结合了kG和GCN的session-based新闻推荐系统，将语义层级的信息通过KG建模与CNN提取出来的新闻文本向量结合，形成语义信息的新闻表示，再通过GCN进一步优化精炼表示，后通过RNN得到session的表达，最后预测输出。虽然运用了较新的图方法但是其本质提升在实验指标不是很大。

Context-aware Graph Embedding for Session-based News Recommendation

Heng-Shiou Sheu
University of Georgia
United States
hs06573@uga.edu

Sheng Li
University of Georgia
United States
sheng.li@uga.edu

ABSTRACT

Online news recommender systems aim to make personalized recommendations according to user preferences, which require modeling users' short-term reading interest. However, due to the limited logged user interactions in practice, news recommendation at session-level becomes very challenging. Existing methods on session-based news recommendation mainly focus on extracting features from news articles and sequential user-item interactions, but they usually ignore the semantic-level structural information among news articles and do not explore external knowledge sources. In this paper, we propose a novel Context-Aware Graph Embedding (CAGE) framework for session-based news recommendation, which builds an auxiliary knowledge graph to enrich the semantic meaning of entities involved in articles, and further refines the article embeddings by graph convolutional networks. Experimental results on a real-world news dataset demonstrate the effectiveness of our method compared with the state-of-the-art methods on session-based news recommendation.

CCS CONCEPTS

• **Information systems** → **Collaborative filtering**; • **Computing methodologies** → *Machine learning algorithms.*

KEYWORDS

Session-based News Recommendation, Graph Embedding

ACM Reference Format:
Heng-Shiou Sheu and Sheng Li. 2020. Context-aware Graph Embedding for Session-based News Recommendation. In *Fourteenth ACM Conference on Recommender Systems (RecSys '20), September 22–26, 2020, Virtual Event, Brazil.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3383313.3418477

1 INTRODUCTION

Recommendation system has become a fundamental technique which selectively recommends highly relevant information to users [2, 13, 15, 16, 26], and it has been widely deployed in many domains, such as search engine, online shopping and streaming service. One of the remarkable applications is news recommendation, which aims at recommending a set of news articles that fit user's preference and interest [26]. Many classical recommendation methods have been applied to news recommendation, such as the content-based methods [26] and collaborative filtering methods [2]. News recommendation has been recognized as a challenging problem because of several remarkable characteristics [3]. For instance, news articles are usually expired quickly after a short time period, and most users are not logged in the applications. Thus, short-term interested concepts have to be modeled by exploiting a few user-article interactions, leading to the session-based recommendation problem [6].

RecSys '20, September 22–26, 2020, Virtual Event, Brazil

ACM ISBN 978-1-4503-7583-2/20/09.
https://doi.org/10.1145/3383313.3418477

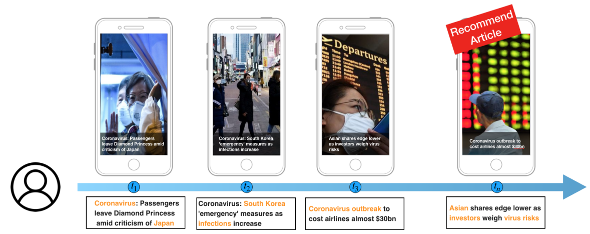

Figure 1: Illustration of session-based recommendation. Exploiting similarities among articles helps personalize recommendations.

The goal of session-based recommendation is to recommend a user the next possible interested item based on sequential information within a short time period [6]. Some session-based recommendation methods have been proposed by exploiting recurrent neural networks (RNN) [5, 6] to model the sequential information. These methods demonstrate that modeling sessions with deep sequential learning could greatly benefit the recommendation tasks. Recently, session-based news recommendation methods [3, 20] have been proposed, which extract textual features from news articles and also take sequential information into account.

Existing work on session-based news recommendation could extract expressive features (i.e., embeddings) from articles and sessions, but they usually ignore the semantic-level structure information of articles. We argue that, by discovering the semantic similarities among articles, the article embeddings could be enriched, which will facilitate the recommendation of highly relevant articles to users within a session. The research challenges are: *How to extract semantically meaningful embeddings for news articles? How to uncover the structural information among news articles?*

To address the above research challenges, we propose a novel context-aware graph embedding (CAGE) framework for session-based news recommendation. CAGE enriches the embeddings of news articles by exploiting an auxiliary knowledge graph. Moreover, it uncovers the structural information and models the similarities among articles using graph convolutional networks, as illustrated in Figure 1. Experimental results on a real-world benchmark dataset demonstrate the effectiveness of our approach.

2 RELATED WORK

Session-based recommendation aims at suggesting a set of highly relevant items to users, by modeling the sequential information based on anonymous user preferences in a short time period. It is an important research topic in user modeling [17, 23]. Recently, several session-based recommendation methods have been proposed by modeling click sequences with deep neural network [19]. Hidasi et al. [6] specifically used the gated recurrent units (GRUs) for prediction within longer sessions. Tan et al. [22] proposed a data augmentation technique and took shifts of user preferences into account. Liu et al. [18] proposed an novel attention priority model to efficiently capture user's long-term interests and short-term interests. Wu et al. [30] modeled session sequences as session-graph to obtain item embeddings and take complex transitions into account. Moreira et al. [3] proposed Chameleon, a deep learning meta-architecture for session-based news recommendation, by extracting textual features from news article and taking sequential information into account. Incorporation knowledge graph as an additional semantic-level knowledge source to improve the recommendation performance has been an emerging research topic. The collaborative knowledge base embedding (CKE) method leverages knowledge graphs to learn the item representations [31]. Huang et al. [7] integrated key-value memory network to RNN-based model for sequential recommendation. Wang et al. [26] proposed a knowledge-aware neural network, which fuses knowledge-level and semantic-level information representations for news recommendation. Wang et al. [28] leveraged path representation by fusing knowledge-level representations of the entities and relationships as sequential information for recommendation. Some recent methods also incorporate graph-structure networks with attention mechanisms [27].

Different from existing work, our approach exploits knowledge graphs to enrich the meaning of the entities in articles, refines the article embeddings by graph convolutional networks, and employs a recurrent neural network with gated recurrent units to learn session embeddings.

3 PROPOSED FRAMEWORK

3.1 Problem Statement

Let $U = \{u_1, u_2, u_3, ..., u_{|U|}\}$ denote a set of unique articles involved in all the sessions, where $|U|$ is the total number of all unique articles. An anonymous session sequence s can be represented by a list $s = [u_{s,1}, u_{s,2}, ..., u_{s,|s|}]$ ordered by timestamps, where $u_{s,i} \in V$ represents a clicked article of the user within the session s, and $|s|$ is the total number of clicked articles within the session s.

In addition to the user-item interactions, we consider semantic-level structural information extracted from an auxiliary knowledge graph defined as KG. KG is a directed graph composed of ⟨subject-property-object⟩ triple facts. Formally, a triple is represented as (e_h, e_t, r), where $e_h, e_t \in \varepsilon$ are entities, and $r \in R$ are relations. Each triple indicates that there exists a relationship r from the head entity e_h to the tail entity e_t.

The task of session-based news recommendation is to predict whether a user will read a news article based on previous user-article interactions within a short time period. Formally, given a set of unique articles U, a number of sessions s with timestamps from 1 to t, and an auxiliary knowledge graph KG, we aim to recommend a list of K articles $\hat{y} = [\hat{y}_1, \hat{y}_2, ..., \hat{y}_K]$ for each session at timestamp $t + 1$, where $1 \leq K \leq |U|$.

3.2 Context-Aware Graph Embedding

We propose a context-aware graph embedding (CAGE) framework for session-based news recommendation, as shown in Figure 2. First, we extract entities from news articles by a pre-trained NLP model. In order to build a sub-knowledge graph, we extract triples from an open knowledge graph based on the discovered entities. Second, we obtain article content embeddings through word embedding and convolutional neural networks. Then, we concatenate these features as article embeddings, and employ graph convolutional networks to further refine them. And we apply RNN to obtain the session-level embeddings. In the end, we calculate the similarity between session embeddings and candidate articles.

3.2.1 Textual-Level Article Embedding. Each article u can be represented as a matrix $\mathbf{V}^u_{1:n}$ consisting of word vectors. To obtain textual-level article embedding, we define a word embedding function $\phi : M \to R^d$, which maps M sized vocabulary in the article u into a d dimensional vector. The word embedding can be any pre-trained word embedding models like word2Vec, Doc2Vec, or Glove. Then, $\mathbf{V}^u_{1:n}$ are represented as: $\mathbf{V}^u_{1:n} = \phi(w^u_1) \oplus \phi(w^u_2) ... \oplus \phi(w^u_n)$, where w^u_k indicates the k-th word in the article u, $\phi(w^u_k)$ is the embedding of w^u_k, and $\oplus$ is the concatenation operation.

We then design a convolutional neural network to learn the textual-level article embeddings based on word vectors $\mathbf{V}^u_{1:n}$. The feature is represented as: $z_j = \sigma(\mathbf{V}^u_{1:n} * k_j + b_j)$, where σ and $*$ denote the activation function and convolution operation, respectively. Finally, the textual-level article embeddings are denoted by

$$\widetilde{\mathbf{V}}^u_{1:n} = \sigma(W \times (z_1 \oplus z_2 \oplus, ..., \oplus z_m) + b), \tag{1}$$

where $W \in R^{m \times n}$ is a weighted matrix, and $b \in R^n$ denotes bias.

3.2.2 Semantic-Level Article Embedding with Sub-Knowledge Graph. In this step, we represent semantic-level embedding for each article with the help of an open knowledge graph, Wikidata [25]. First, to disambiguate the mentioned word w^u_k in knowledge graph KG, we adopt entity-linking to distinguish the meaning in the news article by connecting them with existing entities in KG. Next, based on the identified entities, we extract all triples from the original knowledge graph and construct a sub-knowledge graph G_S. To overcome the sparsity issue and the lack of connections among the identified entities, we expand G_S within one hop of identified entities. Then, we apply TransE [1], a knowledge graph embedding method, to learn a low-dimensional vector for each entity e^u_i in G_S. Finally, we will have a sub-knowledge graph G_S contains all identified entities connected in meaning. Thus, for each article, we will obtain the semantic-level embedding $\widetilde{\mathbf{e}}^u_{1:n}$ by:

$$\widetilde{\mathbf{e}}^u_{1:n} = \psi(e^u_1) \oplus \psi(e^u_2) \oplus ... \oplus \psi(e^u_n), \tag{2}$$

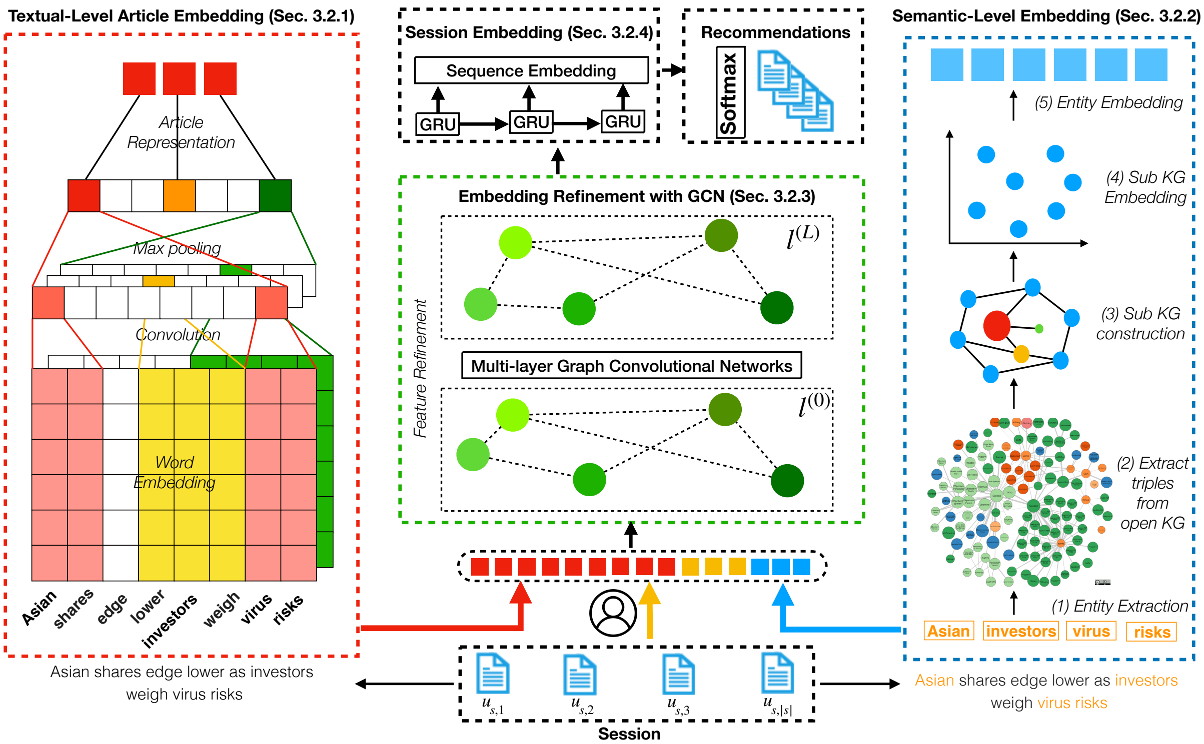

Figure 2: Illustration of our CAGE framework. CAGE extracts textual-level features from news articles with convolutional neural networks (Section 3.2.1) and meanwhile represents semantic-level entities with the help of open knowledge graph (Section 3.2.2). The procedures of extracting semantic-level embedding contains five steps, including entity extraction, triplets extraction, sub-knowledge graph (KG) construction, sub-KG embedding, and entity embedding concatenation. Then, CAGE refines the concatenated embeddings through multi-layer graph convolutional networks (Section 3.2.3). After that, session-level representations are learned by a GRU-based RNN model (Section 3.2.4). Finally, CAGE predicts the next-click article, i.e., the recommendation, for each session.

where e_k^u denotes k-th entity in G_S and $\psi(\cdot)$ denotes the graph embedding model.

3.2.3 Refining Article Embeddings with Graph Neural Networks. The obtained textual-level and semantic-level article embeddings can be integrated with other available side information such as user features. We denote the concatenated article embedding as: $\widetilde{\mathbf{U}} = \widetilde{\mathbf{V}}^u_{1:n} \oplus \widetilde{\mathbf{e}}^u_{1:n} \oplus \mathbf{p}$, where $\mathbf{p}$ represents an one-hot encoding vector of user attributes.

For session-based news recommendation, it is critical to exploit the neighborhood structural information among articles, which would help enrich the article embeddings. For instance, articles with similar concepts shall be close in the embedding space. To this end, we propose to construct article-level graphs and employ graph neural networks [10, 14] to further refine the article embeddings. Articles are nodes on the graph, while the pair-wise similarity values of the article embeddings are weights on edges. We remove the edges with small similarity values (e.g., cosine similarity values that are less than 0.80) and obtain a sparse graph. Then, we use a two-layer graph convolutional network (GCN) [14] for session s:

$$\mathbf{H}^{(l+1)} = \sigma(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}\mathbf{H}^{(l)}W^l) \tag{3}$$

where $\mathbf{H}^{(l)} \in R^{N\times D}$ denotes the input of the l-th layer, $\hat{A}$ is an adjacency matrix with identity matrix, and $\hat{D}$ is a degree matrix. W^l presents trainable variable for the l-th layer. σ presents activation function (e.g. *ReLU*). In our approach, we use $\mathbf{X} = \mathbf{H}^{(2)}$ as the refined article embedding.

3.2.4 Session Embeddings with Gated Recurrent Units. We design a gated recurrent units (GRU)-based recurrent neural network to model the user-article sequential interactions in sessions. GRU is adopted as it can deal with the vanishing gradient problem. We use h_t to denote the session embedding, which is the final hidden state in GRU network. $\hat{h}_t$ is the final session embedding.

3.2.5 Session-based News Article Recommendation. Because of the dynamic nature of news, news articles will be produced and expired rapidly. For this reason, the proposed CAGE framework is trained to maximize the similarity between user's session and the next-click article actually read by user (i.e., positive sample), and minimize the similarity between session and articles that are not read by user (i.e., negative samples). Furthermore, we refer the positive sample and negative samples as $item^+$ and ($item^- \in U^-$), where U^- is a set of articles not read by user in the session. This strategy is also adopted in [8, 20]. We then define the semantic relevance score function R as: $\mathrm{Rel}(session, item) = \varphi(h_t \odot item)$, where $item$ denotes a candidate article from the union of $item^+$ with $item^-$, and φ denotes a neural network with several fully connected layers, and $\odot$ denotes the element-wise product.

In particular, we denote a probability function to evaluate the posterior probability of clicking a candidate news article given the current session embedding $\hat{h}_t$:

$$P(item^+|\hat{h}_t) = \frac{\exp(\gamma \mathrm{Rel}(\hat{h}_t, item^+))}{\sum_{\forall item \in U'} \exp(\gamma \mathrm{Rel}(\hat{h}_t, item))}, \quad (4)$$

where γ is a smoothing factor and U' is the union of $item^+$ and U^-.

In the training stage, our approach maximizes the likelihood of the clicked article given sequential interactions in the current session. Equivalently, we aim to minimize the following loss function:

$$L(\theta) = -\log \prod_{(h_t, item^+)} P(item^+|h_t), \quad (5)$$

where θ denotes the model parameters. Since $L(\theta)$ is differentiable w.r.t to θ, the proposed CAGE framework can be trained using gradient descent based optimization algorithms.

4 EXPERIMENTS

Dataset. We evaluate our framework on a real-world news benchmark dataset, Adressa [4]. Adressa was collected by the Norwegian University of Science and Technology and Adressavisen (a local newspaper in Trondheim, Norway). Specifically, we employ a subsets that contains 16 days of news data for evaluation. This subset is also commonly used in existing work [3, 20]. In addition, we use the open knowledge graph, Wikidata, as the external knowledge. Wikidata is the crowd-sourced knowledge database based on Wikipedia.

Baselines. To evaluate our proposed model, we use the following traditional methods and state-of-the-art methods as our baselines: (1) *Content-Based (CB)*. This method compares the cosine similarity of the article embedding to recommended articles. (2) *Recently Popular (RP)*. It recommends the most viewed articles from the last N user interactions. (3) *Item-kNN*. It employs the cosine similarity to return the most similar items [19]. (4) *Vector Multiplication Session-Based kNN (V-skNN)* [19]. (5) *Sequential Rules (SR)*. It's a variant of associated rule method [12, 19]. (6) *Co-occurrence (CO)*. It recommends the most read articles [11, 19]. (7) *GRU4Rec* [6]. It's the first RNN-based approach for session-based recommendation. In the experiment, we apply the improved version, GRU4RecV2 [5]. (8) *SR-GNN* [30]. It models items within the session as session-graph for recommendation. (9) Chameleon [3]. It's the state-of-the-art method for session-based news recommendation.

Evaluation Metrics. We employ the Hit Rate (HR), Mean Reciprocal Rank (MRR) and Normalized Discounted Cumulative Gain (NDCG) are our primary evaluation metrics. Moreover, to evaluate the novelty and diversity of recommendations, the Expected Self-Information with Rank and Relevance-sensitivity (ESI-RR) metric and the Expected Intra-List Diversity with Rank and Relevance-sensitivity (EILD-RR) metric are also adopted [20, 24]. Following evaluation protocols in existing work [3], we average all results after training on 5-hour data and evaluating on 1-hour data.

Parameter setting. For our framework, we select the dimension of the knowledge graph embedding from {50, 100}, GCN layer1 output from {250, 150, 125, 100}, GCN layer2 output from {150, 70, 60, 50}, and batch size from { 32, 64, 128, 256}. Using a validation set, we finally set the dimensions for GCN layer1, layer2 and the batch size to 125, 60, and 64, respectively. The key parameter settings for baselines are same as configurations reported in Chameleon [3].

Result and Analysis. We show the results of HR@10, MRR@10, NDCG@10, novelty and diversity on the Adressa dataset in Table 1. We summarize the observations as below. First, non-personalized methods, RP and CB, are both obtained poor performance without considering user preferences. Neighborhood-based methods, v-skNN and item-kNN, outperform non-personalized methods by taking similar articles into account. Methods based on association rules, SR and CO, perform better than previous strategies. Chameleon further improves the performance by using neural networks. It shows the advantage of using deep models for session-based news recommendation. As GRU4Rec and SR-GNN are not designed to make recommendations for first seen items during training, they cannot outperform Chameleon. Compared with Chameleon, CAGE achieves better performance on HR@10, MRR@10, and NDCG@10 owing to the semantic-level information and the embedding refinement. As for the novelty and diversity, our CAGE method obtains better performance that its competitors.

Ablation Studies. To investigate the effectiveness of different components in our CAGE framework, we conduct ablation studies and report the results in Table 2. We evaluated the model performance without KG, with different embedding dimensions, and with different graph embedding methods such as TransH [29], TransD [9] and RotatE [21]. The results demonstrate the effectiveness of the proposed CAGE framework.

5 CONCLUSIONS

In this paper, we proposed a context-aware graph embedding framework for session-based news recommendation by considering structured information and refining article embeddings with graph neural networks. We built an auxiliary knowledge graph for capturing semantic-level information to enrich the semantic meaning of entities involved in articles. Extensive experiments on a real-word news dataset demonstrated the effectiveness of the proposed CAGE framework, compared with the state-of-the-art methods on session-based news recommendation.

REFERENCES

[1] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. In *Advances in neural information processing systems*. 2787–2795.

[2] Abhinandan S Das, Mayur Datar, Ashutosh Garg, and Shyam Rajaram. 2007. Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web*. 271–280.

[3] Gabriel de Souza Pereira Moreira, Felipe Ferreira, and Adilson Marques da Cunha. 2018. News session-based recommendations using deep neural networks. In *The 3rd Workshop on Deep Learning for Recommender Systems*. 15–23.

[4] Jon Atle Gulla, Lemei Zhang, Peng Liu, Özlem Özgöbek, and Xiaomeng Su. 2017. The Adressa dataset for news recommendation. In *Proceedings of the international conference on web intelligence*. 1042–1048.

[5] Balázs Hidasi and Alexandros Karatzoglou. 2018. Recurrent neural networks with top-k gains for session-based recommendations. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. 843–852.

[6] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2015. Session-based recommendations with recurrent neural networks. *arXiv*

Table 1: Performance comparison of our framework and baselines (shown in different categories) on Adressa dataset.

Algorithms		Common			Novelty	Diversity
		HR@10	MRR@10	NDCG@10	ESI-RR@10	EILD-RR@10
Context-Aware	CAGE	**70.252**	**34.268**	**42.741**	**60.137**	**1.843**
Deep Learning	Chameleon	70.111	34.211	42.665	60.007	1.838
	SR-GNN	61.296	30.102	39.857	51.340	1.827
	GRU4Rec	48.771	22.100	29.686	37.490	1.416
Association Rules-based	SR	63.358	30.616	38.323	52.661	1.743
	CO	62.332	28.554	36.490	47.376	1.664
Neighborhood based	v-skNN	61.955	27.684	35.733	50.954	1.681
	item-kNN	61.864	28.010	35.946	41.501	1.619
Non-personalized	RP	57.072	25.096	32.584	39.425	0.436
	CB	32.533	11.964	16.718	35.038	1.493

Table 2: Ablation studies of our method on Adressa dataset.

Variants	Common		Novelty	Diversity
	MRR@10	NDCG@10	ESI-RR@10	EILD-RR@10
CAGE (TransE)	34.268	**42.741**	**60.137**	1.843
CAGE w/o KG	34.251	42.710	59.984	1.831
CAGE + dim = 20	34.304	42.733	59.924	**1.844**
CAGE + dim = 50	**34.279**	42.720	60.075	1.839
CAGE + dim = 200	34.210	42.654	59.995	1.840
CAGE + TransH	34.160	42.635	59.907	1.839
CAGE + TransD	34.214	42.703	60.107	1.843
CAGE + RotatE	34.182	42.595	59.907	1.837

preprint arXiv:1511.06939 (2015).
[7] Jin Huang, Wayne Xin Zhao, Hongjian Dou, Ji-Rong Wen, and Edward Y Chang. 2018. Improving sequential recommendation with knowledge-enhanced memory networks. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 505–514.
[8] Po-Sen Huang, Xiaodong He, Jianfeng Gao, Li Deng, Alex Acero, and Larry Heck. 2013. Learning deep structured semantic models for web search using clickthrough data. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*. 2333–2338.
[9] Guoliang Ji, Shizhu He, Liheng Xu, Kang Liu, and Jun Zhao. 2015. Knowledge Graph Embedding via Dynamic Mapping Matrix. In *ACL*. 687–696.
[10] Xiaodong Jiang, Pengsheng Ji, and Sheng Li. 2019. CensNet: Convolution with Edge-Node Switching in Graph Neural Networks.. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*. 2656–2662.
[11] Michael Jugovac, Dietmar Jannach, and Mozhgan Karimi. 2018. Streamingrec: a framework for benchmarking stream-based news recommenders. In *Proceedings of the 12th ACM Conference on Recommender Systems*. 269–273.
[12] Iman Kamehkhosh, Dietmar Jannach, and Malte Ludewig. 2017. A Comparison of Frequent Pattern Techniques and a Deep Learning Method for Session-Based Recommendation.. In *RecTemp@ RecSys*. 50–56.
[13] Donghyun Kim, Sungchul Kim, Handong Zhao, Sheng Li, Ryan A Rossi, and Eunyee Koh. 2019. Domain switch-aware holistic recurrent neural network for modeling multi-domain user behavior. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*. 663–671.
[14] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).
[15] Sheng Li, Jaya Kawale, and Yun Fu. 2015. Deep collaborative filtering via marginalized denoising auto-encoder. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. 811–820.
[16] Sheng Li, Jaya Kawale, and Yun Fu. 2015. Predicting user behavior in display advertising via dynamic collective matrix factorization. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 875–878.
[17] Sheng Li and Handong Zhao. 2020. A Survey on Representation Learning for User Modeling. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*. 4997–5003.
[18] Qiao Liu, Yifu Zeng, Refuoe Mokhosi, and Haibin Zhang. 2018. STAMP: short-term attention/memory priority model for session-based recommendation. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1831–1839.
[19] Malte Ludewig and Dietmar Jannach. 2018. Evaluation of session-based recommendation algorithms. *User Modeling and User-Adapted Interaction* 28, 4-5 (2018), 331–390.
[20] Gabriel de Souza Pereira Moreira, Dietmar Jannach, and Adilson Marques da Cunha. 2019. Contextual hybrid session-based news recommendation with recurrent neural networks. *arXiv preprint arXiv:1904.10367* (2019).
[21] Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang. 2019. Rotate: Knowledge graph embedding by relational rotation in complex space. *arXiv preprint arXiv:1902.10197* (2019).
[22] Yong Kiam Tan, Xinxing Xu, and Yong Liu. 2016. Improved recurrent neural networks for session-based recommendations. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*. 17–22.
[23] Zhiqiang Tao, Sheng Li, Zhaowen Wang, Chen Fang, Longqi Yang, Handong Zhao, and Yun Fu. 2019. Log2Intent: Towards interpretable user modeling via recurrent semantics memory unit. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1055–1063.
[24] Saúl Vargas and Pablo Castells. 2011. Rank and Relevance in Novelty and Diversity Metrics for Recommender Systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems* (Chicago, Illinois, USA) *(RecSys '11)*. Association for Computing Machinery, New York, NY, USA, 109–116. https://doi.org/10.1145/2043932.2043955
[25] Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Commun. ACM* 57, 10 (2014), 78–85.
[26] Hongwei Wang, Fuzheng Zhang, Xing Xie, and Minyi Guo. 2018. DKN: Deep Knowledge-Aware Network for News Recommendation. In *Proceedings of the 2018 World Wide Web Conference* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1835–1844.
[27] Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. 2019. Kgat: Knowledge graph attention network for recommendation. In *SIGKDD*. 950–958.
[28] Xiang Wang, Dingxian Wang, Canran Xu, Xiangnan He, Yixin Cao, and Tat-Seng Chua. 2019. Explainable reasoning over knowledge graphs for recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 5329–5336.

[29] Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. 2014. Knowledge graph embedding by translating on hyperplanes. In *Twenty-Eighth AAAI conference on artificial intelligence*.

[30] Shu Wu, Yuyuan Tang, Yanqiao Zhu, Liang Wang, Xing Xie, and Tieniu Tan. 2019. Session-based recommendation with graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 346–353.

[31] Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, and Wei-Ying Ma. 2016. Collaborative knowledge base embedding for recommender systems. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 353–362.