, and used these new tags to establish the tag correlation graph, and also proposed an adaptive tag algorithm. The experimental results showed that the proposed algorithm achieved superior performance to the other common algorithms regarding the F1, AUC, and MRR values, which proved that taking images in the news into account could be helpful to improve the recommendation effect.

2 Related work

The news recommendation based on the news content and user behavior is to recommend appropriate news to users according to news content and user behavior. Adnan et al. [Adnan, Chowdury, Taz et al. (2014)] proposed a fuzzy information system which based on news content and recommended news with a high relevance. Kompan et al. [Kompan and Bieliková (2010)] proposed a method that the news was represented by a vector, where the correlation between the two vectors was calculated, then the matching news was searched according to the relevant high news. Tasci et al. [Tasci and Cicekli (2014)] proposed a time-based recommendation method that considered the user's recent behavior as content that the user was currently interested in Bansal et al. [Bansal, Das and Bhattacharyya (2015)] proposed the content-based recommendation method, which explored the potential user interest using the user feedback on the recommended content. Wang et al. [Wang and Shang (2016); Desyaputri, Erwin, Galinium et al. (2013)] proposed a recommendation method based on the user's click-based historical data and feedback. Zhu et al. [Zhu, Li, Liang et al. (2019)] proposed a dynamical user profiling model to describe the user's preferences.

News recommendation based on external knowledge embedding mainly refers to embedding content other than news content into the news, which is usually a well-established knowledge map. Joseph et al. [Joseph and Jiang (2019)] propsoed a method of constructing knowledge maps (ECKGs) centered on news events. Rospocher et al. [Rospocher, Van Erp, Vossen et al. (2016)] proposed the SED algorithm, where the knowledge map was used to calculate the entity with the shortest distance from the news entity. Wang et al. [Wang, Zhang, Xie et al. (2018)] proposed the DKN model, and the key point of the DKN model was the use of the knowledge map embedding. Gao et al. [Gao, Xin, Liu et al. (2018)] proposed a multi-dimensional self-attention recommendation model, and the key point of this model was also the context embedding using the entity of the knowledge map.

News recommendation methods based on the deep learning networks use a multi-layer neural network to represent user interest. Kumar et al. [Kumar, Khattar, Gupta et al. (2017)] proposed a model for embedding the semantic similarity between the words into a three-dimensional convolution network. This model included three parts, a 3D tensor, 3D convolution, and score aggregation. Zhu et al. [Zhu, Zhou, Song et al. (2019)] proposed the DAN model mainly composed of three networks: PCNN, ANN, and ARNN. Kumar et al. [Kumar, Khattar, Gupta et al. (2017)] proposed a recommendation model for time-based deep learning, which was based on a level approach, and used it to construct user's interest. Moreira et al. [Moreira, Ferreira and Da Cunha (2018)] proposed a session-based deep learning recommendation framework.

Different from all the mentioned methods and algorithms, this paper proposed to reconstruct the news tags based on news content, and established the global relevance of these tags, and then dynamically adjusted the weights of tags according to user current interest based on user feedback.

3 Our method

3.1 Refactored news tags

In the existing news tags, there are three shortcomings, which are as follows: 1) news tags has no weight indexes, so the recommendation system cannot distinguish which tags represents the main information of the news, 2) the number of news tags is relatively small; the average number of original news tags in the dataset used in this paper is 5.7, which cannot fully reflect the content of the news, 3) creators of the tags and the users may have different concerns about the news. In order to solve these three shortcomings of news tags, we propose a model for news tags reconstructing (MRNT), and it is shown in Fig. 3. The proposed model enhances the original news tag, and the refactored news tag contains the images, contents, and tags in the news. Different from the method proposed in Yin et al. [Yin, Meng, Li et al. (2019)] for extracting entities from text, this method can only extract limited entities.

In this paper, a set $\mathbf{P}=(p_1, p_2, p_3, \cdots)$ includes the tags extracted from the picture, and a set $\mathbf{L}=(l_1, l_2, l_3, \cdots)$ includes the original tags of the news, and $\mathbf{T}=(t_1, t_2, t_3, \cdots)$ is the tags extracted from the news text content. A set $\mathbf{R}=(\mathbf{T}, \mathbf{P}, \mathbf{L})=(t_1, t_2, \cdots, p_1, p_2, \cdots, t_i, t_{i+1}, l_1, l_2, l_3)$ denotes the MRNT output, where $\mathbf{R}$ denotes an ordered collection that follows the order in which the corresponding content appears in the news. In the MRNT, the Baidu Ricture Recognizer[1] is used to extract the information from images, and the Open Source Word Breaker[2] is used to distinguish the vocabulary of news text. After finishing the vocabulary segmentation, tags of news text are extracted by removing the meaningless vocabulary such as stop words and function words.

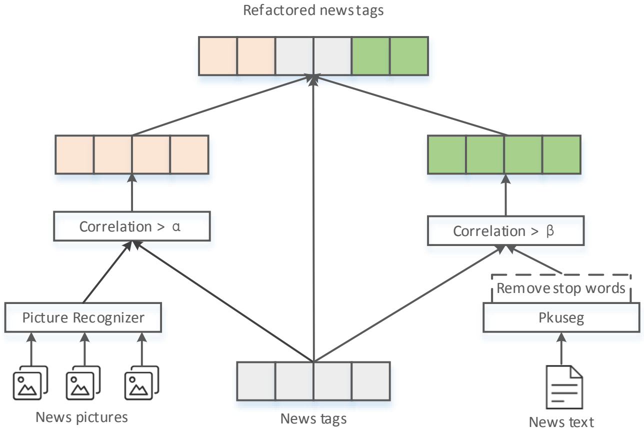

Figure 3: Refactored news tags

3.2 Building tags correlation graph

Although the proposed model enhances the original tags of the news, it only shows the information about the news without reflecting the relationship between the news and the others. On the other hand, these tags denote a lack of extension, which leads to the lack of diversity in news recommendation system. Therefore, to improve the diversity of news recommendations, we established a correlation graph containing the refactored news tags, which consists of directed and undirected graphs. The order in which refactored news tags appear is directed, and the expansion of tags is undirected. The Word2vec model [Mikolov, Chen, Corrado et al. (2013)] was employed to calculate the relevance of tags, and the word vector we used Wikipedia_zh[3]. The refactoring news tag $\mathbf{R}_i^u$ indicates that a user u reads the news i, and assuming user u_1 and user u_2 read the news i and news j respectively, the refactoring tags of the news i and news j can be expressed as $\mathbf{R}_i^{u_1} = (r_{i1}^{u_1}, r_{i2}^{u_1}, r_{i3}^{u_1}, r_{i4}^{u_1})$ and $\mathbf{R}_j^{u_2} = (r_{j1}^{u_2}, r_{j2}^{u_2}, r_{j3}^{u_2}, r_{j4}^{u_2})$, where $r_{i3}^{u_1} = r_{j3}^{u_2}$, The correlation graph of the two news tags is shown in Fig. 4.

---

[1] https://cloud.baidu.com/product/imagerecognition.

[2] https://github.com/lancopku/PKUSeg-python.

[3] https://github.com/Embedding/Chinese-Word-Vectors.

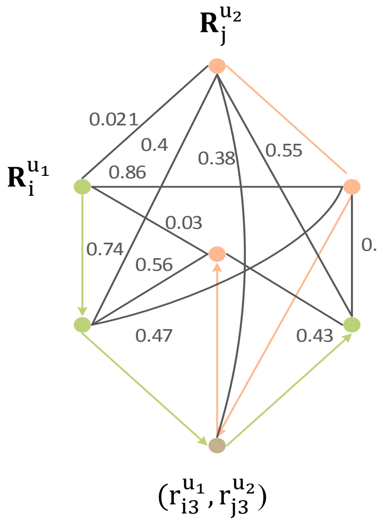

Figure 4: Refactored the news tags graph

3.3 Adaptive tag algorithm

In the online recommendation system, the short-term interest of users has a great influence on the recommendation effect. In this paper, we used a set $\mathbf{A}^u = \left(a_{j1}^u, a_{j2}^u, a_{j3}^u, \cdots\right)$ to represent the collection of short-term interest of a user u, and j=1, 2, 3 to represent the tags extracted from news images, news texts, and original news tags, respectively. Finally, the tag scores in the user short-term interest set were calculated as follows.

I. The proportion of a_{ji}^u tags in the user's browsing history is calculated as:

$$\text{history\_score}(a_{ji}^u) = \alpha_1^{r_i} \sum_{n=1}^{|H(a_{ji}^u)|} log \frac{C\left(a_{ji}^{u^n}\right) \cdot \beta_j}{C(R_n^u)} \tag{1}$$

$$r_i = |\mathbf{A}_i^u| - position(a_{ji}^u) \tag{1.1}$$

where α_1 represents the user interest attenuation value, and $\alpha_1 \geq \alpha_2 \geq \alpha_3$, and $\alpha_1^{r_i}$ indicates that α_1 has experienced r_i attenuation, and $H(a_{ji}^u)$ represents the news collection including the tag a_{ji}^u in the history browsed by user u. $|\cdot|$ indicates the size of the collection, $a_{ji}^{u^n}$ indicates that the user u browses the news n and the news n contains the tag a_{ji}, and $C(\cdot)$ is a statistical function, such as $C\left(a_{ji}^{u^n}\right)$ is the number of times the tag a_{ji} appears in news n, and $position(a_{ji}^u)$ indicates the position of the tag a_{ji}^u in $\mathbf{A}_i^u$.

II. The heat of a_{ji}^u, which combines the current short-term interest calculations of all users, is obtained by:

$$\text{tag\_hot}(a_{ji}^u) = \sum_k^U \sum_{n=1}^{|H(a_{ji}^k)|} \alpha_2^{r_k} log \frac{C\left(a_{ji}^{k^n}\right) \cdot \beta_j}{C(R_n^k)} \tag{2}$$

where $\mathbf{U}$ is the collection of all users, and $k \in \mathbf{U}$.

III. For the future score of a^{u}_{ji}, which selects the news candidate set $\tilde{H}(a^{u}_{ji})$ containing this news tag a^{u}_{ji}, is given by:

$$\text{future\_score}(a^{u}_{ji}) = \alpha_3^{\tilde{r}_i} \sum_{n=1}^{|\tilde{H}(a^{u}_{ji})|} log \frac{C({a^{u}_{ji}}^{n})\cdot\beta_j}{C(R^{u}_{n})} \tag{3}$$

$$\tilde{r_i} = position(a^{u}_{ji}) \tag{3.1}$$

IV. The deepening score of a^{u}_{ji}, which is calculated using the established tag correlation graph, is given by:

$$\text{graph\_score}(a^{u}_{ji}) = \sum_{g}^{\boldsymbol{G}} \text{correlation}(a^{u}_{ji}, a_{i'}) > \text{threshold} \tag{4}$$

where **G** is the tag correlation graph, $\text{correlation}(a^{u}_{ji}, a_{i'})$ is the correlation between the tag a^{u}_{ji} and the tag $a_{i'}$, and threshold$\in(0,1)$ is a super parameter.

The total score of tag a^{u}_{ji} can be obtained by Eqs. (1)-(4):

$$\begin{aligned}\text{score}(a^{u}_{ji}) = &-\lambda\{\text{history\_score}(a^{u}_{ji})+\text{tag\_hot}(a^{u}_{ji})+\text{future\_score}(a^{u}_{ji})\}\\ &+(1-\lambda)\text{graph\_score}(a^{u}_{ji})\end{aligned} \tag{5}$$

Using Eq. (5), we can calculate the score of each tag that a user is interested in. In order to enable $\mathbf{A}^{u}$ to change as the user's reading interest changes, this paper proposes an adaptive tag algorithm, and the pseudo code of this algorithm is given in Tab. 1. The AT algorithm adjusts the tag weight dynamically based on the user feedback, and selects the tag that the user is most interested in.

Table 1: Adaptive tag algorithm

```
Input: current user u and $\mathbf{A}^u$, K, $\mathbf{S}=\mathbf{S}'=\emptyset$ // S, S' are news tags set, K is top-k
Output: $\mathbf{A}^u$
for i=1 to $|\mathbf{A}^u|$ do
    calculate score($a^u_{ji}$),and join the S collection, $\mathbf{S}=\{<a^u_{ji},score(a^u_{ji})>\}$
$a^u_{jmax}$= max(S),$a^u_{jmax}$ is the highest scored tag in the S set
candidate($a^u_{jmax}$)=($n_1,n_2,n_3,\cdots$)// select a candidate set of news according to the tag $a^u_{jmax}$
for i=1 to |candidate($a^u_{jmax}$)| do
calculate - log$\frac{C(a^u_{jmax})\cdot\beta_j}{C(R^u_i)}$, and join the $\mathbf{S}'$ collection, $\mathbf{S}'=\{<n_i,-log\frac{C(a^u_{jmax})\cdot\beta_j}{C(R^u_i)}>\}$
rank($\mathbf{S}'$,K) // The S' collection is selected according to the score, and the score of the top K is selected to recommend it
if user u clicks news $n_{kn}$ ∈rank($\mathbf{S}'$,K)
    add $R_k$ to $\mathbf{A}^u$, and set $r_i = 0$ and $\tilde{r_i} = 0$ corresponding to $a^u_{jmax}$
    $\lambda = \lambda + \Delta\lambda$
    $\beta_j = \sqrt{\beta_j}$
```

```
else
  remove this $a^{u}_{jmax}$ from the $\mathbf{A}^{u}$
  $r_i=r_i+1$, $\tilde{r}_i=\tilde{r}_i+1$
  $\lambda = \lambda - \Delta\lambda$
  $\beta_j = \beta_j \cdot \beta_j$
end
```

4 Experiment

4.1 Datasets description

The dataset consisted of the news collected from March 2019, to May 2019 from the Toutiao News, including a total of 116,752 news, 7,425 users, and the number of users who read the collected news was greater than 50. The dataset mainly contained the user ID, news title, news content, news images, news tags, and so on. As mentioned previously, the main aim of this work is to include news images in the recommendation process; thus, in the native dataset it was needed to remove the news that did not contain images, and to remove the images that were completely unrelated to the content of the news containing the images. The results are given in Tab. 2.

Table 2: Datasets display

Total of users	Total of news	Number of images	Total of original news's tags
7,425	95,960	252,497	546,972

4.2 Parameter settings

In the extraction of the news tags by the proposed model, two hyperparameters, α and β, were used. In order to make the tags extracted from the image and the tags extracted from the news text the same as the original tags, we set α=β, The relationship between the value of the two hyperparameters and the number of extracted tags is shown in Fig. 5. When α or β was equal to 0.774, the number of news tags extracted was 29, and F1, AUC, and MRR had the maximum value; especially, at K=1, the calculation result of each indicator was the accuracy rate. As shown in Fig. 5. the number of news tags corresponding to the peaks of F1, AUC, and MRR was between 20 and 40. In summary, the initial values of α and β were set such that α or β was greater than 0.7. their super parameters are shown in Tab. 3.

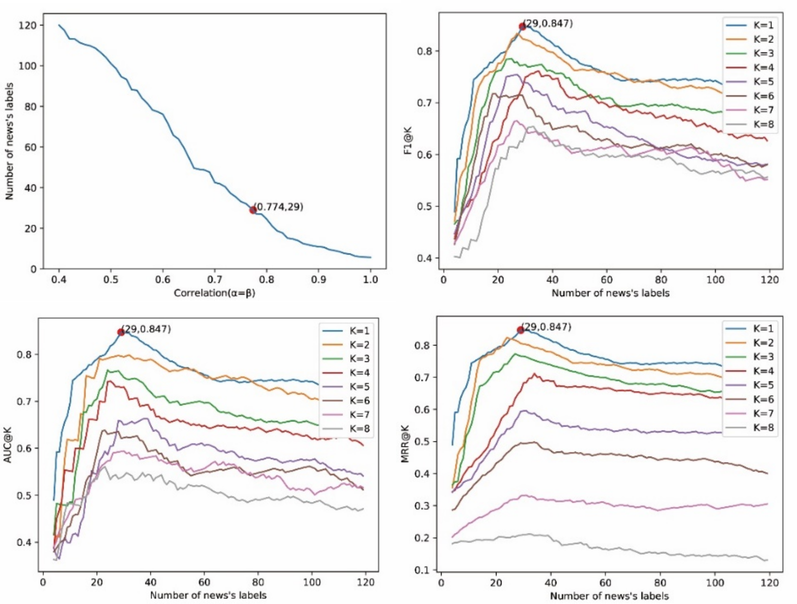

Figure 5: The relationship between the hyperparameters α and β and refactored news tags ad refactored news tags and evaluation index graph

Table 3: Parameter setting

Meaning	Value
Correlation value hyperparameters (α, β)	>0.7
Attenuation value of users' short-term interest $(\alpha_1, \alpha_2, \alpha_3)$	0.95, 0.9, 0.9
Weight of news pictures, text and original tags $(\beta_1, \beta_2, \beta_3)$	0.99
Users' short-term interest weight and increment of changing weight $(\lambda, \Delta\lambda)$	0.99, 0.035

4.3 Comparison method

This paper evaluates the proposed algorithm through the comparison with five baselines algorithms: (1) the PTGA algorithm [Shen, Ai, Xiao et al. (2018)] that determines the tags probability relationship by mining the potential relationships between different objects; (2) the LibFM algorithm [Rendle (2012)] that represents the latest feature-based matrix decomposition algorithm. In this paper, the input of the LibFM algorithm includes two parts: user characteristics (e.g., user's browsing history) and news features (e.g., the TF-IDF value of the refactored news tags); (3) the DeepFM algorithm [Guo, Tang, Ye et al. (2017)] that is composed of the neural network and the factorization machine, which are respectively responsible for the extraction of the low-order and high-order features, these two parts share the same input, the DeepFM adopts the same input as the LibFM algorithm; (4) the DSSM algorithm [Huang, He, Gao et al. (2013)] that denotes a deep semantic matching model, whose aim is to optimizes the probability of clicking on the document. In this paper, the news clicked by the user is used as a query, and then, the most similar candidates are matched to the news; (5) the DFM algorithm [Lian, Zhang, Xie et al. (2018)] that represents a deep fusion model, which mainly consists of two parts: potential representation of each subnet that is used to extract news features, and the attention mechanism that is used to fuse the data of other features.

4.4 Result

In the comparison of the baselines algorithms, we used the F1, AUC, and MRR indicators as the evaluation criteria. The comparison results are given in Tab. 4, where $|\mathbf{A}|$ represents the number of tags of the current short-term interest of a user, which has three value to maximize F1, AUC, and MRR for a given K. For instance, at K=5 and $|\mathbf{A}|$=23, the current short-term user interests set size was 23, and the F1 value was maximized. According to the results presented in Tab. 4. Our algorithm was superior to the other algorithms. Namely, the smaller the value of K was, the more obvious the advantage of our algorithm was. When K was grater than 15, our algorithm was not very advantageous. Further, the F1 value of our algorithm was higher by 3.6%-8.76% than those of the other algorithms; the AUC value was higher by 0.73%-4.18% than those of the other algorithms; and lastly, the MRR value was higher by 0.017-0.0787 than those of the other algorithms.

Table 4: Experimental results of AT algorithm and each baseline

Methods	@K=5, |A|=23, 34, 27			@K=10, |A|=21, 37, 22			@K=15, |A|=17, 32, 29		
	F1	AUC	MRR	F1	AUC	MRR	F1	AUC	MRR
AT	**75.44**	**66.34**	**0.5960**	**64.15**	**55.67**	**0.2108**	**59.24**	**52.81**	**0.1763**
PTGA	69.35	64.96	0.5435	62.15	51.49	0.1321	56.33	50.35	0.1239
LibFM	67.43	65.61	0.5532	61.87	52.37	0.1411	55.84	51.18	0.1321
DeepFM	66.68	64.72	0.5395	60.55	52.65	0.1427	56.44	51.45	0.1356
DSSM	71.76	65.50	0.5773	63.83	53.70	0.1688	57.36	51.66	0.1473
DFM	69.58	64.85	0.5694	62.39	53.26	0.1584	56.82	51.75	0.1464

In the experiments, the LibFM and DeepFM algorithms generally performed poorly, which was because these two algorithms considered all the user characteristics, ignoring the timeliness of the news, so it could happen that these algorithms recommended some outdated news to a user, resulting in poor overall performance. The PTGA algorithm had better performances than the LibFM and DeepFM algorithms regarding F1 value, but they were not as stable as the AUC and MRR, because these algorithms focused on mining the potential relationships between the objects that were not of interest to the user. The DSSM was the second-best algorithm regarding all the indicators, but its stability was better than that of the AT algorithm and the AUC and MRR algorithms, probably because this algorithm was better at mining the news depth features. The possible reason for such results could be that the AT algorithm extracted news features through the word segmentation, which could cause feature lost, resulting in poor algorithm stability. In the experiments, the DFM performance was influenced by the datasets. When different users and news were used in the training set, the test results were also different. The results presented in Tab. 4. which represents the average values of multiple random samples.

Lastly, to verify that AT algorithm had a small K value, the average value of each evaluation index at K=1, 2, ⋯, 8 was determined, firstly. The experimental results were shown in Fig. 6, where it can be seen that the AT algorithm had great advantages regarding the F1 and MRR indicators, especially regarding the MRR indicator, demonstrating that it could capture the short-term user interest well and recommend the news according to the user interest.

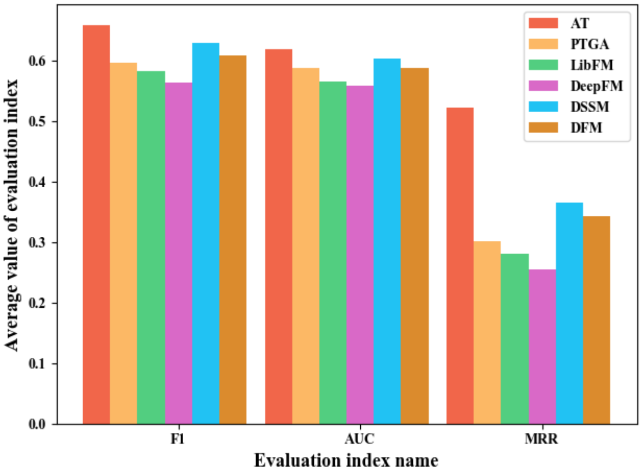

Figure 6: Average results of F1, AUC and MRR for different baselines

Table 5: The average recommended time and memory cost of AT algorithm and each baseline

Methods	@K=5, |A|=23, 34, 27		@K=10, |A|=21, 37, 22		@K=15, |A|=17, 32, 29	
	ARTC (s)	Memory (MB)	ARTC (s)	Memory (MB)	ARTC (s)	Memory (MB)
AT	**1.25014**	**951**	**1.37639**	**974**	**1.40855**	**987**
PTGA	0.91738	843	0.94785	880	0.99347	942
LibFM	0.79366	452	0.81552	466	0.85563	492
DeepFM	0.78771	539	0.80236	550	0.81342	576
DSSM	0.78045	578	0.79861	594	0.82679	635
DFM	0.80191	463	0.82056	484	0.83556	501

Secondly, we also compared the average recommended time consumption (ARTC) and memory cost of each baseline, as shown in Tab. 5. As can be seen from the table, AT algorithm ARTC is 0.33276-0.59513 higher than other baselines, and memory cost is 45-508 higher. But response time is very important for recommendation system, so we optimize AT algorithm. In the experiment, we found that formula (4) is the most time-consuming operation of AT algorithm, because it needs to traverse the established news tag graph, and the time complexity is $O(n^2)$. However, we found that Eq. (4) is needed to calculate the tags that users are interested in, accounting for about 27.58% of tags that can be calculated correctly without Eq. (4), as shown in Fig. 7.

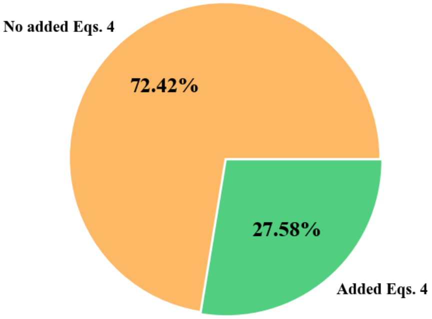

Figure 7: Average results of F1, AUC and MRR for different baselines

Table 6: Experimental results after optimizing the AT algorithm

Methods	@K=5, |A|=23, 34, 27			@K=10, |A|=21, 37, 22			@K=15, |A|=17, 32, 29		
	F1	AUC	MRR	F1	AUC	MRR	F1	AUC	MRR
AT	73.61± 0.335	65.84± 0.152	0.5573± 0.0927	61.59± 0.458	53.58± 0.119	0.1748± 0.0836	55.18± 0.547	49.28± 0.143	0.1397± 0.0718
ARTC	0.89673 (s)			0.92119 (s)			1.03316 (s)		
Memory	946 (MB)			983 (MB)			977 (MB)		

In order to reduce the calculation of Eq. (4) and improve the recommended response, we optimize the AT algorithm by: 1) we calculate Eqs. (1)-(3), get the tags that users are interested in, and then select the candidate news according to this tag; 2) according to the user's feedback, if the user clicks the recommended news, Eq. (4) doesn't need to be calculated, otherwise, Eq. (4) needs to be added to recalculate. Our optimized experimental results were shown in Tab. 6.

According to the results in Tab. 6, through the optimized AT algorithm, when K is less than or equal to 10, the recommended response time is less than 1 second, but the indexes F1, AUC and MRR also decreases by 2.816%, 2.04% and 0.0371 respectively. Therefore, our optimization method is effective in the case of small errors, which can improve the user's reading experience.

5 Conclusion

In this paper, the MRNT algorithm is used to reconstruct the news tags, which consists of the original news tags and tags extracted from the news image and text. The AT algorithm, which considers the tags that a user is currently interested in from four dimensions and adjusts the weight of the tags that the user is interested in based on the user feedback, was proposed. In particular, if the user does not provide feedback, AT algorithm will use the established tag correlation graph to deepen each tag in the current short-term user interest tag set, and then it will recalculate and select the tag with the highest score. The experimental results show that when there are more recommended items in the news, it has no advantage compared with the other baselines algorithms. Conversely, in the case of fewer recommended news items, the performance of AT algorithm is improved significantly. The achieved recommended effect also proves that the AT algorithm can extract the user's current most interesting tags successfully.

Acknowledgement: The authors gratefully acknowledge support from National Key R & D Program of China (No. 2018YFC0831800), National Natural Science Foundation of China (No. 61872134), Natural Science Foundation of Hunan Province (No. 2018JJ2062), Science and Technology Development Center of the Ministry of Education, and the 2011 Collaborative Innovation Center for Development and Utilization of Finance and Economics Big Data Property, Universities of Hunan Province.

Funding Statement: K. Yang would like to thank National Key R & D Program of China (grant No. 2018YFC0831800) for their financial support, and the URLs to sponsors' websites is https://service.most. gov.cn.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Adnan, M. N. M.; Chowdury, M. R.; Taz, I.; Ahmed, T.; Rahman, R. M. (2014): Content based news recommendation system based on fuzzy logic. *International Conference on Informatics*, *Electronics & Vision*, pp. 1-6.

Bansal, T.; Das, M.; Bhattacharyya, C. (2015): Content driven user profiling for comment-worthy recommendations of news and blog articles. *RecSys'15: Proceedings of the 9th ACM Conference on Recommender Systems*, pp. 195-202.

Deldjoo, Y.; Dacrema, M. F.; Constantin, M. G.; Eghbal-Zadeh, H.; Cereda, S. et al. (2019): Movie genome: alleviating new item cold start in movie recommendation. *User Modeling and User-Adapted Interaction*, vol. 29, no. 2, pp. 291-343.

Desyaputri, D. M.; Erwin, A.; Galinium, M.; Nugrahadi, D. (2013). News recommendation in Indonesian language based on user click behavior. *International Conference on Information Technology and Electrical Engineering*, pp. 164-169.

Gao, J.; Xin, X.; Liu, J.; Wang, R.; Lu, J. et al. (2018): Fine-grained deep knowledge-aware network for news recommendation with self-attention. *IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 81-88.

Guo, H.; Tang, R; Ye, Y.; Li, Z.; He, X. (2017): DeepFM: a factorization-machine based neural network for CTR prediction. *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pp. 1725-1731.

Huang, P.; He, X.; Gao, J.; Deng, L.; Acero, A. et al. (2013): Learning deep structured semantic models for web search using clickthrough data. *Proceedings of the 22nd ACM International Conference on Conference on Information & Knowledge Management*, pp. 2333-2338.

Joseph, K.; Jiang, H. (2019): Content-based news recommendation via shortest entity distance over knowledge graphs. *Companion Proceedings of the World Wide Web Conference*, pp. 690-699.

Kompan, M.; Bieliková, M. (2010): Content-based news recommendation. *E-Commerce and Web Technologies. 11th International Conference*, vol. 61, pp. 61-72.

Kumar, V.; Khattar, D.; Gupta, S.; Varma, V. (2017): Word semantics based 3-D convolutional neural networks for news recommendation. *IEEE International Conference on Data Mining Workshops*, pp. 761-764.

Kumar, V.; Khattar, D.; Gupta, S.; Gupta, M.; Varma, V. (2017): User profiling based deep neural network for temporal news recommendation. *IEEE International Conference on Data Mining Workshops*, pp. 765-772.

Lian, J.; Zhang, F.; Xie, X.; Sun, G. (2018): Towards better representation learning for personalized news recommendation: a multi-channel deep fusion approach. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pp. 3805-3811.

Moreira, G. D. S. P.; Ferreira, F.; Da Cunha, A. M. (2018): News session-based recommendations using deep neural networks. *Proceedings of the 3rd Workshop on Deep Learning for Recommender Systems*, pp. 15-23.

Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. (2013): Efficient estimation of word representations in vector space. *Proceedings of the First International Conference on Learning Representations*.

Rospocher, M.; Van Erp, M.; Vossen, P.; Fokkens, A.; Aldabe, I. et al. (2016): Building event-centric knowledge graphs from news. *Journal of Web Semantics*, vol. 37-38, pp. 132-151.

Rendle, S. (2012): Factorization machines with libFM. *ACM Transactions on Intelligent Systems and Technology*, vol. 3, no. 3, pp. 57:1-57:22.

Shen, Y.; Ai, P.; Xiao, Y.; Zheng, W; Zhu, W. (2018): A tag-based personalized news recommendation method. *14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery*, pp. 964-970.

Tasci, S.; Cicekli, I. (2014): Content-based news recommendation: comparison of time-based system and keyphrase-based system. *International Joint Conference on Knowledge Discovery, Knowledge Engineering, and Knowledge Management*, vol. 553, pp. 84-98.

Wang, Y.; Shang, W. (2016): Personalized news recommendation based on consumers' click behavior. *International Conference on Fuzzy Systems & Knowledge Discovery*, pp. 634-638.

Wang, H.; Zhang, F.; Xie, X.; Guo, M. (2018): DKN: deep knowledge-aware network for news recommendation. *Proceedings of the World Wide Web Conference*, pp. 1835-1844.

Yin, L.; Meng, X.; Li, J.; Sun, J. (2019): Relation extraction for massive news texts. *Computers, Materials & Continua*, vol. 60, no. 1, pp. 275-285.

Zheng, E.; Kondo, G. Y.; Zilora, S.; Yu, Q. (2018): Tag-aware dynamic music recommendation. *Expert Systems with Applications*, vol. 106, pp. 244-251.

Zhu, Z.; Li, D.; Liang, J.; Liu, G.; Yu, H. (2019): A dynamic personalized news recommendation system based on bap user profiling method. *Institute of Electrical and Electronics Engineers Access*, vol. 6, pp. 41068-41078.

Zhu, Q.; Zhou, X.; Song, Z.; Tan, J.; Guo, L. (2019): DAN: deep-attention neural network for news recommendation. *The Thirty-First Innovative Applications of Artificial Intelligence Conference, AAAI*, vol. 33, no. 01, pp. 5973-5980.