Smart Targeting: A Relevance-driven and Configurable Targeting Framework for Advertising System

Yong Li, Zihao Zhao, Zhiwei Fang, Kui Ma, Yafei Yao, Changping Peng, Yongjun Bao, Weipeng Yan
{liyong5,zhaozihao3,fangzhiwei2,makui,yaoyafei1,pengchangping,baoyongjun,paul.yan}@jd.com
Business Growth BU, JD.com

ABSTRACT

Targeting system is an essential part of computational advertising. It allows advertisers to select and reach their targeted users. Due to various advertising goals and the demand for making budget plans, advertisers have a strong will to configure the final targeting results, or they can become very cautious in spending money on advertising campaigns. Meanwhile, to guarantee the advertising performance, the targeted users should also be relevant to the ads of the advertisers. Recent targeting methods are mainly based on tags produced by the Data Management Platform (DMP) which is easy for the advertisers to configure the targeting results. However, in such methods, the relevance between the targeted users and ads is not technically evaluated and cannot be guaranteed. The biggest challenge is that it is hard for a machine learning model to both model the relevance and take account of the advertiser's configuration demands. In this paper, we propose a novel relevance-driven and configurable targeting framework called Smart Targeting to solve the problem. Specifically, different from Tag-wise Targeting, we first use a relevance model to retrieve the most relevant users for the ads. To further enable the advertisers to configure the final results, we develop a Delay Intervention Mechanism to leverage the power of DMP. As far as we know, this is the first attempt of combining relevance modeling and advertiser intervention into a unified targeting system. We implement and evaluate our framework on JD.com platform with over 300 million users and the results show that it can bring significant improvements to the core indicators such as CTR and eCPM. The long term monitoring also demonstrates that Smart Targeting gradually becomes the most popular targeting tool after its release.

CCS CONCEPTS

• **Information systems** → **Computational advertising**; **Content match advertising**.

KEYWORDS

Targeted Advertising, Tag-wise Targeting, Smart Targeting

ACM Reference Format:
Yong Li, Zihao Zhao, Zhiwei Fang, Kui Ma, Yafei Yao, Changping Peng, Yongjun Bao, Weipeng Yan. 2020. Smart Targeting: A Relevance-driven and Configurable Targeting Framework for Advertising System. 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/10.1145/3383313.3418481

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

ACM ISBN 978-1-4503-7583-2/20/09.
https://doi.org/10.1145/3383313.3418481

1 INTRODUCTION

Online computational advertising contains three primary entities: ads, advertisers, and users. The advertisers spend money to bid for displaying their ads to their targeted users. Whether an ad can be finally displayed to a given user is determined by three elements: (1) whether the ad has a chance to bid for the display on the user; (2) how much money the advertiser plans to spend on this bidding; (3) the relevance between the ad and the user, i.e., the click-through-rate (CTR). And targeting system is used to deal with the problems of the first element, i.e., it allows and helps the advertisers to select and reach their targeted users on whom they are willing to spend money. Obviously, **targeting system is designed for advertisers instead of users** [12, 16]. Targeting is essential for most of the advertisers. The reason is two-fold: Firstly, most advertisers must make budget plans before they start a campaign and the main content of the budget plan is to configure the bids on different crowds with different user profiles; Secondly, the advertising goals are varied and different goals require different targeted users even when the ads are the same. For example, if an advertiser's goal is to improve the Return On Investment (ROI), he can select users who are his loyal customers. But if he wants to attract more new customers, then the targeted users can be these who have bought the competitive products. The two reasons above show that the targeted users for a given set of ads are not only related to the ads themselves but also deeply affected by the advertisers. This results in that allowing explicit advertiser intervention becomes a primary demand of the advertisers and an essential function of targeting systems [6, 7, 16].

The most widely used targeting system is Tag-wise Targeting (TT), shown in Figure 1 (a), where the advertisers conduct targeting by selecting tags of users. The tags are produced by a Data Management Platform (DMP) [13], where each user is profiled with a variety of statistical tags such as searching/clicking/purchasing, their education level, location, and so on, based on their historical behaviors and static attributes. Then for any combination of tags that are selected by the advertisers, there is a corresponding user group in DMP as the targeted users. Although TT can meet most of the targeting demands of advertisers, the biggest problem is that the relevance between the targeted users and ads of advertisers cannot be guaranteed, which can result in bad advertising performance. In TT, the item-user relevance is evaluated by the advertisers based on their experience and practice. This puts forward a very high demand for the advertisers on expertise if they want to achieve a good advertising result. However, for most of the

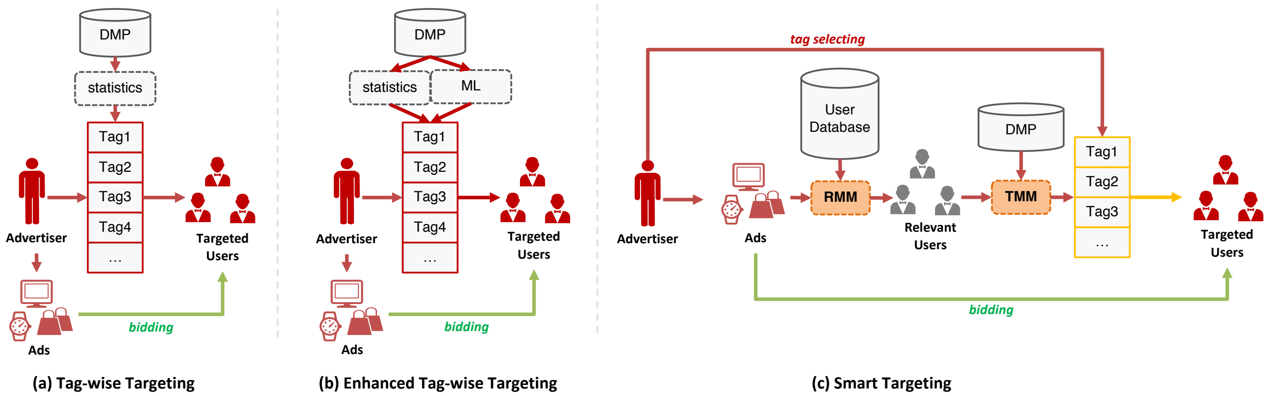

Figure 1: Visualization of Tag-wise Targeting and Smart Targeting. In Tag-wise Targeting, the advertisers select the targeted users by configuring the tags provided by DMP. In Smart Targeting, relevant users will first be retrieved by a Relevance Modeling Model (RMM) and then a delay intervention mechanism enable the advertisers to configure the final targeted users. TMM stands for the Tag Mining Model to product tags from the relevant users.

advertisers, such demand is too harsh and not realistic. Recently, using machine learning technics in recommender systems becomes a trend and shows great power [2, 14, 19]. Some works also attempt to introduce machine learning models into targeting but they still focus on mining more valuable and powerful tags in DMP such as social-network-based targeting [15] and advertising networks [4, 5, 11]. Such enhanced targeting can provide much more abundant tags than TT but shares a similar framework which doesn't fundamentally change the way to select targeted users, so we call it enhanced Tag-wise Targeting (eTT) as shown in Figure 1 (b).

In this paper, we propose a novel targeting framework referred to as Smart Targeting (ST) which can explicitly guarantee the relevance between the ads and the targeted users. The relevance will be technically evaluated by data-driven deep models and the most relevant users will be retrieved. But if we just regard those retrieved relevant users as the final targeted users, the demands of advertisers will not be satisfied. Because deep models are weak in interpretability, it is also hard to allow direct advertiser intervention on the models to achieve their goals. To deal with the challenge, we propose a delay intervention mechanism which re-mines tags from the retrieved users, and these tags are then provided for advertisers to conduct custom configuration on their final targeted users. As shown in Figure 1 (c), our Smart Targeting framework is very flexible. The selection of Relevance Modeling Model (RMM) is nearly arbitrary: most recommendation models can be directly applied such as FM-series [14, 19], deep-network-series [2, 3, 9, 22] and tree-based-series [23, 24]. And most models in DMP can also be reused in Tag Mining Model (TMM). Our contributions can be summarized as three-fold:

- We propose a novel targeting framework called Smart Targeting which can technically evaluate the relevance between targeted users and the ads and enable advertiser interventions on final targeting results, simultaneously. We show that Smart Targeting is very powerful and flexible. As far as we know, this is the first attempt of combining relevance modeling and advertiser intervention into a unified targeting system.
- We provide an implementation of Smart Targeting which is fast enough to serve online in real-time.
- We evaluate Smart Targeting on the advertising platform of JD.com, an online shopping mobile app with over 300 million active users. The results demonstrate that Smart Targeting can bring in significant improvement on the core indicators such as CTR and eCPM. And long term monitoring also shows that it gradually becomes the most popular targeting tool compared with Tag-wise Targeting.

2 BACKGROUND AND RELATED WORK

Advertising system is different from recommendation system although they both deeply rely on recommendation technology. The latter mainly considers the relationship between items and users while the former additionally takes account of the relationship between advertisers and users. Since the advertisers paid money for advertising, they have a strong will and demand to control and configure whom they plan to put the ads to, which is called targeting. The earliest discussions about targeting are from Google's patents [6, 10, 16]. In 2012, Yahoo released a research on the effectiveness of targeted advertising [7]. Then, Chen and Stallaert provided an economic analysis of targeting [1], showing that the revenue can be increased a lot when using targeted advertising. By now, the targeting system has become one of the most primary part in current big advertising platforms, such as Google, Amazon, Facebook, Alibaba, Baidu and JD.com. And the targeting methods are not unique: such as searching-keyword-based targeting [18], mobile-device-based targeting [20], location-based targeting [8], behavioral targeting [1], social-network-based targeting [15] and so on. However, to the best of our knowledge, those methods still serve in the tag-wise targeting framework in Figure 1 (a) and (b).

3 SMART TARGETING

3.1 Framework

Figure 1 (c) illustrates the framework of Smart Targeting (ST). First of all, the ads (or items) in an advertising campaign are determined at the very beginning. Then we develop a Relevance Modeling

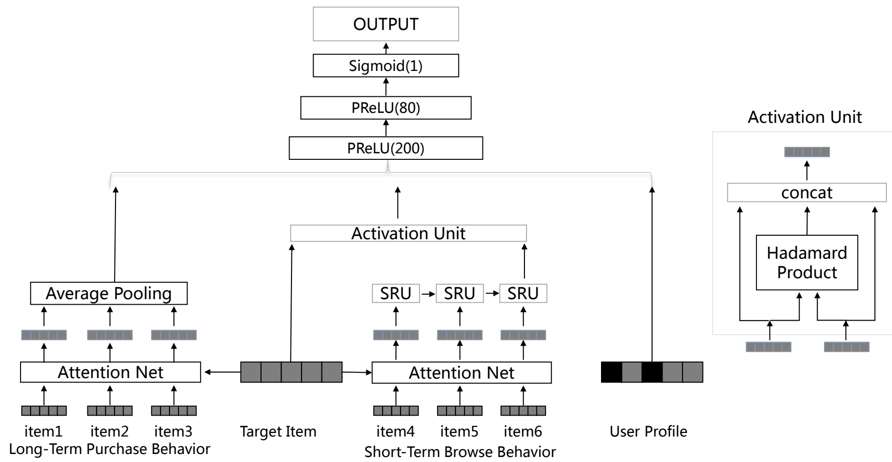

Figure 2: The Architecture of Long Short Behavior Model.

Model (RMM) to predict the relevance between each ad in the campaign and each user from the user database. If time-consuming is not concerned, the selection of RMM is nearly arbitrary. With the relevance scores, we can retrieve the top-K relevant users for each ad. After that, we can get a group of the most relevant users for the campaign. However, these retrieved users may not meet the demands of advertisers. For example, when the ads are some Coca Cola drinks, if the advertiser wants high ROI, the targeted users should be those users who usually buy Coca Cola drinks; while if his goal is to attract more new customers then he should put his ads to those who seldom buy Coca but usually buy competitive drinks such as Pepsi. To allow the advertisers to configure the targeted users, we mine tags again from the retrieved relevant users through Tag Mining Model (TMM) and provide the tags to the advertisers so that they can configure the final targeting results by selecting their interested tags. We call such mechanism Delay Intervention Mechanism compared with TT where intervention occurs at the very beginning. Finally, the advertisers can normally make the budget plans based on the selected tags and put the ads into the following bidding process.

Why Delay Intervention Mechanism. An intuitive question is that why should the intervention is delayed? Can we first conduct Tag-wise Targeting and then measure the relevance? We argue that delay intervention is a better choice. If intervention occurs in the first step, since relevance is not guaranteed in intervention, the number of final reserved users filtered by RMM may be too small to satisfy the advertisers' requirements. That's why we design the Delay Intervention Mechanism.

3.2 Relevance Modeling Model

As mentioned above, the choice of RMM is nearly arbitrary. In this paper, the design of RMM is mainly based on three concerns: (1) As long-term behaviors reflect the stable preferences of users which is very valuable information for targeting, we expect to model such stable preferences. (2) Real-time demand is another key description of user, we tend to model user demand through short-term behaviors. (3) We hope the model is fast enough so that it can satisfy the online serving requirement which means that the network should not be too complex. Based on the three concerns, we propose Long Short Behavior Model (LSBM) to learn the relevance between users and items.

As is shown in Figure 2, to model long-term interest of a user, these long-term purchase behaviors are firstly converted to embedding representations. Then, we introduce an attention net between the target item and the long-term representation to model the relevance between the target item and long-term interest. Following [21], we apply the attention net as

$$Attention(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = softmax(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}})\mathbf{V} \tag{1}$$

where $\mathbf{Q}$ is the embedding representation of target item, $\mathbf{K} = \mathbf{V}$ is the embedding of long-term purchase behaviors and d represents the dimension of embedding.

As the long-term purchase behaviors can be extremely long, we apply an average pooling operation to reduce the computational cost.

As for short-term demands, similarly, the embedding of short-time behaviors will be fed into an attention net with the target item, which selectively emphasizes specific behaviors and suppresses less useful ones for the given target item. Next, we apply SRU operation [17] to model the sequence information of the short-term behaviors. Finally, we combine the information of short-term behaviors and target item through an activation unit, which includes a Hadamard product and concatenation operation.

Besides, some of the user profile features are also concerned, together with the long/short-term features as inputs of LSBM, which could help to learn the potential information among these features. The predicted value $\hat{y}$ for a given sample is:

$$\hat{y} = \sigma(x) \tag{2}$$

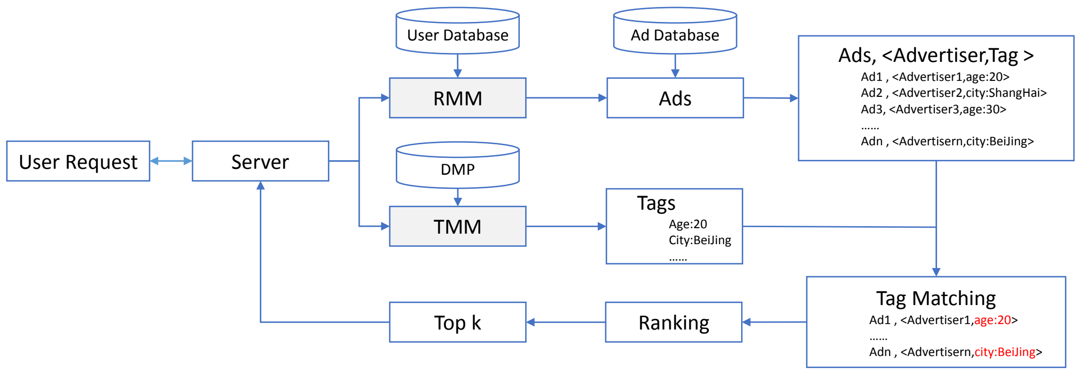

Figure 3: The online serving implementation of Smart Targeting. Different from the Smart Targeting framework in Figure 1, online serving process always starts from a coming user request instead of the advertisers, but the internal logics are the same.

where $\sigma(x) = 1/(1 + exp(-x))$ is the sigmoid function, x is the output of neural network, then the loss function is defined as:

$$loss = -\frac{1}{N}\sum_{i=1}^{N} y_i log(\hat{y}_i) + (1 - y_i) log(1 - \hat{y}_i) \quad (3)$$

where $y_i \in \{0, 1\}$ and $y_i = 1$ stand for click of the given sample i. N is the size of training dataset. Based on LSBM, we can model user static profiles and user interest from long/short-term preferences simultaneously.

3.3 Tag Mining Model

TMM is used to mine tags from the retrieved relevant users. Technically it can be any statistical model or machine learning model or a combination of them in DMP. Since the final implementation of Smart Targeting is expected to serve online and in real-time in this paper, we directly reuse the tags of DMP in TMM, i.e., all the DMP tags related to the relevant users are retrieved as the results of TMM. In this way, the major content of TMM is a look-in-table operation which is fast and efficient.

3.4 Online Serving

To provide real-time serving, we develop an implementation of Smart Targeting as shown in Figure 3. Different from the framework in Figure 1 (c) which starts from the advertisers, when targeting finally takes effects in the advertising recommender system, the beginnings are usually the user requests. Despite this, the internal logics are the same.

When a user u visits the website or app, the workflow of Smart Targeting is as follows:

- *Step1:* Based on the RMM, the most relevant top-M ads are retrieved as candidate ads. Notice that RMM is to model the relevance between users and ads, thus it can be used to not only retrieve relevant users for ads but also retrieve relevant ads for users. Here the RMM is exact our LSBM.
- *Step2:* From the candidate ads, we can get the advertiser of each ad through an inverted indexing operation and know their selected tags, marked as $[\mathrm{ad}_i, <\mathrm{Advertiser}_i, \mathrm{tags}_i>]$.
- *Step3:* Based on TMM, we can also get the tags on the user request.
- *Step4:* To judge whether the user is the targeted user of any advertisers, we conduct a matching operation between the tags of the user and the tags in each of $[\mathrm{ad}_i, <\mathrm{Advertiser}_i, \mathrm{tags}_i>]$.
- *Step5:* The matched ads are reserved and then fed to the following bidding, ranking and displaying stages.

Table 1: A/B test Results of ST+eTT vs online system with eTT

B	A	CTR Gain	CPC Gain	eCPM Gain
ST+eTT	Online	+4.62%	+4.09%	+8.77%

4 EXPERIMENT

To evaluate the effectiveness of Smart Targeting in the online advertising system, we run live A/B experiments on the advertising platform of JD.com[1] with over 300 million active users. The evaluation is based on a couple of aspects, including core indicators CTR(Click Through Rate), CPC (Cost Per Click), eCPM (Effective Cost per Mile), feedback from the advertisers and feedback from the users.

4.1 Experimental Results

We productized the Smart Targeting (ST), along with the online enhanced Tag-wise Targeting (eTT), which are provided to all advertisers. Advertisers can select their preferred products freely. In the A/B test, experiment A is the online system with enhanced Tag-wise Targeting, while experiment B (ST+eTT) further provides Smart Targeting along with the online system. Such a setting is designed to meet product requirements. The online system with eTT should be consistently provided to meet targeting strategies from advertisers. Both A and B get 10% random traffic of users. The output number of ads from LSBM is set to 800. The time cost in the targeting stage is less than 100 milliseconds.

The A/B test lasts for 10 days. And the final result is reported in Table 1. We can find that Smart Targeting further brings 4.62% improvement to CTR, 8.77% improvement to eCPM than the online

[1] www.jd.com

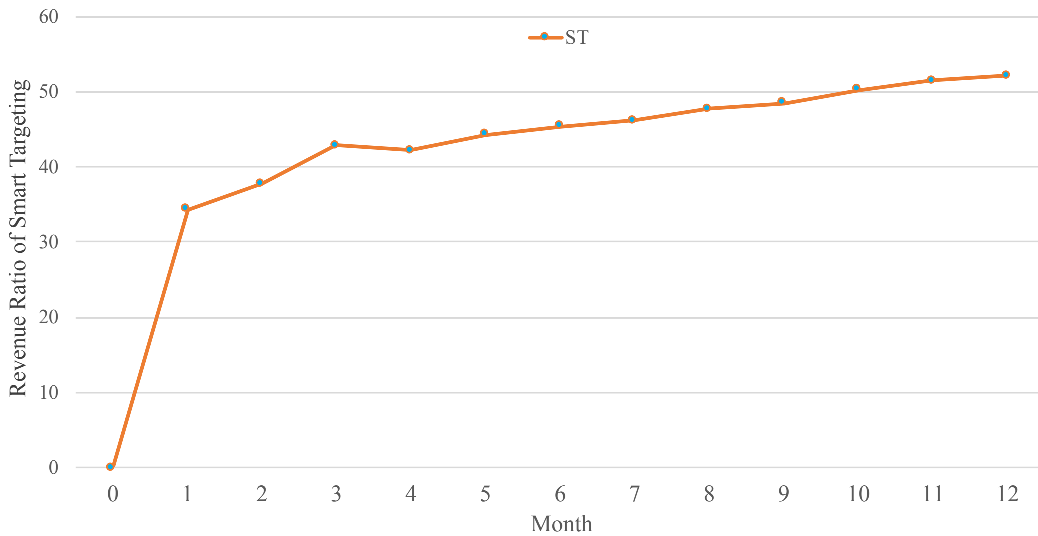

Figure 4: Revenue ratio of Smart Targeting (ST) along with months after production.

Table 2: Comparison between ST and eTT

	CTR Gain
ST vs eTT	+15.26%

system. It is a significant improvement to the online advertising platform. Higher CTR means that more user-relevant items are recalled, and higher CPC means that advertisers tend to set higher biding prices through smart targeting.

A comparison between different targeting methods is provided in Table 2 to show the advantage of Smart Targeting. The CTR of Smart Targeting is 15.26% higher than enhanced Tag-wise Targeting. It is mainly due to the reason that smart targeting explicitly models the relationship between ads and users through LSBM, which helps to recall more user-relevant ads. While in enhanced Tag-wise Targeting, advertisers select targeted users through tags by experience, and the relevance between the targeted users and ads of advertisers cannot be guaranteed.

Feedback from the Advertisers. Smart Targeting is widely used by advertisers and archives a 34.4% in the first month. Meanwhile, Fig. 4 describes the revenue ratio of smart targeting in the next 12 months after the release of the smart targeting product. We can see that the revenue ratio of Smart targeting keeps increasing and reaches 52.2% after the release of ST for one year. Recently, smart targeting becomes the most popular targeting product for advertisers.

To validate the assessment of targeting tools from advertisers, we conducted questionnaire surveys before and after the release of the smart targeting product. The advertisers were randomly selected and were required to give a rating score about the targeting tools. We finally received 350 and 464 feedbacks in the two surveys. We can find that the rating score is improved by 16.82%, which indicates that smart targeting can result in better advertiser experience.

Feedback from the users. On the APP of JD.com, the users are allowed to report their feedback when they receive unsatisfactory advertisements. To see the influence of smart targeting on user experience, we compare the number of the negative-feedback rate from users after/before the deployment of smart targeting. We can find that the negative-feedback rate decreases by 12.71%, and the most significant decrease occurs on the option of "`irrelevant recommendation`". Such a result again indicates that Smart Targeting can effectively improve user-item relevance and user experience.

5 CONCLUSION

Targeting plays an essential role in advertising system. Current Tag-wise Targeting only satisfies the intervention demands of advertisers but ignores modeling the relevance between ads and targeted users, which may result in bad advertising performance. Due to the weak interpretability of deep machine learning models, it is also hard to directly model the relevance in Tag-wise Targeting. In this paper, we propose a novel relevance-driven and configurable targeting framework called Smart Targeting to both model the relevance and satisfy advertisers' intervention demands. We develop a Long Short Behavior Model (LSBM) to learn the relevance and then allow advertisers to configure the targeted users based on the Delay Intervention Mechanism. We productize and evaluate Smart Targeting on the advertising platform of JD.com. The online experimental results show that Smart Targeting brings significant improvements to CTR and eCPM above Tag-wise Targeting. Meanwhile, it leads to positive feedback from advertisers and better user experience, making all-win results to users, advertisers, and the advertising platform.

REFERENCES

[1] Jianqing Chen and Jan Stallaert. 2014. An economic analysis of online advertising using behavioral targeting. *Mis Quarterly* 38, 2 (2014), 429–A7.
[2] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, et al. 2016. Wide & deep learning for recommender systems. In *Proceedings of the 1st workshop on deep learning for recommender systems*. 7–10.
[3] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems*. 191–198.

[4] G Thomas Des Jardins, David Markley, Vinayak Raghuvamshi, and Thomas W Macisaac. 2015. Audio/video advertising network. US Patent 9,104,669.

[5] Pavan Kumar Desikan, Lawrence Ip, Timothy James, Sanjeev Kulkarni, Prasenjit Phukan, Dmitriy Portnov, and Gokul Rajaram. 2010. Reviewing the suitability of Websites for participation in an advertising network. US Patent 7,788,132.

[6] Charles Eldering, John Schlack, Michael Plotnick, and Robert Deitrich. 2003. Targeted advertising in on demand programming. US Patent App. 10/068,771.

[7] Ayman Farahat and Michael C Bailey. 2012. How effective is targeted advertising?. In *Proceedings of the 21st international conference on World Wide Web*. 111–120.

[8] Scott Frank, Robert Starr, John Ruckart, Steven Tischer, and Samuel Zellner. 2013. Location-based targeting. US Patent 8,559,968.

[9] Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuqiang He. 2017. DeepFM: a factorization-machine based neural network for CTR prediction. *arXiv preprint arXiv:1703.04247* (2017).

[10] John S Hendricks and Alfred E Bonner. 2004. Method and apparatus for targeted advertising. US Patent 6,738,978.

[11] Efosa C Idemudia. 2010. The Extended Advertising Network Model.. In *AMCIS*. 185.

[12] Ganesh Iyer, David Soberman, and J Miguel Villas-Boas. 2005. The targeting of advertising. *Marketing Science* 24, 3 (2005), 461–476.

[13] Bernard J Jansen, Kathleen Moore, and Stephen Carman. 2013. Evaluating the performance of demographic targeting using gender in sponsored search. *Information Processing & Management* 49, 1 (2013), 286–302.

[14] Yuchin Juan, Yong Zhuang, Wei-Sheng Chin, and Chih-Jen Lin. 2016. Field-aware factorization machines for CTR prediction. In *Proceedings of the 10th ACM Conference on Recommender Systems*. 43–50.

[15] Ahmad A Kardan and Maryam Hooman. 2013. Targeted advertisement in social networks using recommender systems. In *7th International Conference on e-Commerce in Developing Countries: with focus on e-Security*. IEEE, 1–13.

[16] Kenneth Knowlson. 2003. Targeted advertising. US Patent App. 10/011,953.

[17] Tao Lei, Yu Zhang, Sida I Wang, Hui Dai, and Yoav Artzi. 2017. Simple recurrent units for highly parallelizable recurrence. *arXiv preprint arXiv:1709.02755* (2017).

[18] Anand Madhavan and Shyam Kapur. 2010. Temporal targeting of advertisements. US Patent 7,672,937.

[19] Steffen Rendle. 2010. Factorization machines. In *2010 IEEE International Conference on Data Mining*. IEEE, 995–1000.

[20] Imdad Ullah, Roksana Boreli, Mohamed Ali Kaafar, and Salil S Kanhere. 2014. Characterising user targeting for in-app mobile ads. In *2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 547–552.

[21] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*. 5998–6008.

[22] Guorui Zhou, Chengru Song, Xiaoqiang Zhu, Xiao Ma, Yanghui Yan, Xingya Dai, Han Zhu, Junqi Jin, Han Li, and Kun Gai. 2017. Deep Interest Network for Click-Through Rate Prediction. (2017).

[23] Han Zhu, Daqing Chang, Ziru Xu, Pengye Zhang, Xiang Li, Jie He, Han Li, Jian Xu, and Kun Gai. 2019. Joint optimization of tree-based index and deep model for recommender systems. In *Advances in Neural Information Processing Systems*. 3973–3982.

[24] Han Zhu, Xiang Li, Pengye Zhang, Guozheng Li, Jie He, Han Li, and Kun Gai. 2018. Learning tree-based deep model for recommender systems. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1079–1088.