A Novel Business Model Based on Real Time Bidding and Online Video Interaction Technology

Jiesheng Zhang, Shanghai Jiao Tong University, China
Chengyan Feng, Shanghai Jiao Tong University, China
Vincent Chang, Shanghai Jiao Tong University, China
Jia Tan, Shanghai Jiao Tong University, China
Wayne Wu, AVD Digital Media, China

The European Business and Management Conference 2015
Official Conference Proceedings

Abstract
The aim of this paper is to introduce a novel business model based on real-time bidding (RTB) and an online video interaction technology, and also to demonstrate how the operation of this model could possibly boost the power of advertising. In this context, power is defined to be advertisers' ability to precisely reach their target customers with lowest cost. This business model makes the scenario become possible: when watching a video, in whatever object a viewer feels interested, he or she can simply click the object, and advertisements that tightly relevant to this object will be instantly displayed. On the other hand, if the viewer doesn't click any object, there will be absolutely no advertisements during the entire watching process. This model gets rid of annoying pre-rolls, post-rolls and in-video ads, and brings the most frictionless watching experience to viewers. At the same time, being targeting and cost efficiency, RTB saves advertisers time and money to reach target customers. To prove the technical feasibility, a novel video interaction technology was built around a novel object recognition/tracking technique. Innovative algorithm of scene-splitting was implemented to ensure recognition and tracking accuracy. Through the video-interaction, the Click-Based Viewer Preference Database is established and served as an indispensable price reference during RTB. As a result, advertising becomes more powerful in that advertisers are capable of reaching their target customers more precisely with less time and money spent, and viewers can obtain more information from what they see.

Keywords: Online Video Advertising, Real-Time Bidding (RTB), Video Object Recognition and Tracking, Customer Behavioral Database, Interactive Advertising

Introduction

With the booming use of the internet, people are getting used to doing everything online. Watching videos is undoubtedly one of them. Where there are huge amount of people, there are advertisements. Advertisement space has been the most monetizable assets in the industry of online video, whose revenues are larger than subscription or transactions. According to Russo, IHS Technology (2015), online video generated a global advertising revenue of $11.2 billion in 2014, which has been doubled since 2011. It is forecasted that by 2017, online video advertising revenues will reach $19 billion, which surpasses TV advertising in some markets (Russo, 2015). Take the market in US and China as examples. eMarketer forecasted that the digital advertising spend in US will increase to $5.9 billion in 2014, which achieved a 56% yearly growth (as cited in Blattberg, 2014). Also, iResearch (2014) pointed out that Chinese online video market generated a total revenue of 12.8 billion Yuan in 2013, a growth of 41.9% from 2012.

Currently three formats of online video ads are prevalent in the market, which are pre-roll ads, post-roll ads and in-video ads (VanBoskirk, Li, Katz & Lee, 2007). They are classified by the time they are displayed to audience (before video, after video or in video). Though huge market size, we'd better carefully consider the disturbance that online video ads bring to viewers' watching experience. VanBoskirk et al. (2007) pointed out that 82% audience found the formats of online video advertising annoying. The underlying reason for audience's negative feedback is that ads displayed to audience usually have nothing to do with the video content. In that case, ads become nothing but a kind of interference. Being aware of this embarrassing situation, YouTube, who was bought by Google in 2006, is considering to offer an ad-free, subscription version so that audience can have choice to avoid watching ads (Hopkins, 2006; Etherington, 2014).

Research on Object Tracking

In order to provide ads that are tightly associated with the video content, researchers came up a solution called object tracking. Hare, Saffari and Torr (2011) put forward an adaptive object tracking algorithm whose name is STRUCK. This algorithm relies on a structured output support vector machine with kernels (Hare et al., 2011). Kalal, Mikolajczyk and Matas (2012) proposed Tracking-Learning-Detection (TLD) as a novel tracking algorithm. TLD decomposes the tracking process into tracking, learning, and detection, which enables tracking to be adaptive based on previous frames (Kalal et al., 2012). Henriques, Caseiro, Martins and Batista (2015) proposed a high-speed tracking algorithm, which exploited the properties of circulant matrices and increase the computation efficiency by leveraging Discrete Fourier Transform. In addition, Zhong, Lu and Yang (2012), Kwon and Lee (2011), and Dinh and Medioni (2011) also came up with innovative algorithms for object tracking in videos.

The reason why object tracking is viewed as a promising solution to irrelevant video ads and poor watching experience is that by knowing the timestamp and position information of any object within a video, once a viewer clicks this object, video content providers immediately know the exact object the viewer is clicking through matching with pre-stored tracking results. Therefore, advertisers know completely what the viewer likes, and ads delivered to each viewer could be more personalized, targeting and customer-oriented.

Research has also been conducted on the evaluation of object tracking algorithms. Wu, Lim and Yang (2013) suggested some benchmarks to gauge the performance of different object tracking algorithms in terms of their accuracy and precision. With the help of those benchmarks, the Computer Vision Lab at Hanyang University ("Visual Tracking Benchmark", n.d.) published the evaluation results on 29 tracking algorithms. The results indicated that the average overlap between the tracking result and the ground-truth is around 50% to 60%, and that the average failure rate is approximately 3 to 8 per 100 frames (Hu, Wang, Zeng, Lai & Wang, 2015).

Real-Time Bidding (RTB)

Besides the irrelevant and boring online video ads, the time and money efficiency of advertising is always a major concern for advertisers. Real-Time Bidding (RTB) is brought to the industry promptly. In the online video advertising industry, RTB becomes increasingly essential since its debut in 2009 ("The Arrival of Real-Time Bidding and What it Means for Media Buyers", 2011). RTB is a technology that utilizes computer algorithms to intelligently bid for certain advertisement space based on audience profile (Yuan, Wang & Zhao, 2013).

Figure 1 shows the existing working mechanism of RTB in the industry of online video advertising. Firstly, video producers, either video production companies or individuals, produce videos and upload to various video platforms, such as YouTube from US, Youku-Tudou from China, niconico from Japan, etc. Then, when viewers watch videos on those platforms, RTB is triggered. Video platforms basically tell the Supply-Side Platforms (SSP) the need for advertisements. Next, Ad Exchange, such as Doubleclick from Google, Tanx from Alibaba and Facebook Exchange from Facebook, announces detailed bidding information to the Demand-Side Platforms (DSP). With the help of DSP, advertisers will decide the bidding price for one advertisement to be delivered to one particular viewer (Zhang, Yuan & Wang, 2014). Making this tough decision involves requesting data from some third-party Data-Management Platforms (DMP). DMP provides such information like the geographical and behavioral analysis of a particular viewer to DSP so that DSP could submit a reasonable bidding price. Finally, the advertiser who offers the highest bidding price wins the chance to deliver ads to the particular viewer (Zhang, Feng, Chang, Tan & Qin, 2015).

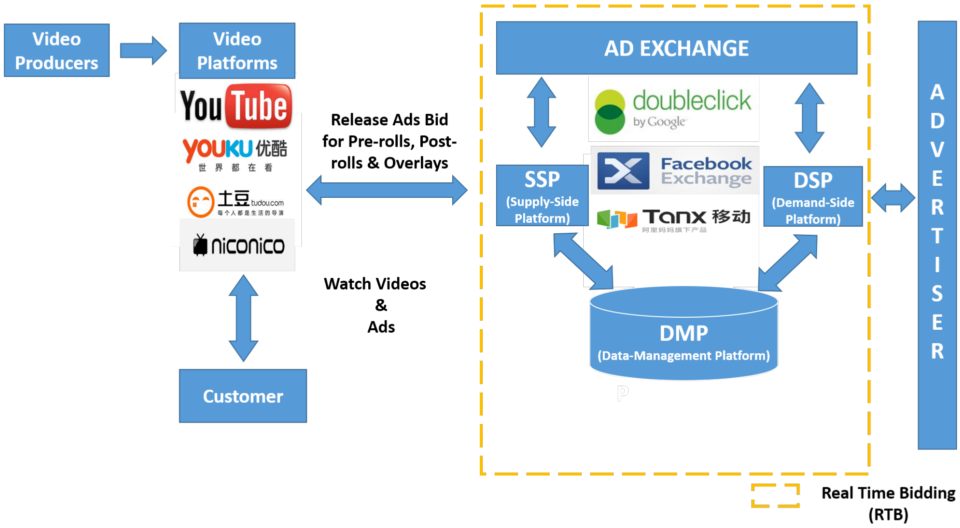

Figure 1: Existing working mechanism of Real-Time Bidding (RTB)

Research effort has been devoted to improving the working efficiency of RTB. For instance, Chen, Berkhin, Anderson and Devanur (2011) maximized publishers' revenue and achieved a good match between advertisers' campaigns and ad impressions with a novel algorithm.

RTB provides a bidding system for advertisers to purchase display inventory by the individual impression. Advertisers will only offer high price for those audience who they believe are most likely to be their target customers. Being targeting and cost efficiency, RTB is gradually becoming revolutionary in the online video advertising. ("What is RTB", n.d.).

Proposed Business Model Based on RTB and Online Video Interaction Technology

To create a frictionless and smooth watching experience for video viewers, we took the advantage of both object tracking algorithms and the RTB technology, and proposed a novel business model based on RTB and online video interaction technology.

Figure 2 shows the proposed business model based on RTB and online video interaction technology. To prove the technical feasibility of the proposed business model, a novel video interaction technology is implemented to pre-process videos. This technology recognizes and tracks objects with high commercial value in the video, and makes them clickable throughout the video. Once viewers click objects in videos, RTB is immediately triggered, and eventually ads which are closely associated with viewers' interests will be displayed. What's worth mentioning is that ads are displayed only when viewers click objects in videos, and viewers won't be forced to watch any boring or irrelevant ads. If viewers click nothing in videos, no ads will show up throughout the entire watching experience.

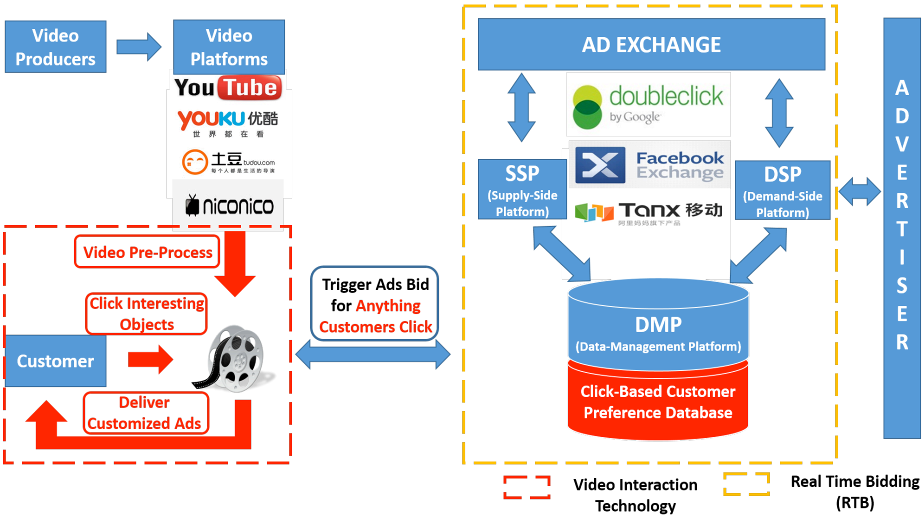

Figure 2: Proposed business model based on RTB and online video interaction technology

Online Video Interaction Technology

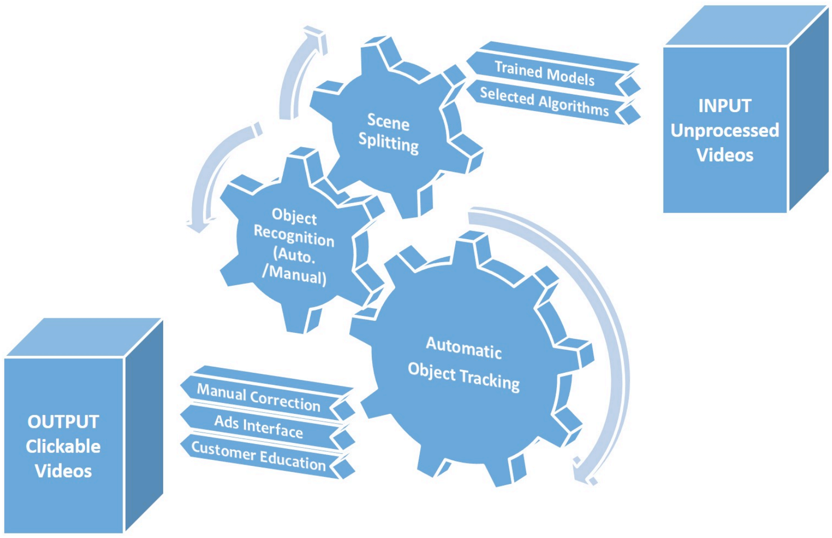

Figure 3: Proposed online video interaction technology

Figure 3 shows the flowchart of the proposed online video interaction technology. When an unprocessed video clip is input, scene-splitting algorithm starts to take action to split the original video into different scenes. Here, the 'scene' is defined to be a collection of consecutive frames which stay almost unchanged within most parts of the frames. Our need for the scene splitting stems from the fact that the tracking algorithm fails whenever there is an abrupt change in frames in the video.

The correlation coefficient (1) is used to decide whether there is an abrupt change between two frames,

$$r = \frac{\sum_m \sum_n (A_{mn} - \overline{A})(B_{mn} - \overline{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \overline{A})^2)(\sum_m \sum_n (B_{mn} - \overline{B})^2)}} \quad (1)$$

where A and B denote the image matrix of consecutive frames in a video. $\overline{A}$ and $\overline{B}$ are the mean of A and B respectively, and r denotes the correlation coefficient. To detect a scene changing using (1), we compare three values r_1, r_2 and r_3. We decide that a scene change happens only when $r_1 - r_2 > \varepsilon$ and $r_3 - r_2 > \varepsilon$, where ε is a predetermined threshold. Different video clips have different optimal ε value. In our 21 tests on car videos, 0.2 is an appropriate value (Zhang, Tan, Qin, Li & Tang, 2014).

Object recognition comes as the next procedure to find out any objects with high value of advertising. The recognition performance relies highly on the training model for certain objects. The object recognition can be finished either automatically or manually. Manual correction is available if automatic recognition fails.

After object recognition is over, the recognition results serve as inputs for automatic object tracking. We utilized the existing algorithms of Consensus-based Matching (CMT) and Tracking-Learning-Detection (TLD) (Nebehay & Pflugfelder, 2014; Kalal et al., 2012). There is a trade-off between overall runtime and accuracy. CMT is more accurate, but TLD can run faster (Wang, Liu, Liang, Guo & He, 2015). So it depends on the type of input video and the balance between runtime and accuracy which tracking algorithm users choose to use.

Finally, manual correction, ads interface and customer education are implemented so that the output video could be directly published to video viewers.

Click-Based Viewer Preference Database (CBVPD)

Another highlighting feature in the proposed model is the CBVPD. Currently, advertisers and DMPs have to speculate viewers' potential purchase intentions by analyzing such indirect information like browsing history, purchase record, etc. But with the help of the proposed technology, once viewers click any object in the video, a json. file with the following format will be transferred to the CBVPD:

(n, t, x, y),

where *n* is the object category (vehicle, cloth, jewelry, celebrity, etc.), *t* is the timestamp at which viewers click this object, and *x* and *y* stands for the location of the center of this object. With those information, viewers directly tell advertisers what they truly feel interested in and there is no need for advertisers to guess at all. CBVPD can significantly boost the efficiency of advertising because it is the most straightforward format of customer behavioral data, helping advertisers to reach their target customers more precisely with lowest cost.

Conclusion

To address the underlying issues of ineffective ads format and unsatisfactory video-watching experience in the online video advertising industry, a novel business model is proposed. The model highlights the use of real-time bidding (RTB), online video interaction technology and the Click-Based Viewer Preference Database (CBVPD).

A novel video interaction technology was developed to technically support the proposed business model, which covers a manual correction mechanism and algorithms including scene-splitting, object recognition and object tracking.

The proposed business model and the corresponding video interaction technology create a significant customer engagement, providing a deeper and more specific customer insight. With the help of RTB, advertisers can use a highly specific data to personalize the video-watching experience and expedite the purchase decision associated with the target customer.

Despite the frictionless watching experience and the cost efficiency that this proposed business model could bring to the industry, applying it to a real business world is still challenging. This is because so far customers are not used to clicking objects when watching videos. To solve this problem and make full use of the proposed model, the industry needs to educate customers with a systematic process and teach them the interactive format of online video advertising. In the near future, customers will be instantly connected to what they want to watch simply by a mouse click, a finger touch or even an eye blink. Be prepared to usher in and embrace a new era of online advertising.

As for future research work, the CBVPD should arouse researchers' interest. Having deep insight into the click pattern from different groups of viewers can certainly contribute to comprehensive understanding of customer behavioral and advertising efficiency. Research in this field will not only improve power in advertising, but also help to better understand human beings.

Acknowledgement

This research was supported by AVD Digital Media (AVD) and conducted at the University of Michigan-Shanghai Jiao Tong University Joint Institute (UM-SJTU). The authors thank Wayne Wu, Cayar Zhang and Andrew Ballen, who mentored, connected, and supported on AVD side. The authors thank Zhengrong Li, Siyu Tang, Qiyu Hu, Yungang Wang, Xinyi Zeng, Li Lai, Wei Wang, Jingkui Wang, Sheng Liang, Jianyu Liu, Xiaoming Guo and Sijia He from UM-SJTU Joint Institute for their contribution and efforts.

References

Blattberg, E. (2014, September 17). 5 Charts: The Growing Market for Digital Video Ads. Retrieved from http://digiday.com/platforms/5-charts-growing-market-digital-video-ads/

Chen, Y., Berkhin P., Anderson B., Devanur, R. N. (2011). Real-Time Bidding Algorithms for Performance-Based Display Ad Allocation. *Proceedings of the ACM Conference on KDD.*

Computer Vision Lab (n.d.). Visual Tracking Benchmark. Retrieved from http://cvlab.hanyang.ac.kr/index.html

Dinh, T. B., Vo, N., & Medioni, G. (2011, June). Context tracker: Exploring supporters and distracters in unconstrained environments. *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 1177-1184.

Etherington, D. (2014, October 28). YouTube Considering an Ad-Free, Subscription-Based Version. Retrieved from http://techcrunch.com/2014/10/28/youtube-considering-an-ad-free subscription-based version/?ncid=rss&utm_source=tuicool

Google (2011). The Arrival of Real-Time Bidding and What it means for Media Buyers. *White paper describes Google's view of real-time bidding (RTB) for online display advertising.*

Hare, S., Saffari, A., & Torr, P. H. (2011, November). Struck: Structured output tracking with kernels. *2011 IEEE International Conference on Computer Vision (ICCV),* 263-270

Henriques, J. F., Caseiro, R., Martins, P., & Batista, J. (2015). High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis & Machine Intelligence 37(3),* 583-596

Hopkins, J. (2006, October 11). Surprise! There's a Third TouTube Cofounder. Retrieved from http://usatoday30.usatoday.com/tech/news/2006-10-11-youtube-karim_x.htm

Hu, Q., Lai L., Wang W., Wang Y., Zeng X., (2015, July 28) Automatic Video Object Tracking Phase II

iResearch (2014, January 26). China Online Video Market Hit an All-time High of 12.81 Bn Yuan in 2013. Retrieved from http://www.iresearchchina.com/views/5423.html

Kalal, Z., Mikolajczyk, K., & Matas, J. (2012). Tracking-learning-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(7),* 1409-1422.

Kwon, J., & Lee, K. M. (2011, November). Tracking by sampling trackers. *2011 IEEE International Conference on Computer Vision (ICCV),* 1195-1202

Main Street RTB (n.d.). What is RTB. Retrieved from http://mainstreetrtb.com/what-is-rtb/

Nebehay, G., & Pflugfelder, R. (2014, March). Consensus-based Matching and Tracking of Keypoints for Object Tracking, *2014 IEEE Winter Conference on Application of Computer Vision (WACV),* 862-869

Russo, A. (May 12, 2015). Online Video Advertising Revenue Doubles in Three Years, Says IHS and Vidiro. Retrieved from http://press.ihs.com/press-release/technology/online-video-advertising-revenue-doubles-three-years-says-ihs-and-vidiro

VanBoskirk, S. (2007). US Interactive Marketing Forecast, 2007 To 2012. *Forrester Research, for Interactive Marketing Professionals.*

Wang J., Liu J., Liang S., Guo X., & He S., (2015, July). A Novel Video-Interactive Technology for Future Advertising.

Wu, Y., Lim, J., & Yang, M. H. (2013, June). Online object tracking: A benchmark. *2013 IEEE Conference on Computer vision and pattern recognition (CVPR),* 2411-2418.

Yuan, S., Wang, J., Zhao, X. (2013). Real-time Bidding for Online Advertising: Measurement and Analysis. *Proceedings of the ACM Conference on ADKDD.*

Zhang, J., Feng, C., Chang, V., Tan, J., & Qin, T. (2015). A Novel Real-Time Bidding Model Based on Video Processing Technology. *Accepted by International Conference on Innovation and Management.*

Zhang, J., Tan, J., Qin, T., Li Z., & Tang, S. (2014). A Video Processing Tool for Car Recognition and Tracking

Zhang, W., Yuan, S., & Wang, J. (2014, August). Optimal real-time bidding for display advertising. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining,* 1077-1086.

Zhong, W., Lu, H., & Yang, M. H. (2012, June). Robust object tracking via sparsity-based collaborative model. *2012 IEEE Conference on Computer vision and pattern recognition (CVPR)*, 1838-1845

Contact email: jerry_zjs@sjtu.edu.cn, vincent.chang@sjtu.edu.cn