공학석사학위논문

뉴스 추천 모델 동향과 맥락형 보조 데이터 활용 방안 제고

Survey on News Recommender Systems and Context-aware Long-/short-term User Interest Modeling

2023년 2월

서울대학교 대학원
컴퓨터공학부
김지연

뉴스 추천 모델 동향과 맥락형 보조 데이터 활용 방안 제고

Survey on News Recommender Systems and Context-aware Long-/short-term User Interest Modeling

지도교수 강 유

이 논문을 공학석사 학위논문으로 제출함

2022년 12월

서울대학교 대학원

컴퓨터공학부

김지연

김지연의 석사 학위논문을 인준함

2023년 1월

위 원 장	문 봉 기	(인)
부 위 원 장	강 유	(인)
위 원	이 상 구	(인)

Abstract

Survey on News Recommender Systems and Context-aware Long-/short-term User Interest Modeling

Jiyun Kim
Department of Computer Science & Engineering
The Graduate School
Seoul National University

How can we recommend news items to countless number of users with unique interest and preference? Harnessing various data that an individual user remains after consuming online articles, news recommender systems aim to predict unobserved his/her behavior on new articles. Since building the overall profile of an user inevitably accompanies processing information from multiple domains, digging into latent representation building in recommendation task is of value. Nevertheless, since interest may change over time and the input data is too sparse, modeling user representation is complicated.

We executed a brief survey for existing news recommender system approaches. The main contribution of the paper is as follows: First, we categorize past researches which particularly deal with systems recommending news articles by several standards tailor-made for the virtues of the item. Second, we look over recommender systems on news items for around thirty years in a comprehensive way. We find the problem they define and technique they use as time goes by.

Keywords : Recommender system, News, Auxiliary data

Student Number : 2020-21659

Contents

List of Tables

Chapter 1

Introduction

How can we offer helpful recommender result from countless number of news items to enormous amounts of unique users? For this issue, diverse recommender algorithms utilizing methods such as collaborative filtering and content-based approaches have been designed for decades. They replaced the role of manual-based recommendations and succeeded to boost the performance significantly.

On the other hand, news recommender systems should be treated as particular since there are multiple distinctive features that news as an item obtains. For example, news item inherently has a short life span unlike movies, books, and groceries. Indeed, users are prone to read new news articles to find what happened today or yesterday. Therefore, news recommender system should deal with such temporality issue and make an attempt to catch swiftly shifting trend.

Based on these features, the thesis proposes usage of deep learning technique, usage of context information, and modeling multiple user interests at the same time as three novel criteria to classify news recommender systems.

The rest of the thesis is organized as follows. In Chapter 2, we explain datasets collected for news recommendation task and metrics commonly used to evaluate news recommender models. In Chapter 3, we explore existing studies on news recommender systems and categorizing them based on novel criteria suitable for the distinctive characteristics of news recommendation. In Chapter 4, we discuss challenges and prospects of news recommendation researches.

1.1 Overview

The need to design an accurate news recommender system is rising due to various reasons such as digital news reading trends nowadays, special characteristics that news items obtain, and complexity of user interest modeling. The thesis deals with the specialty news item has as a recommendation item, and suggests three novel criteria to classify the existing news recommender methods. By grouping methods into eight categories, we can explore the prospect and challenge that news recommender has. Furthermore, we conduct an additional study regarding deep learning based context aware long- and short- user interest modeling.

1.2 Distinctive Features in News Recommender System Design

In this section, we list several distinctive features in news recommender systems. Due to these salient features, special recommender system design choices enhance the recommendation performance.

1.2.1 Temporality of News

One of the most distinctive features of news recommender systems is that the items have short life span comparing to other recommender systems. Articles published long time ago receive less click events by the audience. In contrary, articles dealing with the accident in an hour drag their attention with more ease. For example, it is clear that an article noticing a wildfire on the nearby mountain two year ago is less valuable than the one noticing the same accident occurred this morning. In other words, most of the articles have short life span. Furthermore, short-term interest of user is mostly very fragile. Sometimes breaking news or personal events distract behavior patterns that user has shown for a long time. The fact that meaningfulness of an article decays very rapidly decreases the amount of resource news recommender systems may harness while inferring user's next news choice and it is prone to aggravate the problem of cold-start in recommendation. Plus, this volatile property of news article value emerges the necessity of fast and real-time processing of the system.

1.2.2 Dynamic Shift of User Interest

Since click behavior of users on news article shifts swiftly due to its sensitiveness on popular trend of community or sudden events, user interest modeling should reflect its dynamic nature. Indeed, user interest moves to new news topics relatively fast comparing to other recommendation objects because of the aforementioned short life-span of the item itself. Furthermore, factors that influence the user interest change are various;

1.2.3 Time and Location Dependencies

Researchers in the news recommender system field found that news consumption of individual are highly affected by people around him/her. In other words, considering common preference of the audience or locals he/she belongs matters when designing the system. This tendency is more prone to news recommendation because users are often interested in community issues and local trend. Furthermore, several news topics have shifting popularity depending on temporal condition such as time of the day, weekends, etc.

1.3 Differences between this Survey and Existing Ones

Since there are various kinds of advanced techniques for recommender systems, surveys with specific technical theme have been published. Not only on recommender systems for general items, but also for specific items with distinctive features enormous comprehensive analysis and surveys have been conducted. Due to the emerging trend of online news subscription, news recommender systems are gaining their influence and importance. Therefore, surveys on news recommender systems have been also on the rise. For example, surveys such as [1], [2], [3], [4], [5], [6], [7], [8], [9] handle news recommendation systems with their methods, challenges, and goals in the future. We believe that analysis on recommendation for specific item is helpful to find special characteristics that the item obtains. Therefore, we design this survey concentrating on recommending news. This survey summarizes the background and environment of news recommender systems. Furthermore, it finally shortlists the existing researches and classifies them based on three special criteria.

The contribution of this survey is that it designs novel criteria which consider special traits that a news has as an object of recommendation. By this classification, studies on news recommender system are analyzed regarding their methods, algorithms, architectures, etc. In other words, this survey identifies characteristic features of item and describes existing system designs which suit the features. Moreover, this survey suggests challenges remained and prospects in the future regarding news recommender systems.

Chapter 2

Datasets and Metrics

2.1 Datasets

There are datasets such as MovieLens and Yelp dataset which are widely known for simulating recommendation systems. MovieLens includes a set of movie ratings from a movie recommendation service. Yelp deals with restaurant information and user reviews. However, news articles obtain several exclusive characteristics which are worth being considered when offering a recommendation. Without comprehensive attention to the distinctive features that the news has an item, news recommender system may not yield a good performance. For example, articles usually have shorter life span than movies and restaurants. Their values depreciate very quickly because the vast majority of readers urge to access the freshest news. The popularity of an article diminishes in few days but a movie or a restaurant is an item which can be found interesting after a long time from when it was launched. Consequently, datasets specifically collected for news recommendation have been on demand and released over time to capture these features and reflect news consuming behavior in the real world. In early days, few news recommendation researches conducted a user study by their own or relied on datasets from other domains. Or, they crawled articles from online news portals personally to evaluate the performance. Large publishers which possess colossal amount of article and user interaction data such as Yahoo!, MSN, Google and Bing were frequently collected. APIs from journal websites such as New

York Times were also helpful to generate datasets.

Nowadays, several news datasets are now publicly available so that researchers may access and conduct experiments with ease. Plista[10] dataset is a comprehensive data collection of articles from 13 news portals for a month with user interactions. Articles in the dataset are written in German. Not only user interactions between articles in a single portal, it may track interactions between multiple portals. The dataset separately categorized the users into editors and consumers. The number of users in the dataset is 14,897,978. In addition, it kept four types of news consuming activities which are 70,353 times of creating articles, 5,174,116 times of updating articles, 84,210,795 times of reading articles, and 1,095,323 times of following recommended links. Challenge which utilized Plista dataset to foster advancement of news recommendation algorithms was held in ACM RecSys'13. Adressa[11] dataset is a pile of online news consumption logs from Adresseavisen website for 3 months. It fully takes characteristics of news articles as an item of recommendation into consideration. There are two versions which are large and small and both of them deal with articles written in Norwegian. Adressa 20M consists of 48,486 articles and 15,514 users while Adressa 2M includes data for just one week. Unlike Plista dataset, it contains additional context information on location, time, device, and content. It enables researchers to refer to recency, dependency, and irregularity of news items. Globo[12] dataset contains news consumption patterns of 314,000 users within 46,000 articles recorded by for 16 days. The articles were published in Brazilian news portal Globo and they are written in Portuguese. With a version update, it supports user contextual attributes such as location and device information. Along with metadata about all the articles, Globo dataset provides not the raw content text but the word embeddings of the text. MIND(MIcrosoft News Dataset)[13] collected from the user behavior logs of

Table 1: Details of News Recommendation Datasets

	# users	# clicks	# articles	# portals	Period	Language	Density(%)
Plista	14,897,978	84,210,795	70,353	13	A month	German	0.008
Adressa 20M	3,083,438	27,223,576	48,486	1	10 weeks	Norwegian	0.018
Adressa 2M	15,514	2,717,915	923	1	1 week	Norwegian	0.190
Globo	314,000	3,000,000	46,000	1	16 days	Portuguese	0.021
MIND large	1,000,000	24,155,470	161,013	1	6 weeks	English	0.015
MIND small	94,057	323,419	93,698	1	6 days	English	0.004
MIND demo	50,000	43,150	45,463	1	6 days	English	0.002

Microsoft News. Interactions between 1,000,000 users and 161,013 English news articles during 6 weeks are kept in the dataset. Comparing to other datasets, the dataset offers detail information and samples in a large-scale. Microsoft runs leaderboard to measure model performance with ease using MIND dataset. Microsoft opened the MIND News Recommendation Challenge with the dataset in 2021. In Table 1, we compared the size and detail of the datasets.

2.2 Evaluation Methods

2.2.1 Traditional Accuracy Metrics

Performance evaluation of news recommendation systems is executed in various ways. Ranking is a great virtue of recommender system since recommendation aims for picking out items which are relatively unlikely to be satisfying to users as their next choice and displaying the most promising candidates. Earlier studies often adopted precision@K, recall@K, and F1 as criteria. Assume that set s_1 consists of new articles chosen by recommender system as promising candidate items and s_2 contains new articles that are not chosen by recommendation system. TP refers to the articles which are preferred by user in s_1, FP stands for the rest of s_1. FN implies articles not preferred by user in s_2 and TN refers to the rest of s_2. Precision@K refers to the ratio of articles chosen by the recommender system in top-K most possible candidate articles. Recall@K indicates the proportion of articles chosen by the system in top-k articles which are likely to be preferred by a given user. F1@K finds the balance between precision and recall by:

$$F1@K = \frac{2}{\frac{1}{recall@K} + \frac{1}{precision@K}} \tag{2.1}$$

Several researches set accuracy, HR@K(hit ratio@K) and S@K(success@K)as standard. Accuracy shows how accurately the system estimates what articles will be preferred by a given user.

$$accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{2.2}$$

Assume that a_k refers to the rank of the k-th candidate article recommended by the system, then HR@k is:

$$HR@K = \frac{|\{k|a_k \leq K|\}|}{K} \tag{2.3}$$

S@K(Success@K) refers to the odds of succeeding to estimate the next choice of user by suggesting top-k ranked candidates.

2.2.2 Modern Accuracy Metrics

Experiments conducted recently tend to be based on AUC, MRR, and nDCG as metrics. AUC(Area Under the ROC) Curve is a performance score for classification. It is interpreted as the odds that the systems ranks an article which is likely to be preferred by user higher than an article which is not. N_p and N_n stand for the number of positive and negative samples.

$$AUC = \frac{|\{(i,j)|Rank(p_i) < Rank(n_j)\}|}{N_p N_n} \tag{2.4}$$

nDCG@K(Normalized Discounted Cumulative Gain@K) is an index to represent the performance of recommender system comparing to the most ideal combination of articles. Here, r_i and p_i are the real and predicted ratings of the i-th sample.

$$nDCG@K = \frac{\sum_{i=1}^{K}(2^{r_i}-1)/log_2(1+i)}{\sum_{i=1}^{N_p} 1/log_2(1+i)} \tag{2.5}$$

MRR(Mean Reciprocal Rank) focuses on where the first relevant item appears.

$$MRR = \frac{1}{N_p}\sum_{i=1}^{N_p}\frac{1}{Rank(p_i)} \tag{2.6}$$

AP(Average Precision) measures the performance of recommendation to a single user.

$$AP = \frac{1}{N_p}\sum_{i=1}^{N_p}\frac{|\{k|Rank(p_k) \leq Rank(p_i)|\}|}{Rank(p_i)} \tag{2.7}$$

2.2.3 Other Metrics

Few researches[14],[15],[16] also measure sensitivity, specificity, diversity and novelty as additional criteria with the aforementioned metrics. [16] regards recall as sensitivity measure.

$$sensitivity = \frac{TP}{TP+FN} \tag{2.8}$$

$$specificity = \frac{TN}{TN + FP} \tag{2.9}$$

In [15], diversity is adopted as a metric to offer rich recommendation. The diversity of user u with a recommended list $R(u)$ is calculated as follows:

$$Diversity = 1 - \frac{\sum_{i,j \in R(u), i \neq j} sim(i,j)}{(1/2)|R(u)|(|R(u) - 1|)} \tag{2.10}$$

In [14], novelty is represented as the proportion of fresh items in items recommended by the system.

Chapter 3

Analysis on Existing Methods

3.1 Recommendation System with Non-Deep Learning Methods

On its initial stage, news recommendation system was conceptualized as categorizer. In 1990s, the news recommender systems are either rule-based or collaborative filtering methods. As time went by, they utilize advanced techniques such as clustering, graphs, and genetic algorithms. In this chapter, we briefly introduce each news recommender study chronologically and analyze the problems they solved.

3.1.1 Unified Interest Modeling

3.1.1.1 Models without Context Information

INFOSCOPE[17] is one of the initial trials that capture own semantic interpretations of each user and reflect them for better news browsing experience of the user. In INFOSCOPE, its agents are collections of rule-based heuristics that utilize information resulting from the analysis of user behavior to make suggestions to the user. Tapestry[18] involves humans to make a filtered mail subscription list in addition to the content of mails. The method blends collaborative filtering with existing content-based approach and also covers various data streams including online news articles. GroupLens[19] is a collaborative filtering internet news recommendation system which collects user ratings as source. Unlike Tapestry, it is able to handle multiple news clients. NewsWeeder[20] is a combination of collaborative filtering and content-based filtering method which uses both active user feedback and news text as input. In the study, it shows that machine learning and MDL(minimum description length) are better learning methods than term-frequency/inverse-document frequency weighting(TF-IDF) when categorizing documents. Krakatoa Chronicle[21] is an interactive and personalized news service which builds individual 'user profiles' to display news articles. It shows the articles, receives relevance feedback, and shows other articles by sensing not only personal interest but also community interests. P-Tango[22] predicts potential clicks of users based on a weighted average of the content-based prediction and the collaborative filtering approach. Categorizer[23] harnesses SVM(Support Vector Machine) to classify online news automatically. On the other hand, WebClipping2[24] selects articles of user's interests according to the user profile by Bayesian classifier. Newsjunkie[25] represents articles as words using BoW(Bag of Words)m and named entities by TF-IDF(Term Frequency-Inversed Document Frequency) technique. PersoNews[26] extracts topic of interest from thematic hierarchy adopting incremental Naive Bayes classifier as a news filter. Das et al.[27] utilize MinHash, PLSI(Probabilistic Latent Semantic Indexing) and covisitation counts to conduct a collaborative filtering. News@hand[28] is a news recommender system which populates domain ontologies by semantic information from external source. Hermes[29] proposes a way to classify news items based on domain ontologies. Athena[16] is also a TF-IDF based method which employs user profile to store terms or concepts from news browsed by the user. LinUCB[30] substitute a problem of personalized article recommendation to a contextual bandit problem. Massa et al.[31] use latent semantic analysis to unveil user-news relationships. SCENE[32] clusters articles by using LSH(local sensitive hashing) and probabilistic language models, constructs user profiles based on news topic distribution, similar access patterns, and news entity preference. In SCENE, news selection is treated as budged maximum converge problem. Gao et al.[33] generate semantically meaningful profiles based on time-sensitive TF-IDF(term frequency/inverse document frequency) function. Garcin et al.[14] build a context tree from article topic sequences and their corresponding suffixes. Also, they assign a set of local prediction models to each node in the tree to combine them and find relevance of an article to the context. Getting out of conventional vector space model, PENETRATE[34] puts individual and group user interests into consideration using a consensus hierarchical clustering. Li et al.[35] represent news reading community as a hypergraph to model high-order correlations between users and articles. It sees the recommendation task as a local selection problem on sub-hypergraphs. CCF[36] is a hybrid news topic recommendation approach which utilizes both news content and users-items relationship based on neighborhood models. CCF introduces a kernel to reflect news content and integrates it with collaborative filtering framework. TOCF[37] is a collaborative filtering model which considers time-sequential user behaviors by proposing time-dependent user similarity measure. CCTM[38] executes hierarchical Bayesian modeling with not only news content but also user comments of that news to recommend comment-worthy articles. Shashkin et al.[39] adopt WARP algorithm to calculate loss of factorization model and utilize implicit feedback and temporal dynamics to recommend news.

3.1.1.2 Models with Context Information

On the other hand, Newt[40]is a personalized information filtering framework utilizing relevance feedback and genetic algorithm. Iliveski et al.[41] propose a SWL(select-watch-leave) probabilistic graphical interest model to analyze the context and obtain semantic mapping from user log. NEMAH[15] uses user interest similarity, subclass popularity factor, place, freshness, and key person factors to capture underlying user-news relationship in the microblog. Techniques such as TF/IDF(term frequency/inverse document frequency) and greedy selection algorithm are used to cluster the articles. Hier-UIM[42] proposes a hierarchy user interest modeling with topic model, keyword weight sequence by term frequency/inverse document frequency, and user context information. To reflect temporal dynamics and news taxonomy, Raza et al.[43] propose a biased matrix factorization model which applies both time units and category information as biases.

3.1.2 Long and Short Term Interest Modeling

NewsDude[44] separates user interests into long-term and short-term. The paper assumes that user has plural interests and preferences. This separation is efficient due to the different characteristics that long- and short-term interests obtain. The long-term interest represents universal preference of a user which is stable and less dynamic. On the other hand, the short-term interest represents temporal preference of a user which is changing and dynamic.In NewsDude[44], long-term interest is captured by naive Bayes and short-term interest is captured by nearest neighbor algorithm. YourNews[45] also has long-term and short-term interest representations which have different number of news items. Both are weighted prototype term vectors from news browsing history of a user. LOGO[46] generates a news hierarchy by hierarchical agglomerative clustering algorithm with LDA(Latent Dirichlet Allocation) summarization, and harnesses a time sensitive weighting scheme by time decay factor to represent long-term user interest. Besides, short-term user interest is modeled by greedy algorithm. Epure et al.[47] solve the problem of session-based news recommendation. Furthermore, the paper establishes a notion of medium-term reading behavior between short- and long-term reading behaviors.

3.1.2.1 Location- and Time-Aware Modeling

Liu et al.[48] also model long-term and short-term interest respectively. CROWN[49] models data as a tensor to integrate contextual information such as time and location.

3.2 Recommendation System with Deep Learning Methods

As deep-learning methods rise and are dealt as an almighty solution to many computing problems, their potential was also applied to recommender system field and played a role of key to better performance. Recommender systems mainly utilize neural networks to generate news or user representations. Therefore, what data and which neural model to adopt for training the embedding depending on design choice. In several researches harnessing deep learning techniques, the attention mechanism is largely adopted due to its superior quality of learning relative importance. The attention mechanism is often used to capture topic words and serial user interest.

3.2.1 Unified Interest Modeling

A tendency of news click behaviors that a user has is unique and constant so that news recommender systems describe it as representation. Several researches generate user representation as one unified vector. Okura et al.[50] adopt embedding-based method using distributed representations on news recommendation task to handle synonyms and orthographical variants. Park et al.[51] overcome the challenges of news recommendation such as high number of items and sensitivity on recency or personal interests by adopting modified version of RNN and CNN. DADM[52] proposes a dynamic meta-attention model over multiple deep network to handle unstable characteristics of data and selection criteria. Weave&rec[53] applies 3D CNNs to extract spatial and temporal features from user history. HRAM[54] is a hybrid approach which concatenates generalized matrix factorization network and attention-based user-history component. DKN[55] generates multi-channel news representation by integrating various embeddings from words and entities with knowledge-aware CNN. DeepJoNN[56] is a session-based deep learning news recommendation method which utilizes user-parallel mini-batches. In NPA[57], personalized attention network considers difference on preference between individual users by using user ID queries. DAN[58] is a deep attention network which uses different neural networks to extract news representation, sequential information, and user interest. TANR[59] trains news encoding by attention network with auxiliary topic classification task so that news topic information is identified. NRMS[60] applies multi-head self-attention in a word-level while learning news representations. In addition, the same attention mechanism is executed in a news-level in learning user representation. NAML[61] utilizes various information regarding an article such as categories, title and content and integrates them into one unified news embedding by attentive pooling. PGT[62] is a news recommendation model coalescing personal preference with global temporal preference which consists of popular and fresh articles by attention and multi-layer perceptron. In FIM[63], stacked dilated convolutions are used to build multi-level representations per each news. Also, the model compares consumed news and candidate news at each semantic level. KRED[64] refers to knowledge graph and dynamic contexts such as category, frequency, and position to have better representations on entities. NRNF[65] learns positive and negative news click representations separately by neural network so that two relevance scores with candidate news are calculated and finally combined to make a decision. GNUD[66] uses information propagation along a bipartite graph which consists of user and news nodes to encode high-order relationship between user and news. For better expressiveness and interpretability, the model adopts a single graph convolutional layer with preference disentanglement. CPRS[67] captures user clicking behaviors on news titles and reading behaviors on their contents simultaneously. Each stage of encoding both behaviors and incorporating them harnesses attention mechanism. IMRec[68] is an unique attention-based model which addresses multi-modal understanding of online news articles including images through global and local impression modules. User-as-graph[69] represents each user as a heterogeneous graph built with his/her behaviors to catch the fine-grained relationship between them. Wu et al.[70] adopt pre-trained language models with attention network to learn news embeddings.

3.2.1.1 Location- and Time-Aware Modeling

DRN[71] is a reinforcement learning-based news recommendation framework which adopts deep Q-learning to model user feedback as future reward. DFM[72] is a deep fusion model to handle multiple channels or services with various user features including user ID, gender, age, and locations by inception module and attention mechanism.

3.2.2 Long and Short Term Interest Modeling

GNewsRec[66] describes interactions between users, news, and latent topics with a heterogeneous graph to abate data sparse problem. LSTUR[73] designs long-term interest representation learned from user IDs to GRU network which aims for obtaining short-term interest representation. KOPRA[74] prunes a knowledge graph used for long- and short-term user interest representation so that irrelevant part of it is cut out to focus on important information. KG-LSTUP[75] generates short-term user interest embedding with information from knowledge graph.

Chapter 4

Discussion

In this chapter, we discuss challenges that temporary news recommender systems confront and responding prospects.

4.1 Challenges

As aforementioned, news recommender systems contain several own obstacles which are more critical than recommending other items due to the distinctive features of news as an object of recommendation.

4.1.1 Temporality of Articles

Several news recommender studies such as [30],[14],[41],[49],[50],[11],[47],[39] have been dealt with the problem of short life span that news articles have. Since the object of recommendation obtains short life span, the recommendation should consider it as important issue and reflect such feature so that old news articles achieve less value to the user. Multiple datasets for news recommendation attach ordinal information of news reading behavior to take this matter into consideration. Furthermore, news recommendation approaches [50],[51],[54],[58] adopt methods which keep ordinal information such as RNN(Recurrent Neural Network) and LSTM(Long-Short Term Memory).

4.1.2 Diverse and Dynamic User Interest

User interest on news reading often evolves very quickly because there are various factors which influence news choice of a user. For example, popular trend in the community or cohort and breaking news may affect news reading. Several news recommendation papers such as [52], [43], [58], [76] mention this challenge as an obstacle that the research field must overcome. Diversity and dynamics of user interest lead to swift change of user interest representation and the recommender model should be prepared for this direct shift. For this issue, researches propose multiple interest representations for single user or attention-based interest capturing approaches.

4.1.3 Lack of Background Knowledge

One of the main challenges that news recommender systems now face is lack of background knowledge. Since news articles include rich context through their title, content, and category, relationships between words are valuable but difficult to specify based on the limited information in the article itself. News recommender studies such as [75], [69], [70] point this challenge and show that the relationships between entities not optimally modeled lead to performance loss. The studies come up with several different ways to find the relationships including the adoption of knowledge graph or graph representation.

4.1.4 Ethical Issues

Ethical issues around news recommender system are also on the rise due to the danger of fake news and bias news. In the process of recommending news, effort not to recommend unhealthy news is critical because it is directly related to recommendation quality issue. Regarding this topic, [77], [43] warn the danger of biased news and

[77] adopt bias score to calculate the bias.

4.2 Prospects

The research field of news recommender system has been evolved with the development of platforms and techniques. In the beginning, news recommender was simply treated as a filter based on heuristics. However, the notion of highly personalized news recommendation became widespread and how to model an unique news reading pattern of a single user became important. In this perspective, the rise of deep learning approaches led to skyrocketing performance. The paradigm shift occurred from previous TF-IDF based or Bayes classifier based approaches to neural network based models due to their superior ability to capture critical features in specifying user interests. Among deep learning approaches, attention mechanism becomes one of the major methods for news recommendation. Nowadays, data augmentation skills such as offering knowledge graph or auxiliary information are grabbing attention of the field. Not only generating advanced news representation, but also generating advanced user representation is constantly an important topic for news recommendation. Furthermore, detection on fake and biased news is also an issue to overcome. Since news is often an object of scam and cheat, it is more prone to be abused and taken advantage of. Therefore, more researches on such challenges and issues should be executed in the near future.

Chapter 5

Conclusion

In recent media consuming trend, news recommender system gains more influence and importance as time goes by. Due to the features of news items, challenges such as short life span of news, dynamic nature of user interest, and time/location dependencies of user reading behavior become important. Methods which adopt TF-IDF, clustering, and Bayes classifier have been proposed to deal with the challenges. Furthermore, approaches based on deep learning boost the performance of news recommender with their massive power of neural network. The deep learning approaches often generate news representations and construct user interest representation by the news browsed. To evaluate the performance of news recommender models, accuracy is the major criterion. Since online news publishers display selected news in an order, accuracy metrics considering sequence such as nDCG are popular among researches. For simulation on news recommendation, several online news datasets including user news reading history and auxiliary information around the reading behavior are released. In this survey, we classify the existing studies on news recommender systems based on three criteria and thoroughly observe the history of the field. In recommendation task which handles specific and particular items, such survey will help researchers to understand the geography of the studies.

References

[1] M. Karimi, D. Jannach, and M. Jugovac, “News recommender systems–survey and roads ahead,” *Information Processing & Management*, vol. 54, no. 6, pp. 1203–1227, 2018.

[2] S. Raza and C. Ding, “News recommender system: a review of recent progress, challenges, and opportunities,” *Artificial Intelligence Review*, pp. 1–52, 2021.

[3] R. Burke, “Hybrid recommender systems: Survey and experiments,” *User modeling and user-adapted interaction*, vol. 12, no. 4, pp. 331–370, 2002.

[4] S. Zhang, L. Yao, A. Sun, and Y. Tay, “Deep learning based recommender system: A survey and new perspectives,” *ACM Computing Surveys (CSUR)*, vol. 52, no. 1, pp. 1–38, 2019.

[5] H. L. Borges and A. C. Lorena, “A survey on recommender systems for news data,” in *Smart Information and Knowledge Management*, pp. 129–151, Springer, 2010.

[6] C. Wu, F. Wu, Y. Huang, and X. Xie, “Personalized news recommendation: A survey,” *arXiv preprint arXiv:2106.08934*, 2021.

[7] M. Li and L. Wang, “A survey on personalized news recommendation technology,” *IEEE Access*, vol. 7, pp. 145861–145879, 2019.

[8] L. Li, D.-D. Wang, S.-Z. Zhu, and T. Li, “Personalized news recommendation: a review and an experimental investigation,” *Journal of computer science and technology*, vol. 26, no. 5, pp. 754–766, 2011.

[9] C. Wu, F. Wu, Y. Huang, and X. Xie, “Personalized news recommendation: Methods and challenges,” *ACM Transactions on Information Systems (TOIS)*, 2022.

[10] B. Kille, F. Hopfgartner, T. Brodt, and T. Heintz, “The plista dataset,” in *Proceedings of the 2013 international news recommender systems workshop and challenge*, pp. 16–23, 2013.

[11] J. A. Gulla, L. Zhang, P. Liu, Ö. Özgöbek, and X. Su, "The adressa dataset for news recommendation," in *Proceedings of the international conference on web intelligence*, pp. 1042–1048, 2017.

[12] G. de Souza Pereira Moreira, F. Ferreira, and A. M. da Cunha, "News session-based recommendations using deep neural networks," in *Proceedings of the 3rd Workshop on Deep Learning for Recommender Systems*, pp. 15–23, 2018.

[13] F. Wu, Y. Qiao, J.-H. Chen, C. Wu, T. Qi, J. Lian, D. Liu, X. Xie, J. Gao, W. Wu, *et al.*, "Mind: A large-scale dataset for news recommendation," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 3597–3606, 2020.

[14] F. Garcin, C. Dimitrakakis, and B. Faltings, "Personalized news recommendation with context trees," in *Proceedings of the 7th ACM Conference on Recommender Systems*, pp. 105–112, 2013.

[15] W. Gu, S. Dong, Z. Zeng, and J. He, "An effective news recommendation method for microblog user," *The Scientific World Journal*, vol. 2014, 2014.

[16] W. IJntema, F. Goossen, F. Frasincar, and F. Hogenboom, "Ontology-based news recommendation," in *Proceedings of the 2010 EDBT/ICDT Workshops*, pp. 1–6, 2010.

[17] G. Fischer and C. Stevens, "Information access in complex, poorly structured information spaces," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 63–70, 1991.

[18] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to weave an information tapestry," *Communications of the ACM*, vol. 35, no. 12, pp. 61–70, 1992.

[19] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "Grouplens: An open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pp. 175–186, 1994.

[20] K. Lang, "Newsweeder: Learning to filter netnews," in *Machine Learning Proceedings 1995*, pp. 331–339, Elsevier, 1995.

[21] K. Bharat, T. Kamba, and M. Albers, "Personalized, interactive news on the web," *Multimedia Systems*, vol. 6, no. 5, pp. 349–358, 1998.

[22] M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes, and M. Sartin, "Combining content-based and collaborative filters in an online newspaper," in *Proceedings of ACM SIGIR workshop on recommender systems*, vol. 60, pp. 1853–1870, Citeseer, 1999.

[23] C.-H. Chan, A. Sun, and E. P. LIM, "Automated online news classification with personalization," 2001.

[24] R. Carreira, J. M. Crato, D. Gonçalves, and J. A. Jorge, "Evaluating adaptive user profiles for news classification," in *Proceedings of the 9th international conference on Intelligent user interfaces*, pp. 206–212, 2004.

[25] E. Gabrilovich, S. Dumais, and E. Horvitz, "Newsjunkie: providing personalized newsfeeds via analysis of information novelty," in *Proceedings of the 13th international conference on World Wide Web*, pp. 482–490, 2004.

[26] E. Banos, I. Katakis, N. Bassiliades, G. Tsoumakas, and I. Vlahavas, "Personews: a personalized news reader enhanced by machine learning and semantic filtering," in *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*, pp. 975–982, Springer, 2006.

[27] A. S. Das, M. Datar, A. Garg, and S. Rajaram, "Google news personalization: scalable online collaborative filtering," in *Proceedings of the 16th international conference on World Wide Web*, pp. 271–280, 2007.

[28] I. Cantador, A. Bellogín, and P. Castells, "Ontology-based personalised and context-aware recommendations of news items," in *2008 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology*, vol. 1, pp. 562–565, IEEE, 2008.

[29] F. Frasincar, J. Borsje, and L. Levering, "A semantic web-based approach for building personalized news services," *International Journal of E-Business Research (IJEBR)*, vol. 5, no. 3, pp. 35–53, 2009.

[30] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," in *Proceedings of the 19th international conference on World wide web*, pp. 661–670, 2010.

[31] R. Di Massa, M. Montagnuolo, and A. Messina, "Implicit news recommendation based on user interest models and multimodal content analysis," in *Proceedings of the 3rd international workshop on Automated information extraction in media production*, pp. 33–38, 2010.

[32] L. Li, D. Wang, T. Li, D. Knox, and B. Padmanabhan, "Scene: a scalable two-stage personalized news recommendation system," in *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, pp. 125–134, 2011.

[33] Q. Gao, F. Abel, G.-J. Houben, and K. Tao, "Interweaving trend and user modeling for personalized news recommendation," in *2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*, vol. 1, pp. 100–103, IEEE, 2011.

[34] L. Zheng, L. Li, W. Hong, and T. Li, "Penetrate: Personalized news recommendation using ensemble hierarchical clustering," *Expert Systems with Applications*, vol. 40, no. 6, pp. 2127–2136, 2013.

[35] L. Li and T. Li, "News recommendation via hypergraph learning: encapsulation of user behavior and news content," in *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 305–314, 2013.

[36] Z. Lu, Z. Dou, J. Lian, X. Xie, and Q. Yang, "Content-based collaborative filtering for news topic recommendation," in *Twenty-ninth AAAI conference on artificial intelligence*, 2015.

[37] Y. Xiao, P. Ai, C.-H. Hsu, H. Wang, and X. Jiao, "Time-ordered collaborative filtering for news recommendation," *China Communications*, vol. 12, no. 12, pp. 53–62, 2015.

[38] T. Bansal, M. Das, and C. Bhattacharyya, "Content driven user profiling for comment-worthy recommendations of news and blog articles," in *Proceedings of the 9th ACM Conference on Recommender Systems*, pp. 195–202, 2015.

[39] P. Shashkin and N. Karpov, "Learning to rank for personalized news recommendation," in *Proceedings of the International Conference on Web Intelligence*, pp. 1069–1071, 2017.

[40] B. D. Sheth, *A learning approach to personalized information filtering*. PhD thesis, Massachusetts Institute of Technology, 1994.

[41] I. Ilievski and S. Roy, "Personalized news recommendation based on implicit feedback," in *Proceedings of the 2013 international news recommender systems workshop and challenge*, pp. 10–15, 2013.

[42] M. Lu and J. Liu, "Hier-uim: A hierarchy user interest model for personalized news recommender," in *2016 4th International Conference on Cloud Computing and Intelligence Systems (CCIS)*, pp. 249–254, IEEE, 2016.

[43] S. Raza and C. Ding, "News recommender system considering temporal dynamics and news taxonomy," in *2019 IEEE International Conference on Big Data (Big Data)*, pp. 920–929, IEEE, 2019.

[44] D. Billsus and M. J. Pazzani, "A personal news agent that talks, learns and explains," in *Proceedings of the third annual conference on Autonomous Agents*, pp. 268–275, 1999.

[45] J.-w. Ahn, P. Brusilovsky, J. Grady, D. He, and S. Y. Syn, "Open user profiles for adaptive news systems: help or harm?," in *Proceedings of the 16th international conference on World Wide Web*, pp. 11–20, 2007.

[46] L. Li, L. Zheng, and T. Li, "Logo: a long-short user interest integration in personalized news recommendation," in *Proceedings of the fifth ACM conference on Recommender systems*, pp. 317–320, 2011.

[47] E. V. Epure, B. Kille, J. E. Ingvaldsen, R. Deneckere, C. Salinesi, and S. Albayrak, "Recommending personalized news in short user sessions," in *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pp. 121–129, 2017.

[48] J. Liu, P. Dolan, and E. R. Pedersen, "Personalized news recommendation based on click behavior," in *Proceedings of the 15th international conference on Intelligent user interfaces*, pp. 31–40, 2010.

[49] S. Wang, B. Zou, C. Li, K. Zhao, Q. Liu, and H. Chen, "Crown: a context-aware recommender for web news," in *2015 IEEE 31st International Conference on Data Engineering*, pp. 1420–1423, IEEE, 2015.

[50] S. Okura, Y. Tagami, S. Ono, and A. Tajima, "Embedding-based news recommendation for millions of users," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1933–1942, 2017.

[51] K. Park, J. Lee, and J. Choi, "Deep neural networks for news recommendations," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pp. 2255–2258, 2017.

[52] X. Wang, L. Yu, K. Ren, G. Tao, W. Zhang, Y. Yu, and J. Wang, "Dynamic attention deep model for article recommendation by learning human editors' demonstration," in *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, pp. 2051–2059, 2017.

[53] D. Khattar, V. Kumar, V. Varma, and M. Gupta, "Weave&rec: A word embedding based 3-d convolutional network for news recommendation," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pp. 1855–1858, 2018.

[54] D. Khattar, V. Kumar, V. Varma, and M. Gupta, "Hram: A hybrid recurrent attention machine for news recommendation," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pp. 1619–1622, 2018.

[55] H. Wang, F. Zhang, X. Xie, and M. Guo, "Dkn: Deep knowledge-aware network for news recommendation," in *Proceedings of the 2018 world wide web conference*, pp. 1835–1844, 2018.

[56] L. Zhang, P. Liu, and J. A. Gulla, "A deep joint network for session-based news recommendations with contextual augmentation," in *Proceedings of the 29th on Hypertext and Social Media*, pp. 201–209, 2018.

[57] C. Wu, F. Wu, M. An, J. Huang, Y. Huang, and X. Xie, "Npa: neural news recommendation with personalized attention," in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 2576–2584, 2019.

[58] Q. Zhu, X. Zhou, Z. Song, J. Tan, and L. Guo, "Dan: Deep attention neural network for news recommendation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 5973–5980, 2019.

[59] C. Wu, F. Wu, M. An, Y. Huang, and X. Xie, "Neural news recommendation with topic-aware news representation," in *Proceedings of the 57th Annual meeting of the association for computational linguistics*, pp. 1154–1159, 2019.

[60] C. Wu, F. Wu, S. Ge, T. Qi, Y. Huang, and X. Xie, "Neural news recommendation with multi-head self-attention," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, (Hong Kong, China), pp. 6389–6394, Association for Computational Linguistics, Nov. 2019.

[61] C. Wu, F. Wu, M. An, J. Huang, Y. Huang, and X. Xie, "Neural news recommendation with attentive multi-view learning," *arXiv preprint arXiv:1907.05576*, 2019.

[62] B. Koo, H. Jeon, and U. Kang, "Accurate news recommendation coalescing personal and global temporal preferences," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 78–90, Springer, 2020.

[63] H. Wang, F. Wu, Z. Liu, and X. Xie, "Fine-grained interest matching for neural news recommendation," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 836–845, 2020.

[64] D. Liu, J. Lian, S. Wang, Y. Qiao, J.-H. Chen, G. Sun, and X. Xie, "Kred: Knowledge-aware document representation for news recommendations," in *Fourteenth ACM Conference on Recommender Systems*, pp. 200–209, 2020.

[65] C. Wu, F. Wu, Y. Huang, and X. Xie, "Neural news recommendation with negative feedback," *CCF Transactions on Pervasive Computing and Interaction*, vol. 2, no. 3, pp. 178–188, 2020.

[66] L. Hu, C. Li, C. Shi, C. Yang, and C. Shao, "Graph neural news recommendation with long-term and short-term interest modeling," *Information Processing & Management*, vol. 57, no. 2, p. 102142, 2020.

[67] C. Wu, F. Wu, T. Qi, and Y. Huang, "User modeling with click preference and reading satisfaction for news recommendation.," in *IJCAI*, pp. 3023–3029, 2020.

[68] J. Xun, S. Zhang, Z. Zhao, J. Zhu, Q. Zhang, J. Li, X. He, X. He, T.-S. Chua, and F. Wu, "Why do we click: Visual impression-aware news recommendation," in *Proceedings of the 29th ACM International Conference on Multimedia*, pp. 3881–3890, 2021.

[69] C. Wu, F. Wu, Y. Huang, and X. Xie, "User-as-graph: User modeling with heterogeneous graph pooling for news recommendation," in *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, pp. 1624–1630, 2021.

[70] C. Wu, F. Wu, T. Qi, and Y. Huang, "Empowering news recommendation with pre-trained language models," in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1652–1656, 2021.

[71] G. Zheng, F. Zhang, Z. Zheng, Y. Xiang, N. J. Yuan, X. Xie, and Z. Li, "Drn: A deep reinforcement learning framework for news recommendation," in *Proceedings of the 2018 World Wide Web Conference*, pp. 167–176, 2018.

[72] J. Lian, F. Zhang, X. Xie, and G. Sun, "Towards better representation learning for personalized news recommendation: a multi-channel deep fusion approach.," in *IJCAI*, pp. 3805–3811, 2018.

[73] M. An, F. Wu, C. Wu, K. Zhang, Z. Liu, and X. Xie, "Neural news recommendation with long-and short-term user representations," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 336–345, 2019.

[74] Y. Tian, Y. Yang, X. Ren, P. Wang, F. Wu, Q. Wang, and C. Li, "Joint knowledge pruning and recurrent graph convolution for news recommendation," in

Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 51–60, 2021.

[75] Y. Sun, F. Yi, C. Zeng, B. Li, P. He, J. Qiao, and Y. Zhou, "A hybrid approach to news recommendation based on knowledge graph and long short-term user preferences," in *2021 IEEE International Conference on Services Computing (SCC)*, pp. 165–173, IEEE, 2021.

[76] I. Katakis, G. Tsoumakas, E. Banos, N. Bassiliades, and I. Vlahavas, "An adaptive personalized news dissemination system," *Journal of intelligent information systems*, vol. 32, no. 2, pp. 191–212, 2009.

[77] A. Patankar, J. Bose, and H. Khanna, "A bias aware news recommendation system," in *2019 IEEE 13th International Conference on Semantic Computing (ICSC)*, pp. 232–238, IEEE, 2019.

요 약

각각의 고유한 취향과 흥미를 가지고있는 셀 수 없이 많은 수의 사용자들에게 그들이 원하는 뉴스 기사를 어떻게 효과적으로 추천해줄 수 있을까? 뉴스 추천 시스템은 사용자 개개인이 온라인 뉴스 기사를 열람하면서 남기는 다양한 정보들을 통하여 관찰되지 않은 사용자의 행동 정보를 예측하는 것을 목표로 한다. 사용자에 대한 총체적인 프로필을 생성하려면 필연적으로 서로 다른 여러 도메인의 데이터들을 파고들어야 하므로, 추천이라는 태스크에 있어 잠재 표현을 학습하는 일은 그 가치를 지닌다. 그럼에도 불구하고, 사용자의 흥미는 시간에 따라 변모할 가능성이 있고 입력값 데이터는 산발적으로 분포되어있기 때문에 사용자 모델링은 복잡성을 특징으로 가진다.

이 문제에 대하여, 본 논문에서는 세 가지 분류 기준을 통하여 기존의 뉴스 추천 모델들을 심층 분석한다. 본 논문의 주요한 의의는 다음과 같다. 첫째, 현존하는 뉴스 추천 모델들을 뉴스 아이템의 특성을 고려하여 고안한 세 가지 기준으로 분류한다. 둘째, 30년간의 뉴스 추천 논문을 총망라하여 그 추이와 시대 별 해결하고자 한 문제가 무엇인지 살펴본다.

주요어 : 추천 시스템, 뉴스, 맥락 데이터

학번 : 2020-21659