Recommending Web Advertisements based on Long-Short Term User Interest

Panote Siriaraya
Kyoto Sangyo University
Kyoto, Japan
spanote@gmail.com

Yuriko Yamaguchi
Kyoto Sangyo University
Kyoto, Japan
i1658164@cc.kyoto-su.ac.jp

Mimpei Morishita
Kizasi Company, Inc.
Tokyo, Japan
mimpei@kizasi.jp

Yoichi Inagaki
Kizasi Company, Inc
Tokyo, Japan
inagaki@kizasi.jp

Reyn Nakamoto
Kizasi Company, Inc.
Tokyo, Japan
reyn@kizasi.jp

Jianwei Zhang
Iwate University
Iwate, Japan
zhang@iwate-u.ac.jp

Shinsuke Nakajima
Kyoto Sangyo University
Kyoto, Japan
nakajima@cc.kyoto-su.ac.jp

ABSTRACT

This paper reports the results of a study carried out to develop a system to recommend web advertisements to users based on their latent interests in an online real time bidding environment. As part of this work, we describe an approach which could be used to help predict the latent interest of users by analyzing their long and short term interests based on a large dataset of user web browsing histories. The proposed approach was tested in an experiment study with 32 different websites. Overall, this approach, which separated the user browsing history into sections representing their long and short term interests resulted in significantly higher predictive performance than when a singular section of user browsing history was used to represent the overall interests of users. In addition, we examined the effect of using different category levels as features to represent long and short term interest.

CCS CONCEPTS

• **Information systems** → **Content match advertising**; • **Computing methodologies** → *Supervised learning by classification.*

KEYWORDS

Real-time Web advertising; Real-time bidding; User browsing history analysis

IUI Workshops'19, March 20, 2019, Los Angeles, USA

ACM Reference Format:
Panote Siriaraya, Yuriko Yamaguchi, Mimpei Morishita, Yoichi Inagaki, Reyn Nakamoto, Jianwei Zhang, and Shinsuke Nakajima. 2019. Recommending Web Advertisements based on Long-Short Term User Interest. In *Joint Proceedings of the ACM IUI 2019 Workshops, Los Angeles, USA, March 20, 2019* , 5 pages.

1 INTRODUCTION

The introduction of Real-time bidding (RTB) to the online advertising environment has enabled advertisers to bid for advertisement space in real-time as soon as they become available. Instead of having to negotiate and pre-purchase ad display space from publishers in advance, advertisers are able to decide how much they are willing to pay to market their product to each audience based on real-time information about their characteristics and potential interests. Such a customized advertisement approach has allowed companies to market products in a more precise and cost effective manner [6]. As such, it is not surprising that RTB has shown considerable growth in recent years with a significant number of advertisers have already adopted this system [7].

In regards to the process used in the Real-time bidding, an ad request call is generally triggered when a user first visits a website with an ad content (the publisher site). This call is triggered to the Supply side platform (a platform used by ad publishers to manage their advertisement inventory) which then sends a bid request to different demand side platforms (platforms used by buyers of advertisements to manage and optimize their ad purchases) connected to the system. Information about the characteristics of users and the publisher site (such as their IP address, geographic location etc) is also generally provided to help the demand side platform decide on the potential value of advertisement space. Afterwards, each of the demand side platforms would submit a bid back to the supply side platform indicating how much they would be willing to pay for that particular advertisement space. The supply side platform would then determine the highest bidder and display the advertisement of the winner on the publisher site. Overall, the key challenge for a demand side platform system is therefore to determine which advertisement opportunity would be worth purchasing for each individual user based on their information which is provided by the supply side platform. This necessitates the development of automated systems able to predict user interest in different advertisement contents in near real time.

2 RELATED WORKS

To date, various research studies have been carried out to improve different aspects of the real-time bidding ecosystem (see [7]). For example, research has been carried out to enable advertisers to develop and deploy more effective bidding strategies in such an environment. One study for instance, has been carried out to develop an optimal bidding strategy based on available campaign resources and auction information [1] etc. Another study has looked to develop methods to effectively estimate the winning price of bid requests in a Real-time bidding process [8]. In addition, researchers have also explored topics such as detecting and preventing fraud in Real-time bidding systems[2] and estimating the cost the advertisement provider pays for each user based on the information they exposed [4].

Few studies however have focused on optimizing the process of selecting and presenting advertisements so that they would be appropriate for audiences based on their interests in a RTB environment. Conventional systems, particularly in industry, usually rely on techniques such as interest-match advertising (matching advertisement related keywords with user characteristics derived from analyzing user browsing behavior to determine how well an advertisement fits with a user) or re-targeting (target users who have previously visited the website of the advertisers) to determine if an advertisement is worth purchasing. A key disadvantage of such techniques is that they only emphasize on the current interests of users and thus, it is difficult for advertisers to market to users who might have latent (but not apparent) interest in their products.

Therefore, the overall aim of our research is to develop a system which could recommend advertisements to users based on their latent interests. In our initial study, we have discussed how a classification system could be used to predict latent user interest by analyzing the browsing history of the said user in comparison to the browsing history of users who have and have not previously visited a targeted website [5]. Furthermore, we have shown how the predictive performance of this system could be further improved when

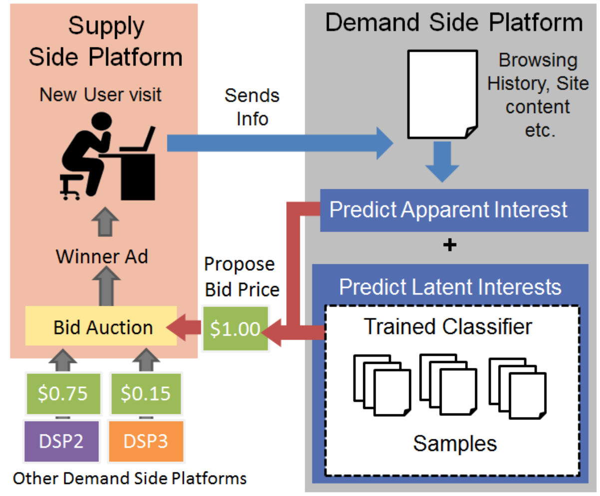

Figure 1: System overview

web categories instead of Fully Qualified Domain Names are used as features [3]. In this study, we predict latent user interests based on the notion of long-short term interest, by dividing the websites in the browsing history of users into separate features based on whether they represent the short or long term interests of the users. Two experiment studies were carried out showing how this approach improves predictive performance in comparison to the methods used in our previous studies [5] [3]. An overview of our proposed system is shown in Figure 1.

3 EXPERIMENT 1: EXAMINING THE EFFECTIVENESS OF LONG-SHORT TERM INTEREST IN PREDICTING THE LATENT INTERESTS OF USERS

In this paper, we describe an approach which aims to improve the predictive performance of latent user interest based on the concept of long and short term interest. The browsing history of users was used to represent their potential interest in different products and services which the advertisements offered. In our previous study, we found that using a longer browsing history acquisition period (7 days vs 1 day etc.) did not necessarily result in a better prediction of latent user interests [3]. This prompted us to hypothesize that perhaps user interests could not be adequately represented using a single section (one week or month etc.) of browsing history. In reality, the objectives, goals and intentions of users would fluctuate and change as time passes. As such, we assumed that the overall interests of the users might be better reflected by separating short term and long term interests in the user model and thus be able to more accurately predict latent user interests. For example, users who have browsed websites related to hiking in recent days, while having looked

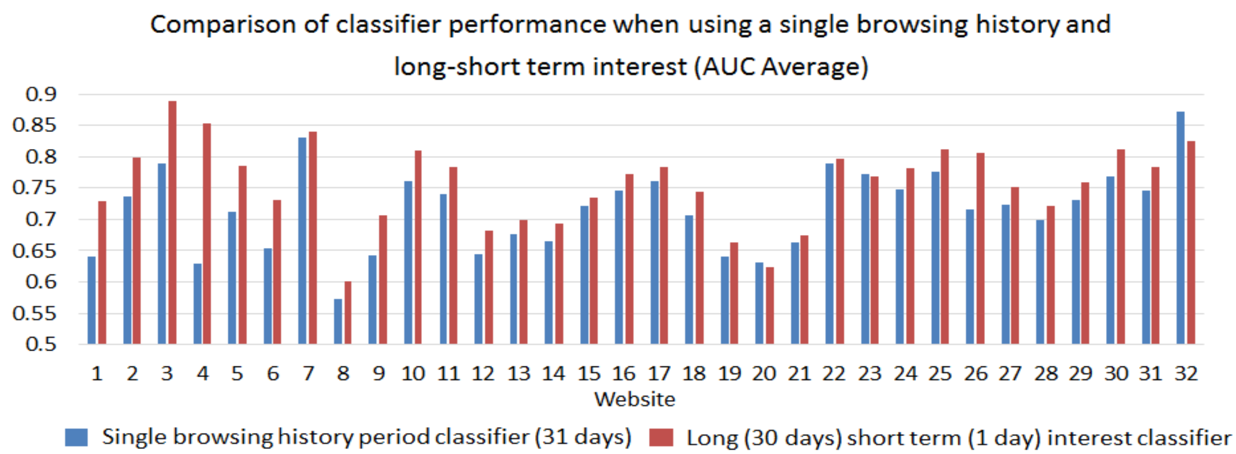

Figure 2: Comparison of classifier performance when using a single browsing history and long-short term interest (using averaged AUC value for 10 experiments)

at websites related to child rearing in the past month might indicate their interests in family oriented tour providers (as opposed to general tourism).

To examine the effectiveness of using the long and short term interests of users to predict their latent interests, an experiment study was carried out using a large dataset of user browsing history. Logistic regression was used to predict whether a user might be interested in a given website (which the user has never visited before) by analyzing their web browsing history in comparison to the web browsing history of users who have and have not previously visited the aforementioned site. Only the website views from the user browsing history was used in this study due to limitations in the information available in our data set. Overall, two different classifiers were constructed, one to represent the new approach (the combined long-short term interest classifier) and the other to represent our previous best approach (the single browsing history period classifier) [3]. The choice of Logistic regression in this study was mainly due to practicality. Other approaches which we had experimented with in our previous study (SVM etc.) had too large of a time cost while not differing considerably in performance [5].

In the single browsing history period classifier, we used the 31 day browsing history of users (Aug 1-31) who have previously accessed the targeted site as positive data and used the 31 day browsing history of users who had not previously visited the targeted site as negative data. For the combined long-short term interest classifier, the web browsing history of users who had accessed the targeted site in the short term (during the previous one day) was used as positive data, while the web browsing history of users who did not access the target site during the short term period was used as negative data (even if they had accessed the site during their long term period). In this classifier, user interests were separated into long term and short term interests. A 1 day segment of browsing history (Aug 31) was used to represent short term interest (as this was found to result in the best performance when representing short term interest in our previous study [5]) and a 30 day segment of browsing history (Aug 1-30) was used to represent long term interest (30 days before the start of the short term interest period). Overall, our data set consisted of user access records of approximately 6.3 million unique websites with an overall size of around 370 GBs.

The categories of the websites within the user browsing histories were used as predictors (i.e. features) in the regression model (as our prior experiments had shown that they resulted in superior performance than using individual domain names). These were determined by analyzing the word frequencies contained within the website (see([3])). Three different levels of categories were obtained for each site: large (such as "Fashion"), medium (such "Fashion accessories") and small (such as "Jewelry"). The total number of large, medium and small categories which were identified in our study were 23, 274 and 837 respectively. A combined category level was created by combining the large, medium and small categories together (resulting in a feature vector space of 1134, the combination of all three possible category types). For example, the feature vector for a web browsing history of a user who had visited a Jewelry store website twice and a Fast food store website four times would be 2 for the categories "Fashion" (large category), "Fashion accessories" (medium category) and "Jewelry" (small category), 4 for the categories "Groceries" (large category), "Restaurants" (medium category) and "Fast food" (small categories) and 0 for all the other categories. This combined category level was used as the input feature for both classifiers in this experiment (as this best represents the overall characteristics of the website, combining the features from all the categories). To differentiate the long term interest features from the short term interest features in the classification process, a suffix was added to the categories. For example, if a fashion website was present in the long term browsing history, then the "fashion_LT" category is used. If the website was present in the short term browsing history, then the "fashion_ST" category is adopted instead.

In the experiment, both classifiers were tested against 32 different target websites based in Japan. These sites include:

- An information site of a professional baseball team
- A social gaming website
- A news website
- A point redemption website
- A website offering tourism and living advice for a city in Japan
- A beauty and fashion information website

The browsing history data from users who had visited the site and those who had not were used as positive and negative training samples. The classifier was trained at a 1:1 sample ratio (The number depended on the number of samples available in our dataset for each website, but generally consisted of around 100 or 200 users). When training the classifier for each target site, browsing history data from 100 users who had previously visited the aforementioned website and 100 users who had not visited the website were used as positive and negative samples. The number of samples used in the test data were also selected to best represent the real world conditions of the data (in most cases, there are generally few positive samples when compared to negative samples). Therefore, the test data consisted of approximately 50 to 100 positive samples not in the training data (the exact number depending on the user data available for each website) and 10,000 negative user samples (browsing history of users who had not visited the website). The experiment on each target site was repeated 10 times and the average Area under the curve (AUC) value was used to measure the performance of the classifier on each target site.

Overall, the results showed that the combined long-short term interest classifier outperformed the single browsing history period classifier (our previous best approach [3]). Figure 2 shows a comparison of the results. The long-short term interest classifier outperformed the single browsing history period classifier on 29 out of the 32 sites. Paired T-tests which were carried out confirmed that average AUC value for the 32 target sites of the combined long-short term interest classifier (Mean=0.759, Min=0.602, Max= 0.889 SD=0.004) was significantly higher than the single browsing history period classifier (Mean=0.711, Min=0.572, Max=0.872 SD=0.004)

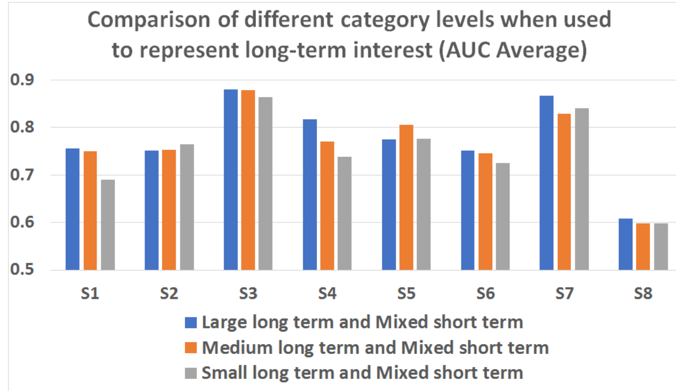

Figure 3: Comparison of classifier performance when using different category levels to represent long term interest (using averaged AUC value for 10 experiments)

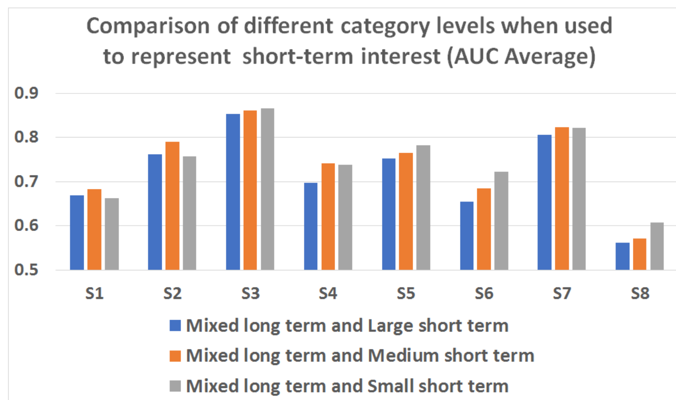

Figure 4: Comparison of classifier performance when using different category levels to represent short term interest (using averaged AUC value for 10 experiments)

($p<0.001$). Websites which showed considerable improvements in performance when using long-short term interest include a car information (+0.10) and a travel website (+0.22). Websites which showed worst performance include a point redemption site (-0.04) and a game news site (-0.003).

4 EXPERIMENT 2: EXAMINING THE EFFECT OF CATEGORY LEVELS IN LATENT INTEREST PREDICTION WHEN USING LONG-SHORT TERM INTEREST

In the previous experiment, we had used the combined category level as the features in our dataset. However, our prior studies also show that different category levels could influence classifier performance in predicting latent user interest. As the combined category level contains a combination of large, medium and small categories, it was not clear how each category levels contributed to the predictive ability of our long-short term interest classifier.

Therefore, we carried out a further experiment study to examine the predictive performance of using either large, medium and short categories to represent long and short term interest. The method used in this experiment was similar to the previous study, in which 32 different target websites were evaluated and a 1:1 sampling ratio was used to train the classifier (the testing sample consisted of approximately 50 to 100 positive samples and 10,000 negative ones). A 30 day browsing history period was used to represent long term interest and a 1 day browsing history period was used to represent short term interest. In addition, similar to the previous experiment, each target site was evaluated 10 times and the average AUC value was used as a measure for classifier performance. This performance measure was used (as opposed to other measures such as the F-score) due to the low positive to negative sampling ratio of users who visited each website.

Figure 3 shows a comparison of the predictive results when different category levels were used to represent long-term interests for the first 8 websites. A Repeated Measures Anova test showed that there was a significant difference in performance between the category levels used to represent long term interest ($F(1.67, 51.96) = 24.50$, $p < 0.001$). Representing long-term interest using the large category level resulted in significantly better AUC score (Mean=0.758, SD=0.07) than using medium categories (Mean=0.743, SD=0.07) ($p<0.01$) and small categories (Mean=0.732, SD=0.06) ($p<0.01$) for 27 out of the 32 websites. Thus, for long-term interest, it seems that large categories provide a better representation of long-term interest than medium and small categories. For short term interests, a Repeated Measures Anova also showed that there was a significant difference in performance when the different category levels were used ($F(1.70, 52.75) = 13.78$, $p < 0.001$). On average, using small categories to represent short term interest resulted in the best performance (Mean=0.725, SD=0.06) when compared to using medium (Mean=0.722, SD=0.06) ($p<0.01$) and large categories (Mean=0.707, SD=0.07) ($p<0.01$) (see Figure 4).

5 CONCLUSION

In this paper, we discuss our work to develop a system which could recommend advertisements to users based on their latent interests in a RTB environment. In particular, we describe an approach which utilizes a user's long and short term interest to predict their latent interest in a target website. We outline the results of an experiment study, showing how this approach results in better predictive performance than when a single browsing history period is used to represent user interests. In addition, we carried out further experimental studies showing that using large category levels to represent long term interest and small category levels to represent short term interest resulted in the best performance. In our future work, we would expand our method and experiment with different classification algorithms (Random Forest, xgboost) and different time-series approaches (RNN, LSTM) to investigate their performance in predicting latent user interest. In addition, we would further examine how additional real-data usage data (ad click through rates etc.) and website features could be used as features to improve advertisement recommendation. Finally, we would experiment with different time periods for short term (2-3 days) and long term interest (45, 60 days etc.).

ACKNOWLEDGMENT

The work carried out in this paper is supported by the JSPS KAKENHI research grant (Grants number 17H01822).

REFERENCES

[1] Han Cai, Kan Ren, Weinan Zhang, Kleanthis Malialis, Jun Wang, Yong Yu, and Defeng Guo. 2017. Real-Time Bidding by Reinforcement Learning in Display Advertising. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. ACM, 661–670.

[2] Brendan Kitts, Jing Ying Zhang, Gang Wu, Wesley Brandi, Julien Beasley, Kieran Morrill, John Ettedgui, Sid Siddhartha, Hong Yuan, Feng Gao, et al. 2015. Click fraud detection: adversarial pattern recognition over 5 years at Microsoft. In *Real World Data Mining Applications*. Springer, 181–201.

[3] Siriaraya Panote, Yuriko Yamaguchi, Mimpei Morishita, Youichi Inagaki, Reyn Nakamoto, Jianwei Zhang, Junichi Aoi, and Shinsuke Nakajima. 2017. Using categorized web browsing history to estimate the user's latent interests for web advertisement recommendation. In *Workshop on Human-Machine Collaboration in BigData, 2017 IEEE International Conference on Big Data*. IEEE.

[4] Panagiotis Papadopoulos, Pablo Rodriguez Rodriguez, Nicolas Kourtellis, and Nikolaos Laoutaris. 2017. If you are not paying for it, you are the product: How much do advertisers pay to reach you?. In *Proceedings of the 2017 Internet Measurement Conference*. ACM, 142–156.

[5] Yuriko Yamaguchi, Mimpei Morishita, Youichi Inagaki, Reyn Nakamoto, Jianwei Zhang, Junichi Aoi, and Shinsuke Nakajima. 2016. Web Advertising Recommender System Based on Estimating Users' Latent Interests. In *Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services (iiWAS '16)*. ACM, New York, NY, USA, 42–49. https://doi.org/10.1145/3011141.3011180

[6] Shuai Yuan, Jun Wang, and Xiaoxue Zhao. 2013. Real-time bidding for online advertising: measurement and analysis. In *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*. ACM, 3.

[7] Yong Yuan, Feiyue Wang, Juanjuan Li, and Rui Qin. 2014. A survey on real time bidding advertising. In *Service Operations and Logistics, and Informatics (SOLI), 2014 IEEE International Conference on*. IEEE, 418–423.

[8] W. Zhu, W. Shih, Y. Lee, W. Peng, and J. Huang. 2017. A gamma-based regression for winning price estimation in real-time bidding advertising. In *2017 IEEE International Conference on Big Data (Big Data)*. 1610–1619. https://doi.org/10.1109/BigData.2017.8258095