

# Transfert entre campagnes de Real-Time Bidding

Charles Monzani \*, Martial Hue\*\*  
Patrick Gallinari\*

\*Sorbonne Université, CNRS, LIP6, 4 place Jussieu, 75005 Paris

\*\*Tradelab, 10 rue Henner, 75009 Paris

**Résumé.** La réussite d'une campagne publicitaire sur internet passe par une modélisation efficace des internautes : pour cela de nombreuses méthodes sont aujourd'hui utilisées, notamment des réseaux de neurones récurrents, qui permettent de prendre en compte les séquences d'action des internautes. Cependant il est nécessaire d'avoir déjà acquis un conséquent volume de données avant qu'ils ne soient vraiment efficaces. Ainsi au début de chaque campagne il y a une phase de pure collecte de données : les dépenses publicitaires sont effectuées de manière aléatoire afin d'acquérir des données non biaisées, de plus on ne sait pas encore quel type d'internautes viser : on parle de problème de Cold-Start.

Pour éviter ce problème, nous proposons de réutiliser dans données acquises sur des campagnes antérieures, en reprenant un réseau de neurones déjà entraîné.

## 1 Introduction

Afficher la bonne publicité, au bon internaute, au bon moment, est la clef de la réussite d'une campagne publicitaire en *Real-Time Bidding*, cette technologie qui permet la mise aux enchères en temps-réel d'un espace publicitaire sur une page (défini par l'adresse de la page, l'emplacement et le format de la bannière, ainsi que l'identifiant anonymisé de l'internaute). Deux métriques sont utilisées a posteriori pour évaluer une campagne : le coût par clic (CPC), qui correspond à la somme totale dépensée en achat de bannières divisée par le nombre de clics obtenus sur ces bannières ; et le coût par acquisition (CPA), qui correspond à la somme totale dépensée divisée par le nombre d'achats réalisés par des internautes ayant vu une bannière publicitaire de la marque (sans qu'ils aient nécessairement cliqué dessus). On parle ainsi de taux de clic (CTR) et de taux de conversion (CVR) pour, respectivement, la probabilité qu'un internaute clique sur la bannière ; et la probabilité qu'il réalise une conversion.

Pour éviter d'afficher une bannière à des internautes à faible probabilité de conversion, il est indispensable d'avoir une méthode efficace de représentation des internautes ainsi qu'une méthode efficace de calcul du taux de conversion. En raison du volume faible de conversion pour une campagne, il est généralement plus simple de calculer le taux de clic. L'association d'un *feature-engineering* poussé et d'une régression logistique a longtemps été la méthode de choix pour le calcul du taux de clic (Agarwal