

Who to show the ad to? Behavioral targeting in Internet advertising

Behavioral
targeting in
Internet
advertising

Wei Xiong

Iona College, New Rochelle, New York, USA

Ziyi Xiong

Stevens Institute of Technology, Hoboken, New Jersey, USA, and

Tina Tian

Manhattan College, Riverdale, New York, USA

15

Received 27 December 2021

Revised 27 January 2022

Accepted 27 January 2022

Abstract

Purpose – The performance of behavioral targeting (BT) mainly relies on the effectiveness of user classification since advertisers always want to target their advertisements to the most relevant users. In this paper, the authors frame the BT as a user classification problem and describe a machine learning-based approach for solving it.

Design/methodology/approach – To perform such a study, two major research questions are investigated: the first question is how to represent a user's online behavior. A good representation strategy should be able to effectively classify users based on their online activities. The second question is how different representation strategies affect the targeting performance. The authors propose three user behavior representation methods and compare them empirically using the area under the receiver operating characteristic curve (AUC) as a performance measure.

Findings – The experimental results indicate that ad campaign effectiveness can be significantly improved by combining user search queries, clicked URLs and clicked ads as a user profile. In addition, the authors also explore the temporal aspect of user behavior history by investigating the effect of history length on targeting performance. The authors note that an improvement of approximately 6.5% in AUC is achieved when user history is extended from 1 day to 14 days, which is substantial in targeting performance.

Originality/value – This paper confirms the effectiveness of BT on user classification and provides a validation of BT for Internet advertising.

Keywords Internet advertising, Machine learning, Behavioral targeting, User modeling

Paper type Research paper

1. Introduction

Targeted advertising has been growing its share of the ad market, as advertisers are gradually shifting their campaign budgets from other media to online advertising and hoping their advertisements to be viewed by the right customers at the right time. There have been targeting techniques that try to leverage the accumulated data of Internet users and show them ads accordingly. For instance, one of the basic techniques is to display ads based on user's geographic data, such as physical location (Bauer and Strauss, 2016). This method is especially useful for advertisers to target a particular location, such as a city or a country. Advertisers may adopt geographic targeting simply because their products or services are only available within a certain area. Similarly, ads can also be delivered to users based on their demographic information, such as marital status, education and gender (Jansen *et al.*, 2013). For instance,

© Wei Xiong, Ziyi Xiong and Tina Tian. Published in *Journal of Internet and Digital Economics*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>.



Journal of Internet and Digital
Economics
Vol. 2 No. 1, 2022
pp. 15-26
Emerald Publishing Limited
e-ISSN: 2752-6364
p-ISSN: 2752-6356
DOI 10.1108/JIDE-12-2021-0023

knowing that Xbox lovers are mostly young males, it would be wise to target them with Xbox ads. This approach allows advertisers to target a small number of people depending on demographics. However, demographic targeting could miss out on some potential customers that do not fall into a specific group. Apparently, a grandma could also be an Xbox buyer if she plans to give an Xbox to her grandchild as a present.

Keywords targeting, contextual targeting and retargeting are another three widely used targeting techniques. Keyword-targeted ads are typically shown on a search result page, targeting the terms users are searching for in search engines. For example, [Google Ads \(Google Ads–Get More Customers With Easy Online Advertising\)](#) is a popular form of keywords targeting, where search advertisements are displayed above or below search results. Unlike keyword-targeted ads, contextual targeting displays ads based on the content of webpages, such as [Google AdSense \(Google AdSense–Earn Money From Website Monetization\)](#) that runs an advertising program allowing publishers to monetize their content and traffic. As another common targeting technique, retargeting “remembers” your website visitors and shows ads to them several times after they leave your website without signing up or buying. Yahoo! Retargeting (*Retargeting ads creation and best practices*), for example, is a targeted advertising network that retargets visitors who have stayed at a publisher’s website before and shows them relevant advertisements when they browse other websites.

As an application of machine learning, behavioral targeting (BT) allows advertisers to display advertisements to users based on their previous web activities, such as browsing history and search queries. This type of advertising is on individual-level, which is more targeted and personalized compared to other targeting techniques mentioned earlier. In BT, when two users open the same webpage, even at the same location and the same time, they may see totally different ads. When a user opens a webpage with space for ads, advertisers compete for it via auctions to display their ads to the user. The bid decisions are made by advertisers based on a fit score between an advertisement and the user. This process finishes in only milliseconds as the communication between publishers and advertisers has to be completed before the webpage is loaded.

As shown in [Figure 1](#), there are two important datasets during this process: training data for learning a classification model and live data for an advertising bid request. The training dataset is accumulated data on users’ online activities, including users’ search queries, clicked URLs, clicked ads, timestamps and IP addresses. It consists of data on both “converters” and “non-converters”. As a pre-defined action, a conversion could be filling out an application (for a quote, for example) or an online purchase, depending on the type of an ad. These data are used to train a classification model using machine learning algorithms. The live data include new users’ online history and the bid request that allows advertisers to bid for an advertisement space. It contains the webpage information, user cookie id, timestamps and the advertisement space properties, including location, size and type. The advertisers then use the trained model to determine whether or not the user should be targeted, along with a bid price.

In general, BT focuses on leveraging users’ historical online activities to understand their interests and predict their future behavior for targeted ads delivery. The performance of BT mainly relies on the effectiveness of user classification since advertisers always want to target their advertisements to the most relevant users. Therefore, in our study, we consider BT as a user classification problem and investigate these two major research questions: how to represent users’ online behavior, and how different representation strategies affect the targeting performance. In addition, we also explore the temporal aspect of user behavior history by investigating the effect of history length on targeting performance.

2. Related work

In today’s digital world, marketers are aggressively collecting users’ Internet usage data and using the aggregated information for targeted advertisements delivery. Based on a report by

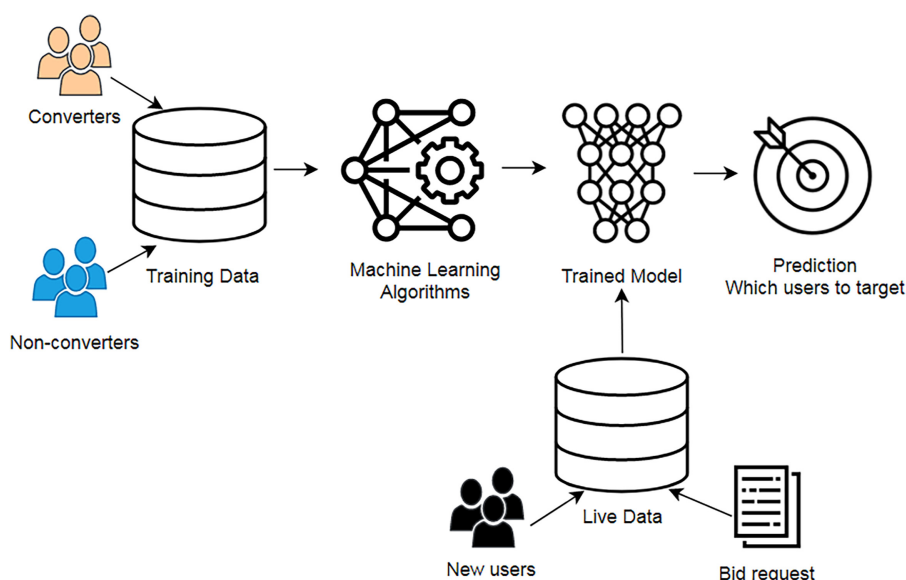


Figure 1.
Behavioral targeting
advertising problem
framework

Wall Street Journal, a single computer in the United States could have more than 3,000 tracking files installed by the 50 most popular webpages (*The New Gold Mine: Your Personal Information and Tracking Data Online—WSJ*). BT is believed to be the most powerful targeting technique for more personalized and individual-level advertising (Keller, 2016). Researchers argue that targeted advertising has become more and more common as the advertising market affirms its effectiveness (Aguirre *et al.*, 2015; Rust, 2016; Schultz, 2016; Ham, 2017). Hoban and Bucklin (2015) state that BT has a positive impact on visitation to a company’s webpage for consumers, according to their online randomized field experiment.

2.1 Targeting strategy

The research adopting a campaign-level perspective mostly focuses on exploring how to best deploy a targeting strategy for online advertising. For instance, Archak *et al.* (2010) model users’ historical online activities with graphic structures to represent local correlation among ad events. In their study, scoring procedures are introduced to measure global roles and the paths of ads in graphs, along with correlations between advertisement impressions and customer conversions. Nottorf (2014) demonstrates that ad click probabilities can be estimated by leveraging a binary logit model across multiple online advertising channels. Comparably, Kagan and Bekkerman (2018) predict users’ purchase behavior by building a classifier on clickstreams of panel members. Recently, by conducting a randomized field experiment, Lian *et al.* (2019) reveal in the restaurant industry that there is a positive correlation between the increase in advertisements click-through rate (CTR) and a variety of factors, including right timing and nearer distances. Their findings demonstrate how different targeting strategies affect online consumers’ response to ads of offline stores.

2.2 User segmentation

Effectively identifying user segments of interest can also help advertisers select the right audience in BT advertising. Daltayanni *et al.* (2018) propose a reputation system and perform

user segmentation using a blend of behavioral, demographic and reputation signal. Their experiment result indicates that the suggested user segments could help refine the advertiser's campaign and increase conversion rate. Kmeans (Kanungo *et al.*, 2002), a classic clustering method, is extensively employed to carry out users segmentation in prior literature because of good scalability, efficiency and high speed when performing large dataset analysis. Empirical research by Yan *et al.* (2009) shows online advertising in search engines could be significantly improved with the help of BT, where they apply Kmeans to user segmentation. Similarly, Zheng *et al.* (2012) use Kmeans to segment consumers by exploring their interest and the properties of Web services. They maintain that when consumers with similar interest are grouped together, it makes the service recommendation more effective.

2.3 Demand-driven taxonomy in behavioral targeting

Taxonomy of topics, which consists of BT categories for capturing a broad class of users' interests, is another popular method to represent user online activities. Budhiraja and Reddy (2015) use a three-level taxonomy to explore the semantic relationship among clustered search terms. To optimize the settings, hybrid taxonomy is proposed to address the occurrence of closely related sub-concepts in the query log. Public ontologies have also been employed to characterize user interests. For instance, Wang *et al.* (2009) construct a hierarchical topic space using Open Directory Project (ODP) ontology with the purpose of matching the tags of users' photos with relevant advertisements, which are projected in a topic space. Nevertheless, topics included in ontologies and taxonomies may require manual updates over time as they cannot be always comprehensive. In addition, some of the topics may be too coarse for representing user online behavior, resulting in information loss in the user data. A taxonomy that works in one advertising system might not work in another, which leads to limited use of demand-driven taxonomy in BT, especially across different domains.

3. User behavior representation

As shown in Figure 2, there are two types of user behavior that can be used for BT: *active* and *passive* activities. The former includes typing search terms and clicking ads where a user displays an interest by actively performing an action on a webpage. The latter includes visiting pages (clicked URLs), which may reveal the user's hidden intention. In this section, we discuss three different user behavior representation methods.

3.1 User search queries

With the fast expansion of digitization and Internet advancement, online search engines are becoming the "go-to" tools for information seeking. The search terms issued by users

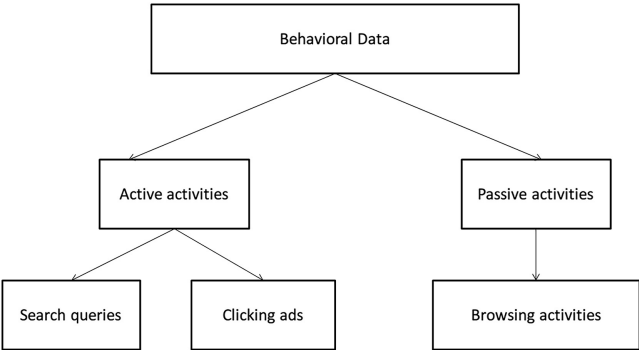


Figure 2.
Types of behavioral
activities

typically provide insights about their interests or information need, as users typically have to manually type in their query in the search box. For example, when a user issues search queries “flight”, “checked baggage allowances” and “visa application”, we know that the user is probably planning for a trip, although the user does not explicitly express his or her intention. Thus, we can characterize users’ online behavior by considering the occurrences of the terms (unigrams) that appear in their queries.

Inspired by the Bag of Words (BOW) model (Salton and Buckley, 1988), we represent each user as a “bag” of distinct terms that appear in the user’s search queries, without considering the order of terms. As such, a user that issued the query “North Carolina offline map” is represented as the same as the user that issued “offline map North Carolina”, as they both are represented as terms “North”, “Carolina”, “offline” and “map”. Hence, the feature space consists of all distinct terms from users’ queries.

We use Term Frequency Inverse Document Frequency (TF-IDF) (Jones, 1972) to assign a weight to each term for each user. It is defined as the term frequency $tf(t, d)$ multiplied by the inverse document frequency $idf(t, D)$, where in this case t is a term, d is a set of queries issued by a user. To be more specific, $tf(t, d)$ is the number of times t appears in d , and $idf(t, D)$ is defined as follows:

$$idf(t, D) = \log \frac{|D|}{1 + |\{d \in D : t \in d\}|}$$

where $|D|$ is the total number of users, and $|\{d \in D : t \in d\}|$ is the number of users whose queries contain term t . Hence, we can calculate the weight for each term using the following equation:

$$[tf * idf](t, d, D) = tf(t, d) \times idf(t, D)$$

In this way, each user’s online behavior is represented as a real-valued feature vector. Obviously, the weight of each term for each user not only depends on the frequency of the term in the user’s queries but also depends on how rare the term is in all users’ queries.

Although representing user behavior based on search query is straightforward, there are two potential disadvantages to queries. First, people tend to issue short and ambiguous queries when seeking information online. A typical query usually has a few words. For example, a user with the query “Java” could be interested in a type of programming language or a brand of coffee. This makes it challenging to distinguish users’ interests and information needs. Second, query feature space is sparse, as a large number of “tail” queries do not appear often. For instance, “laptop”, “Dell laptop” and “XPS 13” are all about laptops. However, “XPS 13” is more specific and it appears much less often. Without knowing “XPS 13” is a laptop model, it can be difficult to target the user with more focused ads.

3.2 User clicked URLs

Another key aspect of users’ online activity is that browsing activity may reveal user’s hidden intention as people are more likely to click URLs that are relevant to their interests. Users clicked URLs provide a comprehensive view of their browsing activity, and therefore, in this study clicked URLs are also used for user behavior representation.

Similar to search queries, clicked URLs also face the problem of high dimensionality and sparsity. In order to address this issue, we adopt Pinboard (*Welcome to Pinboard—Social bookmarking for introverts!*), a social bookmarking web service for returning a list of popular categories for each URL. And then we apply the same Bag of Words model and weighting method described in the previous subsection. Pinboard is one of the best-researched folksonomies, and each URL can be bookmarked and tagged by the entire community. By way of illustration, “https://singaporeair.com/en_UK/us/home” is tagged with categories like

“flight”, “travel”, “airfare”, “airline” and so on. By mapping the URLs into categories, the feature space can be reduced, resulting in a strong generalization of the trained models, although we may lose some granular data. For example, when a user clicks that URL, we know that the user might be purchasing a flight or planning a trip, but we do not know which airline the user is interested in.

3.3 User clicked ads

Another type of active activity that is used to represent user behavior is user clicked ads. When a user clicks an ad, he or she is probably interested in the product or service described in the ad. We assign each ad with a unique id and count the number of times it is clicked by each user. In this way, each user is modeled as a real-valued vector. Like clicked URLs, we also categorize ads based on an existing hierarchical ad categorizer to reduce the feature space.

4. Experimental evaluation

In the previous section, we presented three different user behavior representation methods: search queries, clicked URLs (categorized) and clicked ads (categorized). In this section, we empirically compare and combine them for BT. We also explore the effect of history length on prediction performance.

4.1 Dataset and baseline

In our experiment, a support vector machine (SVM) is used for training conversion models for each advertisement campaign. Each conversion model is trained on a set of positive examples (“converters”) and on a set of negative examples (“non-converters”). We use the data from an online advertising intermediary in the United States, which includes 2 weeks of data for 10 ad campaigns. About 75% of the dataset is used for training the models and the remaining 25% is used for testing. The training and testing split is performed randomly in order to avoid any data dependency between the training and testing datasets. To avoid noise, users who have more than 1,000 clicks within one day are filtered out (they are most likely robots). Furthermore, stemming is performed, and stop words are removed.

We compare the proposed user behavior representation methods using the area under the receiver operating characteristic (ROC) curve (Fawcett, 2006), denoted by AUC as a performance measure. User location information is used as the baseline while we investigate the relative difference as regards the baseline when adding different user behavior representations. We use L_2 normalized frequency for each user representation method. As for the history length, we consider the 1-day history length as a baseline while we examine the effect of history length on prediction performance with respect to this setting.

4.2 Behavior representation methods comparison

Table 1 compares the improvement in average AUC for the 10 campaigns when considering the proposed user behavior representation methods. User location information is treated as the baseline with a 0.00% improvement in AUC. As we can see, relative to user location, each presentation method performs really well. Of all the three presentation methods, search queries appear to be the most informative, despite of the challenges mentioned before. In addition, both search queries and clicked ads outperform clicked URLs. This indicates that browsing activity may provide some insights into the psyche that users would be unable to express, but it does not yield as much information about where they are in the purchase decision-making process as the active activities of search queries and clicked ads. We also see that combining user search queries, clicked URLs and clicked ads show the largest improvement in performance overall.

User presentations	Ad campaigns										Avg
	1	2	3	4	5	6	7	8	9	10	
User location (baseline)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Search queries	14.93%	15.31%	14.90%	14.65%	15.47%	15.20%	15.10%	15.23%	14.65%	15.33%	15.08%
Clicked URLs (categorized)	12.44%	12.51%	11.37%	12.05%	11.91%	12.18%	12.12%	13.11%	11.70%	12.60%	12.20%
Clicked ads (categorized)	14.16%	14.90%	13.70%	14.25%	14.63%	14.09%	14.30%	14.72%	13.37%	14.40%	14.25%
Combination of the above	16.05%	15.91%	16.22%	15.71%	15.28%	16.15%	15.31%	16.54%	15.92%	16.61%	15.97%

Table 1.
Targeting performance
improvement in AUC
compared to baseline

Despite being the most informative, search queries are non-stationary. A massive amount of new queries is generated by online users on a daily basis. Models constructed on user queries might be useful for days or weeks, or forever, or may just last for a few minutes. For instance, a user segmentation model could be trained for tablet PCs ads delivery using historical queries of users that have bought them. If the model was trained before iPad hit the market, it might not be very useful for recognizing users that are interested in iPads as potential tablet customers after iPad was introduced. The reason is that “iPad” cannot be one of the features in the model before it was introduced by Apple. But, customers who issue queries related to iPad would probably be interested in other tablet products as well. They would have responded to other tablet PC ads.

4.3 Effect of history length

Short recent history can help with short-term conversions, such as ordering a meal online, which do not take much thought, whereas long user history is essential when predicting other conversions, such as applying for a mortgage online, which typically take a longer time for users to make up their minds. We note that in Figure 3, an improvement of approximately 6.5% in AUC is achieved when we extend user history from 1 day to 14 days, which is substantial in targeting performance. This is not surprising given that longer history gives us more data about users, which allows a better understanding of user interests and hence better-targeting performance. However, in a practical setting, one might have to take into consideration many other factors, such as computational cost, scalability of the algorithm, and the cost of collecting and storing user data.

4.4 Ads click-through rate

The effectiveness of BT can also be measured using another popular metric, namely ads CTR. The CTR of ad a_i before user segmentation is denoted by $CTR(a_i)$. After users are grouped into segments, we define CTR of ad a_i of segment k as

$$CTR(a_i|k) = \frac{\text{number of users that clicked } a_i \text{ in } k}{\text{number of users in } k}$$

$$\text{while } CTR(a_i) = \frac{\text{number of users that clicked } a_i}{\text{number of users that are served with } a_i}$$

When determining $CTR(a_i|k)$, we choose the segment with maximum CTR because advertisers always tend to set their ad campaigns for targeting user segments with the highest ads click probability. In addition, the segment chosen must have a more than an average number of users in it, as it helps reduce some special situations. For instance, if the only user in a segment clicked the ad, that segment would have the highest CTR of 100%, but apparently it is not useful. Furthermore, we define the CTR improvement as follows:

$$\Delta CTR(a_i) = \frac{CTR(a_i|k) - CTR(a_i)}{CTR(a_i)}$$

In this experiment, we group similar users into segments based on their behavior: search queries and clicked URLs. Since clicked ads are used for evaluation purpose, we do not include them as one of the user representation methods in this section. We use CLUTO (Karypis, 2002), a classic clustering algorithm, to group users into 10, 20, 40 and 80 segments. As we see in Figure 4, advertisements CTR can be improved by as much as 144% via clustering users based on their online behavior. We see that using search queries can achieve better CTR improvement than using clicked URLs for user representation, which aligns with

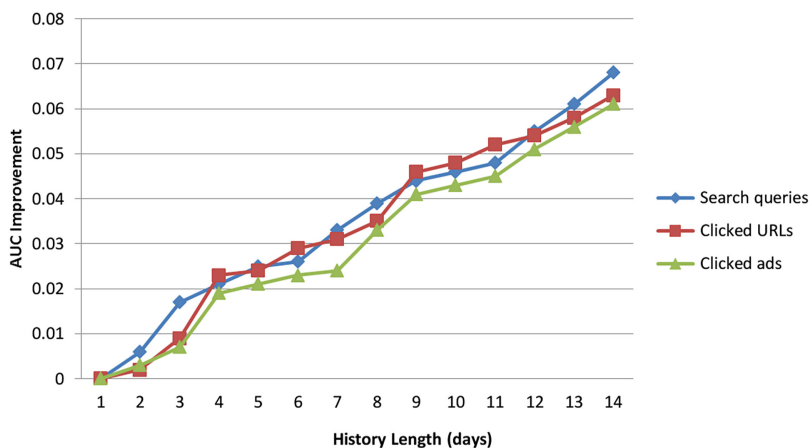


Figure 3.
Effect of user history
length

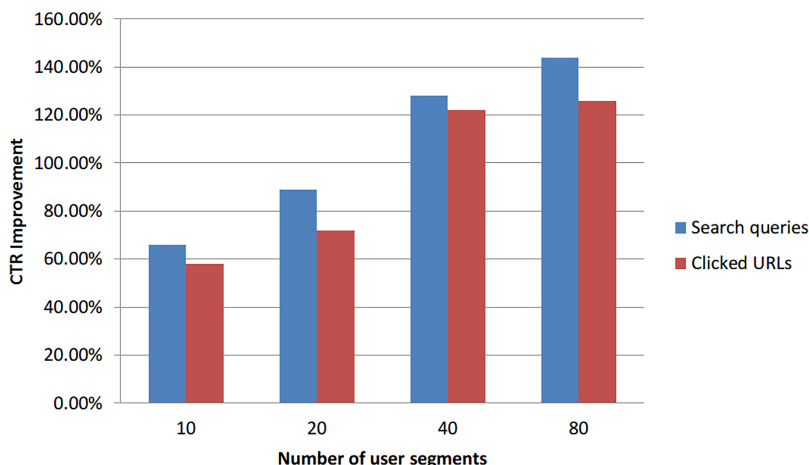


Figure 4.
CTR improvement by
user segmentation

our previous observation that search queries appear to be the most informative for predicting user conversions. In addition, we note that when we increase the number of segments, the CTR improvement increases as well. Yet, pushing a number of segments to the extreme would be unwise, as some segments might end up with only a few users.

5. Discussion

5.1 Research implications

This work has implications for the problem of “cold start” (Perlich *et al.*, 2014; Mo *et al.*, 2015; Zhao *et al.*, 2019), where the machine learning-based BT requires relatively large training data to achieve desirable performance. Before an advertising campaign is started, there are no data at all on whether or not users would convert after having been shown the advertisement. Labeled data (“converters” and “non-converters”) only start to accumulate after a campaign has started. Actually, the cold start is fairly common in the field of machine learning.

In our experiments, the positive examples are users who have actually converted in a targeted window. Conversions are so rare that trained classifiers based on them often do not perform well on the test sets due to highly imbalanced classes. To tackle this challenge, we slightly modify the SVM classifier by giving more weightage to the positive class in the cost function of the algorithm. In our study, a class weighting parameter of 20:1 for the positive and negative class provided optimal overall performance.

5.2 Practical implications

Traditional advertising, such as TV ads and outdoor ads, does not allow advertisers to monitor user interaction with ads, which makes it very hard to estimate the effectiveness of their ad campaigns and understand the characteristics of the customers who have responded to their ads. Internet ads, on the other hand, have begun to reduce those uncertainties. With the ever-increasing user behavior data in today's age, there is a dire need to model user profiles systematically and effectively in BT advertising. This research will be valuable to advertisers and marketers, in that it proposes and empirically compares three different user modeling methods. It helps maximize user conversions in BT, which directly translate to the advertiser's revenue.

6. Conclusions

This study contributes to the field of online advertising in the following aspects. First of all, we describe the BT problem in the context of machine learning. The experimental results in this study confirm the effectiveness of BT on user classification and provide a validation of BT for online advertising. Second, three different user behavior representation methods are proposed and compared empirically. We see that combining user search queries, clicked URLs and clicked ads show the largest improvement in performance when compared to the baseline of user location. Finally, we explore the temporal aspect of user behavior history by investigating the effect of history length on targeting performance.

References

- Aguirre, E., Mahr, D., Grewal, D., De Ruyter, K. and Wetzels, M. (2015), "Unraveling the personalization paradox: the effect of information collection and trust-building strategies on online advertisement effectiveness", *Journal of Retailing*, Vol. 91 No. 1, pp. 34-49.
- Archak, N., Mirrokni, V.S. and Muthukrishnan, S. (2010), "Mining advertiser-specific user behavior using adfactors", *Proceedings of the 19th International Conference on World Wide Web*, pp. 31-40.
- Bauer, C. and Strauss, C. (2016), "Location-based advertising on mobile devices", *Management Review Quarterly*, Vol. 66 No. 3, pp. 159-194, doi: [10.1007/s11301-015-0118-z](https://doi.org/10.1007/s11301-015-0118-z).
- Budhiraja, A. and Reddy, P.K. (2015), "An approach to cover more advertisers in adwords", *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, IEEE, pp. 1-10.
- Daltayanni, M., Dasdan, A. and de Alfaro, L. (2018), "Automated audience segmentation using reputation signals", *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Association for Computing Machinery (KDD '18), New York, NY, USA, pp. 186-195, doi: [10.1145/3219819.3219923](https://doi.org/10.1145/3219819.3219923).
- Fawcett, T. (2006), "An introduction to ROC analysis", *Pattern Recognition Letters*, Vol. 27 No. 8, pp. 861-874, doi: [10.1016/j.patrec.2005.10.010](https://doi.org/10.1016/j.patrec.2005.10.010).
- Google Ads - Get More Customers With Easy Online Advertising (2021), available at: <https://ads.google.com/home/> (accessed 22 April 2021).
- Google AdSense - Earn Money From Website Monetization (2021), "Google AdSense", available at: <https://www.google.com/adsense/start/> (accessed 18 April 2021).

-
- Ham, C.-D. (2017), "Exploring how consumers cope with online behavioral advertising", *International Journal of Advertising*, Vol. 36 No. 4, pp. 632-658.
- Hoban, P.R. and Bucklin, R.E. (2015), "Effects of internet display advertising in the purchase funnel: model-based insights from a randomized field experiment", *Journal of Marketing Research*, Vol. 52 No. 3, pp. 375-393.
- Jansen, B.J., Moore, K. and Carman, S. (2013), "Evaluating the performance of demographic targeting using gender in sponsored search", *Information Processing and Management*, Vol. 49 No. 1, pp. 286-302, doi: [10.1016/j.ipm.2012.06.001](https://doi.org/10.1016/j.ipm.2012.06.001).
- Jones, K.S. (1972), "A statistical interpretation of term specificity and its application in retrieval", *Journal of Documentation*, Vol. 28, pp. 11-21.
- Kagan, S. and Bekkerman, R. (2018), "Predicting purchase behavior of website audiences", *International Journal of Electronic Commerce*, Vol. 22 No. 4, pp. 510-539.
- Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R. and Wu, A.Y. (2002), "An efficient k-means clustering algorithm: analysis and implementation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24 No. 7, pp. 881-892.
- Karypis, G. (2002), *CLUTO - A Clustering Toolkit*, Minnesota Univ Minneapolis Dept of Computer Science, Minneapolis, available at: <https://apps.dtic.mil/sti/citations/ADA439508> (accessed 14 September 2021).
- Keller, K.L. (2016), "Unlocking the power of integrated marketing communications: how integrated is your IMC program?", *Journal of Advertising*, Vol. 45 No. 3, pp. 286-301, doi: [10.1080/00913367.2016.1204967](https://doi.org/10.1080/00913367.2016.1204967).
- Lian, S., Cha, T. and Xu, Y. (2019), "Enhancing geotargeting with temporal targeting, behavioral targeting and promotion for comprehensive contextual targeting", *Decision Support Systems*, Vol. 117, pp. 28-37.
- Mo, K., Liu, B., Xiao, L., Li, Y. and Jiang, J. (2015), "image feature learning for cold start problem in display advertising", *Twenty-Fourth International Joint Conference on Artificial Intelligence*, AAAI Press (IJCAI'15), Buenos Aires, Argentina, pp. 3728-3734.
- Nottorf, F. (2014), "Modeling the clickstream across multiple online advertising channels using a binary logit with Bayesian mixture of normals", *Electronic Commerce Research and Applications*, Vol. 13 No. 1, pp. 45-55.
- Perlich, C., Dalessandro, B., Raeder, T., Stitelman, O. and Provost, F. (2014), "Machine learning for targeted display advertising: transfer learning in action", *Machine Learning*, Vol. 95 No. 1, pp. 103-127, doi: [10.1007/s10994-013-5375-2](https://doi.org/10.1007/s10994-013-5375-2).
- Rust, R.T. (2016), "Comment: is advertising a Zombie?", *Journal of Advertising*, Vol. 45 No. 3, pp. 346-347.
- Salton, G. and Buckley, C. (1988), "Term-weighting approaches in automatic text retrieval", *Information Processing and Management*, Vol. 24 No. 5, pp. 513-523, doi: [10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0).
- Schultz, D. (2016), "The future of advertising or whatever we're going to call it", *Journal of Advertising*, Vol. 45 No. 3, pp. 276-285.
- The New Gold Mine: Your Personal Information & Tracking Data Online - WSJ (2021), available at: <https://www.wsj.com/articles/SB10001424052748703940904575395073512989404> (accessed 30 July 2021).
- Wang, X.J., Yu, M., Zhang, L., Cai, R. and Ma, W.Y. (2009), "Argo: intelligent advertising by mining a user's interest from his photo collections", *Proceedings of the Third International Workshop on Data Mining and Audience Intelligence for Advertising*, pp. 18-26, available at: <http://dl.acm.org/citation.cfm?id=1592752> (accessed 5 November 2012).
- Welcome to Pinboard—Social bookmarking for introverts! (2021), available at: <https://pinboard.in/> (accessed 7 July 2021).

- Yan, J., Liu, N., Wang, G., Zhang, W., Jiang, Y. and Chen, Z. (2009), "How much can behavioral targeting help online advertising?", *Proceedings of the 18th international conference on World Wide Web*, pp. 261-270.
- Zhao, Z., Li, L., Zhang, B., Wang, M., Jiang, Y., Xu, L., Wang, F. and Ma, W. (2019), "What you look matters? Offline evaluation of advertising creatives for cold-start problem", in *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM '19)*, Association for Computing Machinery, New York, NY, pp. 2605-2613, doi: [10.1145/3357384.3357813](https://doi.org/10.1145/3357384.3357813).
- Zheng, K., Xiong, H., Cui, Y., Chen, J. and Han, L. (2012), "User clustering-based web service discovery", in *2012 Sixth International Conference on Internet Computing for Science and Engineering*, IEEE, pp. 276-279.

Corresponding author

Wei Xiong can be contacted at: wei.xiong2008@gmail.com