Intention-aware User Modelling for Personalized News Recommendation*

Rongyao Wang[1], Shoujin Wang[2], Wenpeng Lu[1](✉), Xueping Peng[3], Weiyu Zhang[1], Chaoqun Zheng[1], and Xinxiao Qiao[1]

[1] School of Computer, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China
[2] Data Science Institute, University of Technology Sydney, Sydney, Australia
[3] Australian Artificial Intelligence Institute, University of Technology Sydney, Sydney, Australia

honorw@foxmail.com, {wenpeng.lu, zwy, qxxyn}@qlu.edu.cn, cqzhengwork@163.com, {shoujin.wang, xueping.peng}@uts.edu.au

Abstract. Although tremendous efforts have been made in the field of personalized news recommendations, how to accurately model users' reading preferences to recommend satisfied news remains a critical challenge. In fact, users' reading preferences are often driven by his/her high-level goal-oriented *intentions*. For example, in order to satisfy the intention of traveling, a user may prefer to read news about national parks or hiking activities. However, existing methods for news recommendations often focus on capturing users' low-level preferences towards specific news only, neglecting to model their intrinsic reading intentions, leading to insufficient modelling of users and thus suboptimal recommendation performance. To address this problem, in this paper, we propose a novel intention-aware personalized news recommendation model (IPNR), to accurately model both a user's reading intentions and his/her preference for personalized next-news recommendations. In addition to modelling users' reading preferences, our proposed model IPNR can also capture users' reading intentions and the transitions over intentions for better predicting the next piece of news which may interest the user. Extensive experimental results on real-world datasets demonstrate that IPNR outperforms the state-of-the-art news recommendation methods in terms of recommendation accuracy [4].

Keywords: News recommendation · Intention-aware user modelling · User preference · Graph convolutional network.

* The work is partly supported by Key R&D Program of Shandong under Grant No.2020CXGC010901, Nature Science Foundation of Shandong under Grant No.ZR2022MF243, National Nature Science Foundation of China under Grant No.61502259, Program of Science and Technology of Qilu University of Technology under Grant No.2021JC02010. Wenpeng Lu is the corresponding author.

[4] The source code is available at: https://github.com/whonor/IPNR

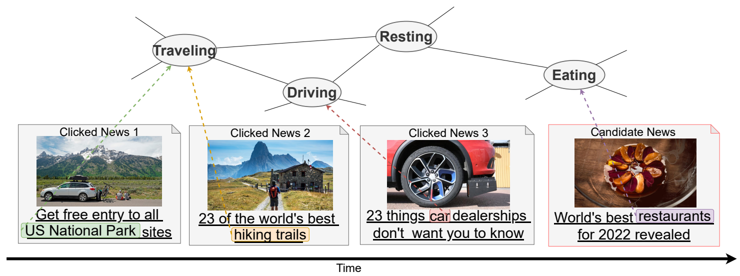

Fig. 1. An example of user reading behaviours with three clicked news and a candidate one. Each news reading behaviour is driven by a specific intention which is denoted by the word in the corresponding circle.

1 Introduction

With the prevalence of online news platforms, such as Apple News and Google News, users are overwhelmed with a large amount of online news covering various topics every day. This makes it difficult for users to quickly find out interesting news. To alleviate such an information overload problem, it is essential to recommend a small set of news that interests users according to their preferences for saving their time and improving their reading experience. Therefore, news recommender systems (NRS) have become a critical component of online news reading platforms, and they have attracted much attention from both industry and academia in recent years [1–3].

Many efforts have been devoted to news recommender systems and thus different research directions have been formed. Earlier methods strive to utilize topic models and collaborative filtering to represent news and users to recommend suitable news [4, 5]. In recent years, deep learning has gained exceptional success in news recommendations, which focuses on modeling user reading preferences from different perspectives for recommendations. For example, some works focus on generating accurate news representations with convolutional neural networks (CNN) and attention mechanisms [6, 7]. However, they only consider users' single and static preferences hidden in the user's reading history. To learn users' dynamic preferences over time, some works attempt to model users' long-term and short-term preferences with recurrent neural networks (RNN) [8] and graph neural networks (GNN) [9] respectively. However, they neglect to consider the diversity of users' preferences. More recently, to solve such a problem, some works have been done to model the potential multiple preferences with a parallel network [10] or poly attention mechanism [2]. Although these existing works have achieved better performance, they merely focus on modelling users' preferences, while neglecting to model users' intrinsic intentions which essentially drive users' reading behaviours and affect users' reading preferences. In practice, for a given user, his/her reading behaviours are often jointly determined by both his/her intentions and preferences [11]. Users' intentions are intrinsic and a high-level signal to indicate users' behaviour direction (e.g., to read political news or read economic news) while preferences are specific and a low-level signal to decide which specific news piece may be of the user's interest [12]. In most cases, users often first have a goal to read some topic/event-related news, such as economic news, which indicates their reading intentions. Then, for the same intention (e.g., to read economic news), different users may have different preferences towards specific news and thus may read different pieces of economic news. For instance, some users may like to read economic news in Wall Street Journal while others may read economic news in The Economist. Although both intentions and preferences are important for determining users' reading behaviours, most existing works fail to capture users' reading intentions, which inevitably results in insufficient modelling of users and thus sub-optimal news recommendation performance [13, 14]. Actually, how to accurately model users' reading intentions is a very critical problem for NRSs.

Some pioneering works have attempted to capture user intentions in next-basket recommendations [15, 16]. Moreover, some other related works strive to construct knowledge graphs to model interest transitions in session-based recommendations [17,18]. Although these works have achieved great success, they are devised for product recommendations only in the e-commerce domain. They often use a relatively simple model to handle product IDs without modelling any semantic and descriptive information [19]. This greatly differentiates them from news recommendations. Different from products, in addition to news IDs, a piece of news also contains an article with rich content and semantic information, which provide essential clues for us to capture users' reading intentions and preferences. Such an informative textual article should be well-modelled with more complicated and powerful models. This prevents the aforementioned works for product recommendations from being applied to news recommendations directly. Therefore, it is an urgent demand to devise novel news recommender systems with the capability of good modelling of both users' reading intentions and preferences.

Based on our observations, users' reading behaviours often show the following two unique characteristics. First, a user Tom may click multiple preferred news articles which are fit his/her high-level reading intentions. For example, as shown in Fig. 1, in the beginning, a given user has the intention of *traveling* and thus he preferred to click travel-related news, namely News 1 and 2, corresponding to national parks and hiking trails respectively. Apparently, the preferences towards news are driven by high-level intentions. Second, a user's reading intentions keep changing over time. For example, at the current time, Tom's intentions may be *traveling* and *driving*. However, the next time, his intention may transit to *resting* or *eating* as shown in Fig. 1. As a result, some news articles about restaurants may be preferred by Tom for his next click. Apparently, modelling the transitions of intentions is critical for accurate news recommendations.

To address the unique challenges triggered by the aforementioned unique characteristics in news recommendations, we propose a novel news recommendation framework called Intention-aware Personalized News Recommendation (IPNR). Thanks to the careful and unique design, IPNR is able to accurately and effectively model each user's reading intentions and his/her reading preferences for accurate next-news recommendations. In IPNR, first, a novel *reading intention module* based on graph neural networks is devised to generate the representation of the user's current reading intention by comprehensively modelling his/her intention transitions over time. At the same time, a novel *reading preference module* is devised to learn the representation of the user's reading preference from the historical news which has been read by the user. Then, a *gate network* is carefully designed to smartly aggregate the learned reading intention representation and reading preference representation together to form the informative user's representation, called intention-aware user representation in this work. Finally, a *prediction module* is designed to predict the click probability of each candidate news in the user's next click action. In this module, a special *candidate-aware attention* network was designed to more accurately select the useful information for prediction via taking the candidate news information as a guidance signal.

We summarize the main contributions of this work below:

- We propose modelling users' reading intentions for accurate next-news recommendations. As far as we know, this is the first work to comprehensively model users' reading intentions and their transitions for news recommendations.
- We devise a novel intention-aware personalized news recommendation model called IPNR, to simultaneously and effectively model users' reading intentions and reading preferences. IPNR not only models reading intention transitions over time but also detects reading preferences from users' reading history.
- A novel *reading intention module* is particularly designed to first detect possible reading intentions of a given user and then to model complex transitions of reading intentions over time to infer his/her next reading intention(s).

2 Related Work

2.1 Personalized News Recommender Systems

Personalized news recommendation is critical to improving the reading experience of users [20]. Many researchers have made a great effort to enhance the performance of news recommendations with various deep learning methods, including convolutional neural networks (CNNs), attention mechanisms, recurrent neural networks (RNNs) and graph neural networks (GNNs). For example, CNNs and 3D convolutions were utilized to encode fine-grained user representations and capture interactions between users and candidate news [21,22]. The self-attention mechanism was applied to select important information from reading history to model user preferences [13,23,24]. Multi-head self-attention was employed to detect potential multiple interests in parallel so as to model the diversity of user preferences [7,10]. RNNs were used to capture users' long- and short-term preferences from their recently browsed news, to learn the transitions of user preferences over time [8]. GNNs were utilized to model semantic interactions of news content and cluster-structural representation of users' reading history for further modelling the complex transitions of users' reading preferences [9,25].

Although these existing works have achieved better performances, they merely focus on modelling user preferences, while neglecting to model the intrinsical high-level intentions that drive users to click preferred news articles. This actually inspires us to explore users' reading intentions to recommend suitable news more accurately.

2.2 Intention-aware Recommender Systems

Recently, how to capture users' potential intentions has received much attention in next-basket/session-based recommendations [26]. Some earlier works utilized RNNs together with specially-devised intention recognizers to capture user intentions in next-basket recommendations [15,16]. Recently, GNN-based methods are popular in recommendation communities. The LP-MRGNN model constructed multi-relational-item graphs over all sessions and employed GNNs to model interest transitions of users in session-based recommendations [17]. The ISRec model extracted user intentions from historical sequences and constructed an intention graph to model intention transitions [18]. The Satori model first constructed a heterogeneous graph with users, items, and categories as the user intention graph, then leveraged a graph attention network to model user intentions and preferences respectively [27]. Although the existing works have achieved great success by exploring the role of user intentions, they are carefully devised for product recommendations, whose inputs are IDs of products without any semantic information. Apparently, there is abundant text in news articles, which conveys enough semantic information to represent users' reading intentions. If the intentions are captured, they are beneficial to improve the performance of personalized news recommender systems. However, the existing works ignore the importance of users' reading intentions. To the best of our knowledge, there is no existing work that leverages both users' reading intentions and preferences simultaneously to recommend news articles.

3 Problem Formulation

A user-news interaction dataset consists of the interaction sequence of each user, which records users' reading or clicking behaviours with news articles. Let $\mathcal{D} = \{\mathcal{S}_1, \cdots, \mathcal{S}_u, \cdots, \mathcal{S}_{|\mathcal{U}|}\}$ denotes a user-news interaction dataset, where $\mathcal{U}$ refers to the set of all users, $\mathcal{S}_u$ means the interaction sequence of the user u. $\mathcal{S}_u = \{v_1, \cdots, v_t\}$ consists of t pieces of news which are sequentially interacted by the user u, where $v \in \mathcal{A}$, $\mathcal{A}$ refers to the set of all news articles.

For each user u, given the $(t-1)$ pieces of clicked news, denoted as a reading history $C_u = \{v_1, \cdots, v_{t-1}\}$, the goal of our proposed model $\mathcal{M}$ (i.e., IPNR) is to learn users' reading intentions and preferences from C_u and predict the click probability of each candidate news.

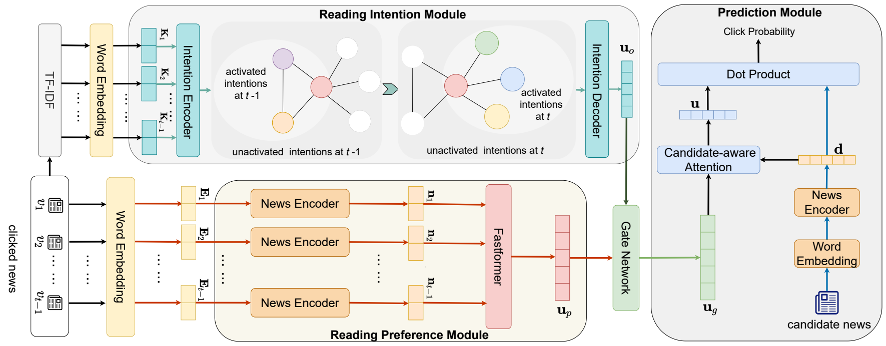

Fig. 2. Framework of IPNR, which mainly consists of four modules: a *reading intention module*, a *reading preference module*, a *gate network* and a *prediction module*. In the *reading intention module*, the nodes with/without colours indicate activated/inactivated intentions respectively.

4 The IPNR Model

4.1 Framework of IPNR

As illustrated in Fig. 2, our proposed IPNR model mainly contains four modules: (1) a *reading intention module*, which leverages GNNs to model the transitions of user intentions to generate user intention representations; (2) a *reading preference module*, which leverages CNNs and transformer networks to capture user preferences to generate user preference representations; (3) a *gate network*, which aggregates the two former representations to generate the intention-aware user representation; (4) a *prediction module*, which leverages a special *candidate-aware attention* network to incorporate candidate news features into the intention-aware user representation and predict the click probability of each candidate news.

IPNR is fed with a user's reading history $C_u = \{v_1, \cdots, v_i, \cdots, v_{t-1}\}$. We first map the content of each news v_i consisting of the news title, abstract and category, into the news embedding $\mathbf{E}_i$, which is initialized with the pre-trained Glove embeddings [28]. In addition, we utilize a TF-IDF component to extract keywords from the user's reading history, and then map the keywords of each news v_i into the keyword embedding $\mathbf{K}_i$. Afterwards, we feed the news embedding $\mathbf{E}_i$ and the keyword embedding $\mathbf{K}_i$ into the *reading preference module* and the *reading intention module*, respectively.

Users' intention graphs are critical for IPNR to leverage GNNs to model users' intentions. We denote a user's intention graph as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ denotes the set of nodes consisting of all available keywords and $\mathcal{E}$ denotes the set of edges including all directed edges. The keywords are extracted by TF-IDF from news content. Once a keyword can be matched with a concept in ConceptNet [5], it will be taken as a node in $\mathcal{G}$. The edges are connected and are set weights

[5] http://conceptnet.io/

according to the semantic relations and weights in ConceptNet. We refer to these keyword nodes in the user's intention graph as possible intentions.

4.2 Reading Intention Module

To model a user's potential intentions from the browsed news sequences, we devise a *reading intention module* to infer the possible reading intentions at the current time $(t-1)$ and then predict reading intentions at the next time (t) by modelling the intention transitions over the intention graph. This module contains an *intention encoder*, a *graph neural network* and an *intention decoder*.

Intention encoder The *intention encoder* aims to infer the possible intentions from a user's reading history. Taking the keyword embeddings of the user's clicked news $\{\mathbf{K}_1, \cdots, \mathbf{K}_i, \cdots, \mathbf{K}_{t-1}\}$ as the input, we first employ a CNN to encode these embeddings, described as:

$$\mathbf{c}_i = \mathrm{ReLU}\left(\mathbf{W}' * \mathbf{K}_{(i-f):(i+f)} + \mathbf{b}'\right), \tag{1}$$

where $\mathbf{K}_{(i-f):(i+f)}$ is the concatenation of the keyword embeddings from the position $(i-f)$ to $(i+f)$, $\mathbf{W}'$ is the learnable parameter of CNN filters, $\mathbf{b}'$ is the bias, $*$ indicates the convolutional operator, ReLU denotes a non-linear activation function. The output $\mathbf{c}_i$ is the convolutional keyword vector.

To infer the possible intentions from the keyword vectors of clicked news articles, we introduce a keyword embedding matrix $\mathbf{H}$ which is initialized by all node embeddings in the user's intention graph $\mathcal{G}$. Since only a part of keywords extracted from the news content belongs to the intention graph, we need to infer the possible keywords by capturing the relevance between the convolutional keyword vectors $\mathbf{C}_i = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{t-1}]$ and the keyword embedding matrix $\mathbf{H}$ to generate the possible keyword embedding matrix. Inspired by the work of He et al. [29], we first filter the possible keywords through a transformation matrix $\mathbf{M}$ which can learn the relations between the convolutional keyword vectors $\mathbf{C}_i$ and the keyword embedding matrix $\mathbf{H}$. Then we are able to obtain the possible keywords embedding matrix by calculating the distribution of the possible keywords in the keyword embedding matrix. The operations are specified as follows:

$$\mathbf{W}_p = \mathrm{softmax}\left(\mathbf{HMC}_i\right), \quad \mathbf{C}_p = \mathbf{W}_p\mathbf{H}, \tag{2}$$

where $\mathbf{M}$ is the learnable parameter, softmax is the normalized operator, $\mathbf{W}_p$ is the learnable weight matrix, $\mathbf{C}_p$ indicates the possible keyword matrix.

Graph Neural Network Once the possible keyword matrix is ready, we feed it into a graph neural network. To learn the transitions of user intentions on his/her intention graph, we employ a graph convolutional network (GCN) inspired by Li et al. [18]. Specifically, the operation of the l^{th} GCN layer is specified as follows:

$$\mathbf{H}^{l+1} = \mathrm{ReLU}\left(\widetilde{\mathbf{D}}^{-\frac{1}{2}}\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{H}^l\mathbf{W}^l\right), \tag{3}$$

where $\mathbf{H}^l$ is the node representation of the l^{th} GCN layer, $\mathbf{W}^l$ is a learnable matrix in the l^{th} layer, $\widetilde{\mathbf{D}}$ is a diagonal degree matrix, $\widetilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$, $\mathbf{I}$ is the identity matrix. To be specific, the input of the first layer is the keyword matrix $\mathbf{C}_p$, i.e., $\mathbf{H}^0 = \mathbf{C}_p$. Through n GCN layers, the user intention at the time (t) can be represented with $\mathbf{C}_t = \mathbf{H}^n$.

Intention Decoder After $\mathbf{C}_t$ is built by GCN, we employ the self-attention mechanism to devise the decoder to generate the representation of the user's reading intentions, as follows:

$$\alpha_j = \frac{\exp\left(\varphi\left(\mathbf{c}_j\right)\right)}{\sum_{l=1}^{R} \exp\left(\varphi\left(\mathbf{c}_l\right)\right)}, \quad \mathbf{u}_o = \sum_{j=1}^{R} \alpha_j \mathbf{c}_j, \tag{4}$$

where $\mathbf{c}_j$ indicates the j^{th} row of $\mathbf{C}_t$, R is the number of keywords. The output $\mathbf{u}_o$ is the representation of the user's reading intentions.

4.3 Reading Preference Module

Since the intention is the high-level representation which is not able to model the specific user preferences, we devise the *reading preference module* to learn fine-grained user preferences that are also an important part of user representations. This module consists of a *news encoder* and a *Fastformer*. The former learns contextual news representations for each piece of clicked news, and the latter learns the representation of the user's reading preferences from a sequence of clicked news representations.

News Encoder. As shown in Fig. 3, once the i^{th} news embedding $\mathbf{E}_i$ is obtained from the word embedding layer, we take it as the input of the *news encoder*. A news embedding contains various meta-news information, e.g., title, abstract, category and subcategory.

For the news title, we denote its embedding as $\mathbf{E}^t = \left[\mathbf{e}_1^t, \mathbf{e}_2^t, \ldots, \mathbf{e}_{|\mathbf{E}^t|}^t\right]$. To reserve position information in the sentence, we apply the positional embedding [30], e.g., $\mathbf{P}^w = \left[\mathbf{p}_1^w, \mathbf{p}_2^w, \ldots, \mathbf{p}_{|\mathbf{P}^w|}^w\right]$. Specifically, we firstly concatenate the title embedding and the positional embedding for each position x, e.g., $\mathbf{h}_x^w = \mathbf{e}_x^w + \mathbf{p}_x^w$, then feed its embedding into CNN aiming to capture important words:

$$\mathbf{t}_x = \mathrm{ReLU}\left(\mathbf{W}_c * \mathbf{h}_{(x-f):(x+f)}^w + \mathbf{b}_c\right), \tag{5}$$

where $*$ indicates a convolutional operator, $\mathbf{W}_c$ is a learnable parameter, $\mathbf{h}_{(x-f):(x+f)}^w$ is the aggregation of word embeddings from the position $(x-f)$ to $(x+f)$.

Afterwards, an attention network is utilized to learn the final title representation $\mathbf{h}^t$ from the convolutional title representation $\mathbf{t}_x$ as follows:

$$\alpha_x^t = \mathrm{softmax}\left(\mathbf{v}_t^\top \tanh\left(\mathbf{W}_t \mathbf{t}_x + \mathbf{b}_t\right)\right), \quad \mathbf{h}^t = \sum_{x=1}^{T} \alpha_x^t \mathbf{t}_x, \tag{6}$$

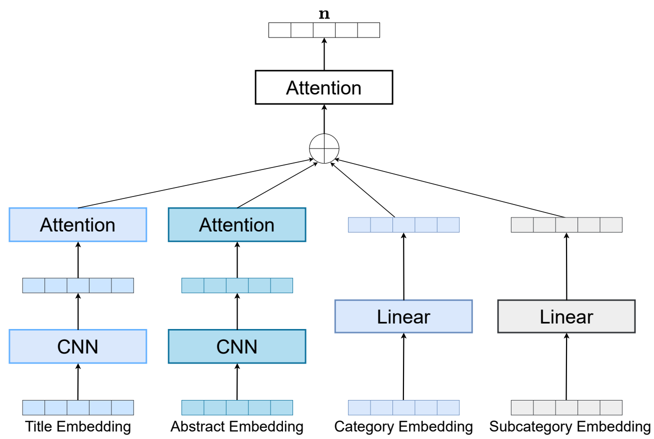

Fig. 3. Architecture of news encoder.

where $\mathbf{v}_t$, $\mathbf{W}_t$, $\mathbf{b}_t$ are learnable parameters, T is the length of a title. For an abstract, we can adopt the similar process mentioned above to generate the abstract representation $\mathbf{h}^a$. Inspired by the work of Wu et al. [31], a simple linear layer is applied to learn the category representation $\mathbf{h}^c$ and subcategory representation $\mathbf{h}^{sc}$. Finally, we aggregate all information and employ the attention mechanism to obtain the final news representation $\mathbf{n}$ as follows:

$$\mathbf{h} = \left[\mathbf{h}^t; \mathbf{h}^a; \mathbf{h}^c, \mathbf{h}^{sc}\right], \quad \alpha_x^h = \mathrm{softmax}\left(\mathbf{v}_h^\top \tanh\left(\mathbf{W}_h \mathbf{h}_x + \mathbf{b}_h\right)\right), \quad \mathbf{n} = \sum_{x=1}^{L} \alpha_x^h \mathbf{h}_x, \quad (7)$$

where $\mathbf{v}_h$, $\mathbf{W}_h$, $\mathbf{b}_h$ are learnable parameters, L is the length of meta news information, i.e., 4, $\mathbf{h}_x$ indicates one word embedding in a news embedding $\mathbf{h}$.

Fastformer. Aiming to model the informative behaviour interactions from a long news document, we utilize a state-of-the-art transformer network called *Fastformer* [32]. To be specific, we take the operation of an arbitrary attention head in *Fastformer* as example [33]. The *Fastformer* first aggregates global contexts into a query embedding $\mathbf{q}$. Next, it transforms the embedding of each token according to their relatedness with global contexts. Specifically,

$$\mathbf{q}_i = \mathbf{W}_q \mathbf{e}_i^w, \quad \mathbf{q} = \mathrm{Att}\left(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N\right), \quad (8)$$

$$\mathbf{k}_i = \mathbf{W}_k \mathbf{e}_i^w, \quad \mathbf{k} = \mathrm{Att}\left(\mathbf{q} \odot \mathbf{k}_1, \mathbf{q} \odot \mathbf{k}_2, \ldots, \mathbf{q} \odot \mathbf{k}_N\right), \quad (9)$$

$$\mathbf{v}_i = \mathbf{W}_v \mathbf{e}_i^w, \quad \hat{\mathbf{e}}_i = \mathbf{W}_o\left(\mathbf{k} \odot \mathbf{v}_i\right), \quad (10)$$

where $\mathbf{W}_q$, $\mathbf{W}_k$, $\mathbf{W}_v$, $\mathbf{W}_o$ are learnable parameters, $\mathbf{e}_i^w$ indicate the i^{th} token embedding, Att indicates the attention pooling network, $\odot$ indicates the element-wise product. $\hat{\mathbf{e}}_i$ indicates the output of i^{th} token embedding in the sequence, which is generated by the current attention head. Afterwards, we build the reading preference representation by concatenating the outputs of all attention heads:

$$\mathbf{d}_k = [\hat{\mathbf{e}}_1^k; \hat{\mathbf{e}}_2^k; \cdots; \hat{\mathbf{e}}_M^k], \quad \mathbf{u}_p = [\mathbf{d}_1; \mathbf{d}_2; \cdots; \mathbf{d}_N], \tag{11}$$

where $[;]$ indicates the concatenation operation, M is the number of attention heads, N indicates the length of a reading history sequence, $\mathbf{u}_p$ indicates the representation of the user's reading preferences.

4.4 Gate Network

The *gate network* is devised to select the important information and aggregate the representations of a user's reading intentions and preferences. Once the representations of the user reading intentions and preferences (i.e., $\mathbf{u}_o$ and $\mathbf{u}_p$) are ready, we feed them into the *gate network* to generate the intention-aware user representation. The operations are specified as follows:

$$\mathbf{g} = \mathrm{ReLU}\left(\mathbf{W}_g\left[\mathbf{u}_o; \mathbf{u}_p\right] + \mathbf{b}_g\right), \tag{12}$$

$$\mathbf{u}_g = \mathbf{g} \odot \tanh\left(\mathbf{V}\mathbf{u}_o + \mathbf{v}\right) + (1 - \mathbf{g}) \odot \mathbf{u}_p, \tag{13}$$

where $\mathbf{g}$ is a gate embedding, $\mathbf{u}_g$ is the intention-aware user representation.

4.5 Prediction Module

Before predicting the next news, we devise a *candidate-aware attention* to incorporate candidate news features into the intention-aware user representation. And then, take the candidate news representation $\mathbf{d}$ and the final user representation $\mathbf{u}$ as inputs, and we employ the dot product to predict the next piece of news. The operation is specified as follows:

$$\widehat{\alpha} = \mathrm{Att}\left(\mathbf{W}_Q\mathbf{d}, \mathbf{W}_K\mathbf{u}_g\right), \quad \mathbf{u} = \sum_{i=1}^{U} \widehat{\alpha}_i \mathbf{u}_{g,i}, \quad \hat{y} = \mathbf{u}^\top \cdot \mathbf{d}, \tag{14}$$

where $\mathbf{W}_Q$, $\mathbf{W}_K$ are learnable parameters, $\mathbf{d}$ indicates the representation of the candidate news generated by the *news encoder* shown in Fig. 3, U is the length of a user's reading history, the output $\mathbf{u}$ is the final user representation, $\hat{y}$ indicates the click probability of the candidate news.

4.6 Model Training

We utilize negative sampling strategy to train our model and employ the log-likelihood function as a loss function:

$$\mathcal{L} = -\sum_{j=1}^{\mathcal{S}} \log \frac{\exp\left(\hat{y}_j^+\right)}{\exp\left(\hat{y}_j^+\right) + \sum_{i=1}^{\mathcal{N}} \exp\left(\hat{y}_{j,i}^-\right)}, \tag{15}$$

where $\hat{y}_j^+$ indicates the probability of the j^{th} positive sample, $\hat{y}_{j,i}^-$ indicates the probability of the i^{th} negative sample w.r.t the j^{th} positive sample, $\mathcal{S}$ is the number of the training positive samples, $\mathcal{N}$ is the number of negative samples.

5 Experiment and Evaluation

5.1 Dataset and Experimental Settings

Table 1. Performance Comparison with Baselines.

Model	MIND-small				MIND-200k			
	AUC	MRR	nDCG@5	nDCG@10	AUC	MRR	nDCG@5	nDCG@10
libFM	0.6001	0.2764	0.2992	0.3595	0.6116	0.2788	0.3006	0.3644
DKN	0.6394	0.2999	0.3246	0.3941	0.6543	0.3030	0.3316	0.3993
LSTUR	0.6611	0.3100	0.3419	0.4066	0.6752	0.3238	0.3610	0.4242
NRMS	0.6682	0.3184	0.3517	0.4158	0.6701	0.3185	0.3534	0.4175
TANR	0.6455	0.3107	0.3367	0.4017	0.6611	0.3148	0.3467	0.4114
NAML	0.6588	0.3092	0.3411	0.4058	0.6765	<u>0.3269</u>	0.3623	<u>0.4270</u>
NPA	0.6613	0.3174	0.3510	0.4140	0.6734	0.3259	0.3598	0.4228
NNR	<u>0.6771</u>	<u>0.3239</u>	<u>0.3592</u>	<u>0.4222</u>	<u>0.6828</u>	0.3252	<u>0.3634</u>	0.4266
IPNR	**0.6825**	**0.3247**	**0.3615**	**0.4236**	**0.6995**	**0.3401**	**0.3785**	**0.4414**
Improv.* (%)	0.80%	0.25%	0.64%	0.33%	2.45%	4.04%	4.16%	3.37%

* The improvement over the best-performing baselines which is underlined.

The real-world news recommendation dataset MIND [6] is utilized to conduct our experiments, which contains two versions: MIND-small and MIND-large. Due to MIND-large being quite large-scale and hard to process, following the previous works [25], we randomly sample 200,000 users' behaviour logs to build a new version named MIND-200k. Besides, as the limitation of licences, we can not obtain the labels of samples in the test set of MIND-large. Therefore, we randomly split half of the original validation set into a new validation set and a new test set respectively. For experimental settings, we apply Adam optimizer to optimize the process of training. The learning rate and dropout rate are set to 2e-5, 0.2 respectively. The ratio of negative sampling $\mathcal{S}$ is set to 4 and the batch size is 64. The number of GCN layers is set to 5. We evaluate the performance of our model in terms of ROC curve (AUC), mean reciprocal rank (MRR) and normalized discounted cumulative gain (NDCG).

5.2 Baselines

We select eight state-of-the-art methods to compare with our model: [7] libFM [34], a classical matrix factorization model for news recommendations. DKN [35], a deep news recommender system, which enriches news content with external entities in a knowledge graph and employs a knowledge-aware CNN to generate news representations. TANR [36], a deep news recommender system with topic-aware news representations, which employs CNNs and attention networks to learn news representations, jointly optimized with an auxiliary topic classification task.

[6] https://msnews.github.io

[7] We adopt the official code to re-implement all baselines on the datasets.

NAML [31], an attentive multi-view recommendation model, which learns news representation from multiple kinds of news information with CNNs and attention mechanisms. NRMS [7], a deep news recommendation model, which utilizes multi-head self-attention to model news representations from news titles, and employs multi-head self-attention to capture the relatedness of browsed news to generate user representations. LSTUR [8], a neural news recommendation model based on short- and long-term user interests, which employs gated recurrent network (GRU) and user IDs to generate the representations of user's short- and long-term interests. NPA [6], a neural news recommendation model based on personalized attention, which devises a personalized attention network to recognize the important words in news content according to user preferences. NNR [25], a deep recommender system based on collaborative news encoding (CNE) and structural user encoding (SUE), which employs biLSTM and cross-attention to realize CNE and utilizes GCNs to implement SUE.

5.3 Performance Comparison with Baselines

The recommendation performance of our proposed IPNR and those of eight baselines are reported in Table 1. We have the following observations.

First, the traditional method based on matrix factorization (i.e., libFM) perform worse significantly than the other deep neural methods. This demonstrates the superiority of deep models in handling the news and users data, which can capture more sophisticated potential features in news articles and reading history.

Second, among the deep models, DKN perform worst. Although DKN introduces external entities to enrich news representation, it employs a news-level attention network which can not capture the important word-level information within a news article, and thus it performs badly.

Third, out of the last six baselines, LSTUR employs GRUs and user IDs to model users' short- and long-term preferences. It merely utilizes IDs to embed long-term preferences, which is hard to capture enough user information. NRMS utilizes multi-head self-attention to capture word and news interactions to generate news and user representations. However, it merely utilizes news title information, which inevitably misses the important features contained in news abstract and content. TANR, NAML and NPA employ various CNNs and attention mechanisms to capture important information to generate news and user representations. However, they rely on CNNs, which are not good at modelling sequential features contained in reading history and thus limit their performance. NNR utilizes biLSTM and cross-attention to realize collaborative news encoding and employs GCNs to implement structural user encoding. With the support of GCNs, NNR can capture more structural user features and can show more powerful performance than the other baselines. Interestingly, NAML performs better than NNR in terms of MRR and nDCG@10 on the MIND-200k dataset. This may be because multi-view news information adopted by NAML is efficient to model accurate news representations over a large number of news articles.

Finally, our proposed IPNR achieves the best performance on both datasets. This verifies the superiority of IPNR, which not only effectively models users' reading preferences and efficiently learns users' reading intentions to perform personalized news recommendations. Specifically, IPNR significantly outperforms the best-performing news recommendation methods (e.g., NNR) with an average of 1.98% in terms of all metrics on two datasets.

5.4 Ablation Study

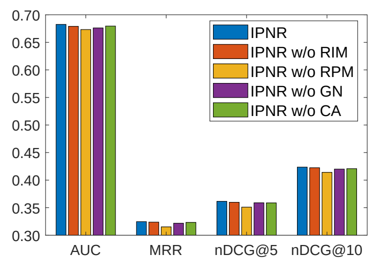

Fig. 4. Performance comparison with variants.

We design the variants of our proposed IPNR to analyze the effectiveness of each key module. According to Fig. 2, we remove the reading intention module (RIM), reading preference module (RPM) and gate network (GN) respectively to obtain three variants, i.e., *IPNR w/o RIM*, *IPNR w/o RPM* and *IPNR w/o GN*. Further, we remove the component of the *candidate-aware attention* in the prediction module to obtain *IPNR w/o CA*.

Because training on the MIND-200k dataset requires too expensive GPU cost, we only utilize the MIND-small dataset to conduct the ablation study. As shown in Fig. 4, we can see that removing each module leads tosuboptimal performance. This demonstrates that each module is critical and effective in IPNR. Comparing *IPNR w/o RIM* and *IPNR w/o RPM*, the former is better than the latter. This demonstrates that modelling preferences are more important than modelling intentions for IPNR. This is reasonable because that RPM captures more abundant news information, including titles, abstracts, etc., while RIM merely depends on the keywords extracted from news content. RIM can be applied as the complement of RPM, but can not replace RPM. Besides, the performance decrease of *IPNR w/o GN* and *IPNR w/o CA* means that the aggregation method of different representations also plays a key role in IPNR.

5.5 Hyperparameter Analysis

Influence of Number of Keywords. Due to the keywords are taken as the input of *intention encoder* in IPNR, we evaluate the influence of different numbers of keywords as shown in the upper-left panel in Fig. 5. According to these results, when the number is set to 5, our model can achieve the best performance. When the number of keywords is too small, the performance is worse. This is because a

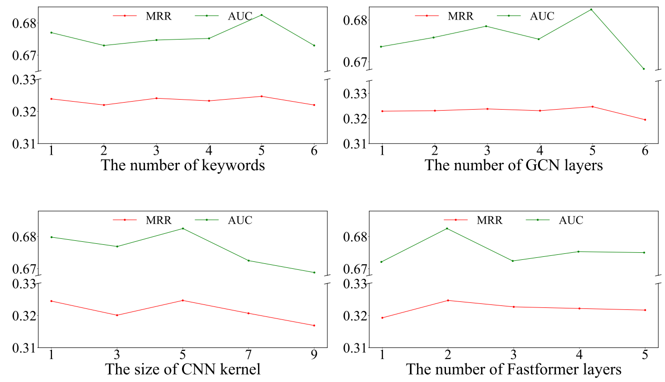

Fig. 5. The influence of hyperparameters.

few keywords can not provide enough key semantic information to the *intention encoder*, and thus hurt the modelling of potential intentions. When the number of keywords is too large, the performance begins to decline. This is because too many keywords are easy to induce noise intention information.

Influence of Number of GCN Layers. In order to explore the influence of GCN layers, we conduct several experiments as shown in the upper-right panel in Fig. 5. When the number of layers is set to 5, our model can achieve the best performance. This is probably because GCN can effectively aggregate neighbour information through 5 layers, which benefits the transitions of user intentions. When the number of layers is too small or too large, the performance will decrease. This is probably because GCN fails to efficiently capture the transitions of user intentions under the inappropriate settings on layers.

Influence of CNN Kernel Size. Because CNN is utilized in the *news encoder* in IPNR, we explore the influence of different kernel sizes in convolutional networks. As shown in the lower-left panel in Fig. 5, when the kernel size is set to 5, our model can achieve the best performance. When the kernel size is too small, our model performs worse. This is because useful information can not be fully captured, and thus it is difficult to learn accurate news representations. However, when the kernel size is set to too large, the performance of our model consistently declines. This may be because some noisy information hurts news representations.

Influence of Number of Fastformer Layers. As Fastformer is a key component in IPNR, we analyze the influence of the number of Fastformer layers. As shown in the lower-right panel in Fig. 5, when the number of Fastformer layers is set to 2, our model achieves the best performance. When the number of Fastformer layers becomes too large, the performance of our model begins to decline. The reason may be that too many layers cause the over-smoothing issue.

In the case of the model with one Fastformer layer, the performance is worse. This is because one Fastformer layer can not support IPNR to accurately learn the representation of a user's reading preferences.

6 Conclusion

In this work, we pay attention to a novel and important research problem: how to effectively model a user's reading intentions for next-news recommendations. In order to solve this problem, we have proposed an intention-aware personalized news recommendation (IPNR) model to accurately model both a user's reading intentions and his/her preferences. Extensive experimental results on real-world datasets demonstrate IPNR outperforms the state-of-the-art news recommendation methods in terms of recommendation accuracy.

References

1. Qiu, Z., Hu, Y., Wu, X.: Graph neural news recommendation with user existing and potential interest modeling. ACM Transactions on Knowledge Discovery from Data **16**(5), 1–17 (2022)
2. Li, J., Zhu, J., Bi, Q., et al.: MINER: Multi-interest matching network for news recommendation. In: ACL. pp. 343–352 (2022)
3. Wang, S., Pasi, G., Hu, L., Cao, L.: The era of intelligent recommendation: Editorial on intelligent recommendation with advanced ai and learning. IEEE Intelligent Systems **35**(5), 3–6 (2020)
4. Bansal, T., Das, M., Bhattacharyya, C.: Content driven user profiling for comment-worthy recommendations of news and blog articles. In: RecSys. pp. 195–202 (2015)
5. Ge, S., Wu, C., et al.: Graph enhanced representation learning for news recommendation. In: WWW. pp. 2863–2869 (2020)
6. Wu, C., Wu, F., An, M., Huang, J., Huang, Y., Xie, X.: NPA: Neural news recommendation with personalized attention. In: KDD. pp. 2576–2584 (2019)
7. Wu, C., Wu, F., Ge, S., Qi, T., Huang, Y., Xie, X.: Neural news recommendation with multi-head self-attention. In: EMNLP. pp. 6390–6395 (2019)
8. An, M., Wu, F., Wu, C., Zhang, K., Liu, Z., Xie, X.: Neural news recommendation with long-and short-term user representations. In: ACL. pp. 336–345 (2019)
9. Hu, L., Li, C., et al.: Graph neural news recommendation with long-term and short-term interest modeling. Information Processing & Management **57**(2), 102142 (2020)
10. Wang, R., Wang, S., Lu, W., Peng, X.: News recommendation via multi-interest news sequence modelling. In: ICASSP. pp. 7942–7946 (2022)
11. Alley, S.J., Schoeppe, S., Rebar, A.L., et al.: Age differences in physical activity intentions and implementation intention preferences. Journal of Behavioral Medicine **41**(3), 406–415 (2018)
12. Albarracin, D., Wyer Jr, R.S.: The cognitive impact of past behavior: influences on beliefs, attitudes, and future behavioral decisions. Journal of Personality and Social Psychology **79**(1), 5 (2000)
13. Qi, T., Wu, F., Wu, C., et al.: HieRec: Hierarchical user interest modeling for personalized news recommendation. In: ACL. pp. 5446—5456 (2021)

14. Wang, J., Chen, Y., Wang, Z., Zhao, W.: Popularity-enhanced news recommendation with multi-view interest representation. In: CIKM. pp. 1949–1958 (2021)
15. Wang, S., Hu, L., et al.: Intention nets: Psychology-inspired user choice behavior modeling for next-basket prediction. In: AAAI. pp. 6259–6266 (2020)
16. Wang, S., Hu, L., Wang, Y., et al.: Intention2Basket: A neural intention-driven approach for dynamic next-basket planning. In: IJCAI. pp. 2333–2339 (2021)
17. Wang, W., Zhang, W., Liu, S., et al.: Incorporating link prediction into multi-relational item graph modeling for session-based recommendation. IEEE Transactions on Knowledge and Data Engineering (Early Access) **1**(1), 1–1 (2021)
18. Li, H., Wang, X., Zhang, Z., Ma, J., Cui, P., Zhu, W.: Intention-aware sequential recommendation with structured intent transition. IEEE Transactions on Knowledge and Data Engineering **34**(11), 5403–5414 (2022)
19. Wang, S., Cao, L., Hu, L., et al.: Hierarchical attentive transaction embedding with intra-and inter-transaction dependencies for next-item recommendation. IEEE Intelligent Systems **36**(4), 56–64 (2020)
20. Wang, S., Xu, X., Zhang, X., et al.: Veracity-aware and event-driven personalized news recommendation for fake news mitigation. In: WWW. pp. 3673–3684 (2022)
21. Wang, H., Wu, F., Liu, Z., Xie, X.: Fine-grained interest matching for neural news recommendation. In: ACL. pp. 836–845 (2020)
22. Lu, W., Wang, R., Wang, S., et al.: Aspect-driven user preference and news representation learning for news recommendation. IEEE Transactions on Intelligent Transportation Systems **23**(12), 25297–25307 (2022)
23. Qi, T., Wu, F., Wu, C., Huang, Y.: News recommendation with candidate-aware user modeling. In: SIGIR. p. 1917–1921 (2022)
24. Lu, W., Zhao, P., Li, Y., Wang, S., Huang, H., Shi, S., Wu, H.: Chinese sentence semantic matching based on multi-level relevance extraction and aggregation for intelligent human-robot interaction. Applied Soft Computing **131**, 109795 (2022)
25. Mao, Z., Zeng, X., Wong, K.F.: Neural news recommendation with collaborative news encoding and structural user encoding. In: EMNLP. pp. 46–55 (2021)
26. Wang, S., Zhang, X., Wang, Y., et al.: Trustworthy recommender systems. arXiv preprint arXiv:2208.06265 **1**(1), 1–16 (2022)
27. Chen, J., Cao, Y., Zhang, F., Sun, P., Wei, K.: Sequential intention-aware recommender based on user interaction graph. In: ICMR. pp. 118–126 (2022)
28. Pennington, J., Socher, R., Manning, C.D.: Glove: Global vectors for word representation. In: EMNLP. pp. 1532–1543 (2014)
29. He, R., Lee, W.S., Ng, H.T., Dahlmeier, D.: An unsupervised neural attention model for aspect extraction. In: ACL. pp. 388–397 (2017)
30. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, , Polosukhin, I.: Attention is all you need. In: NeurIPS. pp. 6000–6010 (2017)
31. Wu, C., Wu, F., An, M., Huang, J., Huang, Y.: Neural news recommendation with attentive multi-view learning. In: IJCAI. pp. 3863–3869 (2019)
32. Wu, C., Wu, F., Qi, T., Huang, Y., Xie, X.: Fastformer: Additive attention can be all you need. arXiv preprint arXiv:2108.09084 **1**(1), 1–11 (2021)
33. Qi, T., Wu, F., Wu, C., Huang, Y.: Fum: Fine-grained and fast user modeling for news recommendation. In: SIGIR. p. 1974–1978 (2022)
34. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer **42**(8), 30–37 (2009)
35. Wang, H., Zhang, F., Xie, X., Guo, M.: DKN: Deep knowledge-aware network for news recommendation. In: WWW. pp. 1835–1844 (2018)
36. Wu, C., Wu, F., An, M., Huang, Y., Xie, X.: Neural news recommendation with topic-aware news representation. In: ACL. pp. 1154–1159 (2019)