P

Performance Measures and RTB Optimization

Wenxue Huang
Guangzhou University, China

Yuanyi Pan
InferSystems Corp., Canada

Jianhong Wu
York University, Canada

INTRODUCTION

Rare events are the events occurring with very low probabilities but with significant consequences. It often refers to the preferred category in a target, mostly binary, variable. Statistical performance measures for binary variable targeting have been studied for many years (Aitchison, 1975; Efron, 1978; Hosmer & Lemeshow, 1989; Landwehr, Pregibon & Shoemaker, 1984; Menard, 2000).

A fundamental issue in rare events targeting is how to evaluate its performance. The Gini coefficient, or Gini index, is widely used to measure the performance of rare event targeting models. We have found that the Gini index is not quite consistent with the Pareto 80-20 rule, which is critical in most rare event targeting based business. Two new indices and their generalization were porposed in our article (Huang, Pan & Wu, 2012) to correct the Gini coefficient's inherent shortbacks from business application point of view. (Pan, 2012) proposed its application in real time bidding (RTB) optimization system. This chapter is in general a reciting to these two articles plus certain additional information.

BACKGROUND

The industrial applications of rare events targeting, besides online advertising, include banking and insurance modeling (King & Zeng, 2001; Pednault, Rosen & Apte, 2000; Vasu and Ravi, 2007), credit card fraud detection (Chan & Stolfo, 1998), network intrusion detection (Dreger, Feldmann, Paxson, & Sommer, 2004), and direct marketing (Ling, C., Ling, C.X. & Li, 1998) among others.

The industry of online advertising is growing fast. According to emarketer.com, US online ad spending will reach $62 billion in 2016 from $32 billion in 2011. Meanwhile, traditional media ad spending is descending or of no significant increasing. The newspaper and magazine ad speding was approxinately $36 billion and the TV ad spending was $60.7 billion in 2011, and these are expected to be $32.3 billion and $72 billion respectively in 2016.

The display advertisements, including online video, banner ads, rich media and sponsorships seems to have a prospective future, though the search advertising seems to have a bigger market for now, as shown in Figures 1 and 2 that illustrate the spending and the growth of spending in each format.

The demand side of the display advertising ecosystem, as shown in Figure 3, can be generalized as *advertisers* while the supply side being as the *publishers*. The advertisers have the demand to display ads and the publishers meet these demands. In the following, we adopt the abbreviation DSP and SSP for the platform for the demand side and the platform for the supply side respectively.

Two business models in the online advertising market were suggested in (Asdemir, Kumar,

DOI: 10.4018/978-1-4666-5202-6.ch165

Figure 1. Spending

US Online Ad Spending, by Format, 2010-2015
billions

	2010	2011	2012	2013	2014	2015
Search	$12.00	$14.38	$17.03	$18.85	$20.19	$21.53
Banner ads	$6.23	$7.61	$8.94	$9.93	$10.97	$11.73
Classifieds and directories	$2.60	$3.00	$3.35	$3.65	$3.98	$4.29
Video	$1.42	$2.16	$3.09	$4.20	$5.64	$7.11
Rich media	$1.54	$1.66	$1.73	$1.74	$1.73	$1.68
Lead generation	$1.34	$1.42	$1.45	$1.47	$1.50	$1.52
Sponsorships	$0.72	$0.91	$1.05	$1.18	$1.32	$1.47
Email	$0.20	$0.16	$0.16	$0.17	$0.17	$0.18
Total	**$26.04**	**$31.30**	**$36.80**	**$41.20**	**$45.50**	**$49.50**

Note: eMarketer benchmarks its US online ad spending projections against the IAB/PwC data, for which the last full year measured was 2010
Source: eMarketer, June 2011

128163 www.eMarketer.com

Figure 2. Spending growth

US Online Ad Spending Growth, by Format, 2010-2015
% change

	2010	2011	2012	2013	2014	2015
Video	39.6%	52.1%	43.1%	35.9%	34.3%	26.0%
Sponsorships	87.5%	26.4%	16.0%	12.3%	11.6%	11.0%
Banner ads	23.1%	22.1%	17.6%	11.0%	10.4%	7.0%
Search	12.2%	19.8%	18.4%	10.7%	7.1%	6.6%
Classifieds and directories	15.2%	15.7%	11.4%	9.0%	8.9%	7.8%
Rich media	2.2%	7.9%	4.3%	0.8%	-0.8%	-2.7%
Lead generation	-7.7%	6.1%	1.8%	1.7%	2.1%	1.2%
Email	-33.2%	-16.5%	-0.5%	3.3%	3.4%	3.1%
Total	**14.9%**	**20.2%**	**17.6%**	**12.0%**	**10.4%**	**8.8%**

Note: eMarketer benchmarks its US online ad spending projections against the IAB/PwC data, for which the last full year measured was 2010
Source: eMarketer, June 2011

128167 www.eMarketer.com

& Jacob, 2011). The first one is an input based cost model, where the number of impressions (i.e. the number of times the ads are displayed in front of the viewers) is the base for the charges. This is commonly referred to as a CPM (cost per thousand impressions) model. The second model is performance based, and it charges by the times of activities that the viewers behave after the ads are displayed. This is called a CPC (cost per click) model if the targeted behavior is clicking the ad. In contrast, the model is called a CPA (cost per action) model if the targeted behavior needs further action after the click, mostly submiting a transaction or the viewer's personal information. This kind of model is also called a CPL (cost per lead) model (pontiflex.com, 2008).

A very popular method by which DSP acquires impressions from SSP is real time bidding (RTB)