An Intelligent System for Effective Mobile Application Advertising

[1]Gene P. K. Wu, [2]Yi-Cheng Chen, [3]Wen-Yuan Zhu and [4]Keith C. C. Chan
[1,4]Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong
[2]Department of Computer Science and Information Engineering, Tamkang University, New Taipei City, Taiwan
[3]Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan
E-mail: cspkwu@comp.polyu.edu.hk ycchen@mail.tku.edu.tw wyzhu@cs.nctu.edu.tw cskcchan@comp.polyu.edu.hk

***Abstract*—Due to the advance of big data and increasing availability of smartphones equipped with various sensors and networking capability, it provides new opportunities for innovating mobile application advertising services. Unlike Web, where cookies for identifying users store in web browsers, there is a challenge for mobile apps to track users so acquiring sufficient data for training from ideal sampling distribution is computationally and economically expensive. Thus, a massive-scale intelligent system for targeted mobile app display advertising is developed to capture data for the learning task and transfer extracted knowledge back to the target task. The proposed system is evaluated by real world data and deployed to an advertising agency, illustrating the practicability and applicability.**

Keywords-mobile app advertising; inductive learning; intelligent system; big data analytics

I. INTRODUCTION

In Internet economy, mobile advertising is a fast growing industry. This creates hype for application developers to deliver mobile services and applications (apps). Mobile app advertising is a key subfield of the industry where advertisers are keen on well-targeted ads. It is promising as recent work has shown that data collected from mobile sensors from handheld devices such as geo-locations, speed of moving can be used to predict the current context of the people [1]. For instance, an online advertiser can show ads to the users who are more likely to be influenced by the ads based on the user context revealed by statistics collected from mobile sensor data. It is however also challenging as the mobile advertising ecosystem is complex where obtaining mobile data and showing the ads to the target audiences involve a large number of players and transactions as comparing to those of more traditional search engine advertising.

This paper proposes a novel system to tackle this targeted mobile app display-advertising problem. In particular, it is to deliver advertisements to consumers who have no known observed interaction with the advertisers but may be interested to the brand that are said to be likely to become customers after they read the advertisement.

II. MOBILE APP ADVERTISING SYSTEM

The proposed system consists of 2 phases on inductive learning that is 1) the learning is performed for a task (we call it source task) that is different from the target task in terms of the data samples, the attributes, the class label and the functional dependence between the attributes and the class label, and 2) the knowledge extracted from this learning is transferred back to the target task in order to improve the performance of the target task. In this problem, we have a number of data sets $D_1, \ldots, D_z$. Each data set refers to an advertising campaign. For each advertising campaign after it launches for a while, it accumulates a large number of data samples. Each data sample as denoted by x consists of N discrete valued attributes and a class label c. The set of attributes $A = \{A_1, \ldots, A_N\}$ is data collected from mobile sensors such as geo-location, telecommunication operator, device models, mobile app category, mobile app and data enriched by data analytics partners such as app language, app popularity, app developer and etc. The class label c is a binary variable: 0 for failure conversion or 1 for successful conversion. The number of data samples, M, of this campaign is the impression. The conversion rate is defined as

$$Conversion = \frac{the\ number\ of\ data\ samples\ where\ c=1}{M}.$$

The problem is to maximize $Conversion$ subject to a constant M. Now given a number of advertising campaigns $D_1, \ldots, D_z$, any one of them, whether used for training or testing, represents a learning task with data samples, attributes, class label and model f. A source task includes a source dataset $D_{i,\ i \in S}$ and a source model f_S. A target task includes a target dataset D_T and a target model f_T. The goal is to build models from different source tasks D_S, which predict well for the target task that is to estimate f_T as good as possible. Inductive learning aims to improve the learning of f_T by transferring knowledge of D_S and f_S into the estimation of f_T. In inductive learning principle, f_S is not observed but is learned approximately from D_S.

A. Dimensionality Reduction and Model Learning

This phase in the inductive process is to reduce the feature space in order to make the learning possible. The design of the system adopts the Mixed-mode Attribute Clustering Algorithm (MACA) by [2, 3] to cluster the attribute set $A = \{A_1, \ldots, A_N\}$ of a source dataset $D_{i,\ i \in S}$ into k attribute clusters such that attributes within a cluster should have high interdependence whereas attributes in different clusters are less correlated. Using this algorithm, the most representative attribute A^c from each attribute cluster can be found. The system takes the set of the most representative attributes from different attribute clusters, $\{A^{c_1}, \ldots, A^{c_k}\}$, as the reduced set of attributes. This procedure repeats until every source data set $D_{i,\ i \in S}$ is pre-processed. We then obtain a set of dimension reduced data set $D_{\bar{S}}$. Given the rare positive sample, p.s. class label is 1, the sampling strategy is to pick 1 positive sample and then to pick 1 negative sample randomly drawn from the population. This can eliminate > 99.95% negative samples since the conversion rate for an impression is typically < 0.05%. This procedure repeats until all dimension reduced data set $D_{\overline{S}}$ are sampled. We then obtain a set of sampled dimension reduced data set $D_{\overline{S}}'$. The operations can be executed in parallel so that different source tasks can start working simultaneously. Once the $D_{\overline{S}}'$ is prepared, the learning process can begin. The system adopts pattern discovery [4] approach in supervised learning to find rules for app users to convert an impression to a click. After all the rules are discovered, the source model f_S is ready to use.

B. *Model Evolution*

Suppose that a set of rules that come up with the source model f_S is produced by pattern discovery in phase A. For launching a new ad campaign (target task), the system applies the source model f_S to predict the chance whether or not this app user will convert this ad impression to a click. If the predicted outcome is higher than the pre-defined threshold, the system will place a bid. Assuming that the system decides to place a bid and the bid wins at an appropriate price, the actual outcome of whether or not the app user will click on the display ad banner can be measured and the target dataset D_T will increment a new data sample. Ultimately, the target data set D_T can again be an instance of the inductive learning by adaptively treating some of its positive samples as the source data set D_S and re-running the phase A. From time to time, the source model f_S will keep evolving to approximately estimate the target model f_T by improving the lift of the model trained on phase 1.

III. Experimental Results

To the best of our knowledge, the proposed system is the first discussing the attribute clustering among multiple advertising campaigns with inductive learning concept. The UI of the proposed system is shown in Fig. 1. Users can input requirements for ad campaigns. The system will optimize the goal and report the results automatically.

The first experiment applied on a real advertising campaign aims at separating the source task model and the target task model to test whether or not the source task model is able to effectively predict the outcome of the target task. The metric used is lift using the top 5% of the population. The lift is calculated by dividing the number of positive samples in the top 5% of the trained model by the number of positive samples in the top 5% of a random model. The higher the lift, the better the performance of the model. The data plotted in the chart is the average of 1,000 bootstrap estimates for the sake of minimizing the variance. Fig. 2 compares model performance of 10 campaigns between inductive learning on source and target task. The y-axis is for the lift of source task while the x-axis is for that of target task. Data points above the regression line mean the lift of source task is better. 7 out of 10 for the model trained in source task has better performance than those trained in target task. It is attributed to the fact that more information is stored in source task and knowledge extracted by the model is applicable to predict outcome of target task.

The second experiment analyzes performance by conversion rate that is the number of clicks divided by the number of impressions. The advertising agency is looking for potential customers to drive sales and online membership registration with limited budget. The goal is to maximize the conversation rate. In week 1, the agency wanted to test the market by random audiences. As revealed from Fig. 3, the conversion rate is near 0%. Beginning from week 2, the agency launched the campaign at scale without wasting impression. The source task model immediately boosted the conversation rate up to around 0.04%. After the target model accumulated sufficient positive samples and the source model re-trained and evolved, in week 7 the conversion rate began to improve and reached 0.055% steadily in week 8.

IV. Conclusion

The proposed system solves the mobile app target advertising problem by an inductive learning approach by employing attribute clustering algorithm in dimensionality reduction and pattern discovery in model learning, producing promising results. The model evolution phase demonstrates that a good modeling scheme directly influences the performance of a learning task. All in all, this framework provides a means to its users for effective mobile app target advertising with empirical supports that help the growth and reliable use of big data knowledge.

References

[1] E. Miluzzo, C. T. Cornelius, A. Ramaswamy, T. Choudhury, Z. Liu, and A. T. Campbell, "Darwin phones: the evolution of sensing and inference on mobile phones," Proc. ACMMobiSys, 2010.

[2] G. P. K. Wu, K. C. C. Chan, and A. K. C. Wong, "Unsupervised fuzzy pattern discovery in gene expression data," BMC Bioinformatics, vol. 12 (Suppl 5), 2011.

[3] A. K. C. Wong, B. Wu, G. P. K. Wu, and K. C. C. Chan, "Pattern discovery for large mixed-mode database," Proc CIKM, 2010.

[4] Y. Wang, and A. K. C. Wong, "From association to classification: inference using weight of evidence," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 3, pp. 914-925, 2003.

Figure 1. UI of the mobile app advertising system.

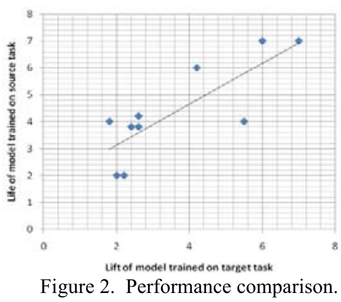

Figure 2. Performance comparison.

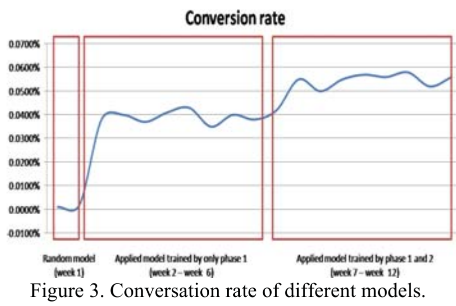

Figure 3. Conversation rate of different models.