Impact of Ad Impressions on Dynamic Commercial Actions: Value Attribution in Marketing Campaigns

Joel Barajas*
UC, Santa Cruz
Santa Cruz CA, USA
jbarajas@soe.ucsc.edu

Ram Akella
UC, Santa Cruz
Santa Cruz CA, USA
akella@soe.ucsc.edu

Marius Holtan
AOL Research
Palo Alto CA, USA
marius.holtan@teamaol.com

Jaimie Kwon
AOL Research
Palo Alto CA, USA
jaimie.kwon@teamaol.com

Aaron Flores
AOL Research
Palo Alto CA, USA
aaron.flores@teamaol.com

Victor Andrei
AOL Research
Palo Alto CA, USA
victor.andrei@teamaol.com

ABSTRACT

We develop a descriptive method to estimate the impact of ad impressions on commercial actions dynamically without tracking cookies. We analyze 2,885 campaigns for 1,251 products from the *Advertising.com* ad network. We compare our method with A/B testing for 2 campaigns, and with a public synthetic dataset.

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*; G.3 [**Mathematics of Computing**]: Probability and Statistics—*Time Series Analysis*

General Terms

Algorithms, Economics, Management, Measurement

Keywords

Attribution, DLM, Marketing, Online Display Advertising

1. INTRODUCTION

Evaluating the effectiveness of marketing campaigns is a key problem in Online Display Advertising. Under the Cost-Per-Action (CPA) model, advertisers share the number of online commercial actions with the ad network. However, a significant number of users are not tracked because they either delete or reject cookies outright. Approximately 17% of *Advertising.com* users are not tracked via cookies[1].

We develop an interpretative method, based on Dynamic Linear Models (DLM)[3], to estimate the impact of ad impressions on actions without cookies. We incorporate *persistence* of campaign effects on actions assuming a decay factor. We relax the assumption of a linear impact of ads on actions using the log-transformation. We account for outliers with long-tail distributions fitted automatically for each observation[3]. Our method uses aggregate data and is simple to implement without expensive infrastructure. We measure model fitting and prediction with 4 model variants for 2,885 campaigns and 1,251 products from the *Advertising.com* network. We compare our results with A/B testing, and validate our model with PROMO[2] dataset.

2. RELATED WORK

For reliable cookies, running experiments (A/B testing)[2], and correcting observational data[1] have been proposed to evaluate campaigns. These methods rely heavily on cookies, human intervention or user features. In contrast, we model the impact of ad impressions on actions to provide daily estimates without cookies or user features.

3. METHODOLOGY

We use the indices: t=1:T for time, c=1:N campaigns, s for samples, and k for steps ahead in forecast. We define: Y_t number of actions; X_t number of ad impressions; ξ_t cumulative effect of impressions on Y_t; ψ_t impact contribution per impression; $\lambda \in \{0, 0.88\}$ decay rate of ξ_t (persistence). DLM variables: θ_t latent state; V_t noise variance for ν_t; W_t covariance matrix of state evolution w_t. We define a DLM:

$$\begin{array}{ll} Y_t = F_t'\theta_t + \nu_t & \nu_t \sim N(0, V_t) \\ \theta_t = G_t\theta_{t-1} + w_t & w_t \sim N(0, W_t) \end{array} \quad (1)$$

We define the model for N campaigns:

$$\begin{aligned} Y_t &= \sum_{c=1}^{N} \xi_t^{(c)} + \nu_t \\ \xi_t^{(c)} &= \lambda^{(c)}\xi_{t-1}^{(c)} + \psi_t^{(c)}X_t^{(c)} + w_t^{(\xi,c)} \\ \psi_t^{(c)} &= \psi_{t-1}^{(c)} + w_t^{(\psi,c)} \end{aligned} \quad (2)$$

This is expressed by combining DLMs for each campaign c:

$$\theta_t'^{(c)} = [\xi_t^{(c)}, \psi_t^{(c)}], \qquad F'^{(c)} = [1, 0]$$

$$G_t^{(c)} = \begin{bmatrix} \lambda^{(c)} & X_t^{(c)} \\ 0 & 1 \end{bmatrix}, \quad W_t^{(c)} = \begin{bmatrix} W_\xi^{(c)} + \left(X_t^{(c)}\right)^2 W_\psi^{(c)} & X_t^{(c)} W_\psi^{(c)} \\ X_t^{(c)} W_\psi^{(c)} & W_\psi^{(c)} \end{bmatrix}$$

We expand this model, $M^{(0:N)}$, using 2 base models: a random walk and a seasonal weekly model. Mlog model uses the log transformation to relax the linear relationship between actions and impressions. Algorithm 1 shows the model $M\omega$ that handles outliers. Here, Γ and $Mult$ are the Gamma and Multinomial distributions. Algorithm 2 shows the Gibbs sampling steps used to fit the model, and defines the variable sets of interest. We sample $\theta_{1:T}|\Phi, \Omega, D_{1:T}$ based on Forward Filtering Backward Sampling [3]. For $\Phi|\theta_{1:T}, \Omega, D_{1:T}$

*Main contact.
[1]AOL Research and Development internal memo.

[2]Available online at: `http://www.causality.inf.ethz.ch/repository.php?id=2`

WWW 2012 Companion, April 16–20, 2012, Lyon, France.
ACM 978-1-4503-1230-1/12/04.

Algorithm 1 Generative Model to Handle Outliers

Draw $p|\alpha \sim Dirichlet(\alpha)$
for $t \leftarrow 1$ to T **do**
  Draw $\eta_t|p \sim Mult(1,p)$, $\omega_t|\eta_t \sim \Gamma(\frac{\eta_t}{2}, \frac{\eta_t}{2})$, $V_t{=}\omega_t^{-1}V$
end for

Algorithm 2 Gibbs Sampling Algorithm

Define $D_{1:T} = \left\{Y_{1:T}, X_{1:T}^{(1:N)}\right\}$, $\Omega = \{\omega_{1:T}, \eta_{1:T}, p\}$
Define $\Phi = \left\{\lambda^{(1:N)}, W_{\psi}^{(1:N)}, W_{\xi}^{(1:N)}, W^{(0)}, V\right\}$
for $s \leftarrow 1$ to $N_0 + N_s$ **do**
  Draw $\theta_{1:T}^s \sim p\left(\theta_{1:T}|\Phi^{s-1}, \Omega^{s-1}, D_{1:T}\right)$
  Draw $\Phi^s \sim p\left(\Phi|\theta_{1:T}^s, \Omega^{s-1}, D_{1:T}\right)$
  Draw $\Omega^s \sim p\left(\Omega|\theta_{1:T}^s, \Phi^s D_{1:T}\right)$
end for

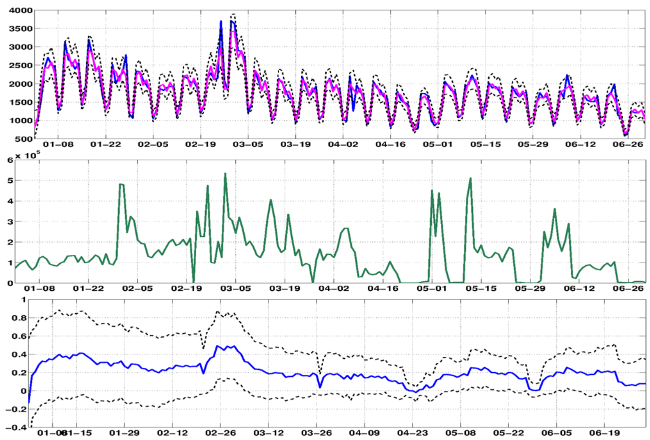

Figure 1: From top to bottom: model fitting results, daily number of impressions, proportion of actions attributed to impressions. *X*-axis is time in dates.

we use Inverse Gamma conjugate priors and Truncated Normals for λ. $\Omega|\theta_{1:T}, \Phi, D_{1:T}$ is sampled based on Algorithm 1. To sample η_t, we estimate the normalization constant. For evaluation, we estimate one step ahead forecast, $Y_t^{k=1}|D_{t-1}$. $\hat{Y}_t$ and $\hat{\omega}_t$ are the posterior medians. We use mean relative squared error, MRSE, $e{=}(Y_t{-}\hat{Y}_t)/Y_t$. For attribution, we find the proportion of actions described by campaign c. We also estimate the variability attributed to a campaign, $R^2(c)$, respect to data variance, base model squared error $M^{(0)}$, and full model squared error without campaign c, $M^{(0:N\neg c)}$.

4. RESULTS

We test the models $M\omega$ and $M\omega$log for the 2 base models defined. We analyze 2,885 campaigns associated with 1,251 products during six months. Fig 1 shows the model fitting and the proportion of actions attributed to campaign impressions. We use $M\omega$log with weekly seasonal base model. We observe that the peak in the first half of the action series is attributed to a gradual increase of daily impressions. Table 1 shows the fitting results. Better performance is reported when the log transformation is included suggesting a non-linear relationship between actions and impressions. This model reports the highest campaign percentage with non-significant effect according to Table 2. Table 3 depicts the mean and variance over campaigns for R^2. Here, $R^2(c|\text{var}(Y_t))$ is lower for the weekly seasonal base model because actions are attributed to the day of the week.

We compare our method with A/B testing, which is expensive and requires significant human intervention, for 2 campaigns in Table 4. Here, confidence intervals for A/B testing are not tight due to the sparsity of actions. Both methods report one positive significant campaign at 90% confidence level and one leaning towards positive effect. We also test our method with the PROMO dataset to compare with a ground truth. We use products with less than 6 relevant campaigns for 365 days. We detect **84.6%** of effective campaigns correctly, and **73.2%** of the days a campaign is effective per product.

Table 1: Model evaluation results, scaled by 10^{-2}, averaged over products. 95% confidence intervals are shown.

Model	Fitted	Forecast	Fitted $\omega_t = 1$	Forecast $\omega_t = 1$
Random Walk Base model, MRSE				
Mω	7.91 ± 1.85	61.77 ± 7.13	14.87 ± 2.40	72.13 ± 7.58
Mω log	1.33 ± 0.32	**13.25±2.21**	5.49 ± 1.02	20.00 ± 2.57
Weekly Seasonal Base model, MRSE				
Mω	8.26 ± 2.13	61.13 ± 7.34	12.65 ± 2.11	70.75 ± 7.67
Mω log	**0.72±0.14**	15.51 ± 2.81	**3.92±0.92**	21.20 ± 3.12

Table 2: Averaged campaign evaluation results. Distribution of campaign effect significance (%).

Model	% attributed	Campaign Significance		
		(+)	(-)	(±)
Random Walk Base model				
Mω	14.07 ± 1.36	23.13	0.71	76.09
Mω log	21.31 ± 1.63	18.65	0.58	**80.71**
Weekly Seasonal Base model				
Mω	10.39 ± 1.23	19.66	1.31	78.96
Mω log	19.84 ± 1.64	14.83	0.60	**83.98**

Table 3: Attributed variability results.

	Random Walk		Weekly Seasonal	
Measure	Mean	Std Dev	Mean	Std Dev
$R^2(c\|\text{var}(Y_t))$	0.1241	0.2704	**0.0667**	0.1750
$R^2(c\|M^{(0)})$	0.2804	0.3827	0.3002	0.3701
$R^2(c\|M^{(0:N\neg c)})$	0.4967	0.4114	0.4703	0.3729

Table 4: A/B testing results compared to the attribution given by Mωlog for the RandWalk base model.

Method	Campaign 1			Campaign 2		
	Low	Med	High	Low	Med	High
A/B	0.009	0.199	0.458	-0.034	0.115	0.312
Mωlog	0.013	0.051	0.107	-0.049	0.347	0.809

5. DISCUSSION AND FUTURE WORK

We have presented a descriptive DLM based approach to measure the effects of campaigns on actions without tracking users. Our experiments show that a model in the log-scale is more suitable to describe the behavior of actions, and a seasonal base model gives less attribution to campaigns. We observed several campaigns with non-significant average effect on actions, which is consistent with A/B testing results. Our ultimate goal is to provide daily significant estimates of the effects of campaigns on sparse actions.

6. ACKNOWLEDGMENTS

This work is partially funded by CONACYT UC-MEXUS grant 194880, CITRIS and AOL Faculty Award.

7. REFERENCES

[1] D. Chan, R. Ge, O. Gershony, T. Hesterberg, and D. Lambert. Evaluating online ad campaigns in a pipeline: causal models at scale. In *Proceedings of SIGKDD*, pages 7–16. ACM, 2010.

[2] R. A. Lewis, J. M. Rao, and D. H. Reiley. Here, there, and everywhere: correlated online behaviors can lead to overestimates of the effects of advertising. In *Proceedings of WWW2011*, pages 157–166. ACM, 2011.

[3] G. Petris, S. Petrone, and P. Campagnoli. *Dynamic Linear Models with R*. use R! Springer-Verlag, 2009.