Conception d'un modèle généraliste pour l'évaluation d'un test A/B

Emmanuelle Claeys*, Pierre Gançarski *
Myriam Maumy-Bertrand** Hubert Wassner***

*ICube – Université de Strasbourg – 67412 – Illkirch – France
{claeys, gançarski}@unistra.fr

**IRMA – Université de Strasbourg – 67084 – Strasbourg – France
mmaumy@math.unistra.fr

***Entreprise AB Tasty – 3 Impasse de la Planchette – 75003 – Paris – France
hubert@abtasty.com

1 La problématique d'un test A/B

Un test *A/B* (ou *A/B testing*¹), consiste à créer plusieurs versions d'un même objet, dans le but de vérifier une hypothèse marketing. Par exemple, il s'agit de tester auprès d'un échantillon d'internautes, nommés visiteurs, durant une période donnée, différentes variations (généralement deux) d'une même page web (nommées variation *A* et variation *B*) proposées par un e-commerçant. Le but est d'identifier celle qui donne les meilleurs résultats par rapport à un objectif fixé par l'e-commerçant en amont du test. Une contrainte commune, lors d'un test *A/B*, est que lorsqu'un visiteur est affecté à une variation, elle lui est affectée pour toute la durée de sa visite, jusqu'à la fin du test. Il est donc impossible de connaître de façon sûre le comportement qu'il aurait eu sur l'autre variation. De plus, les visiteurs testés sur la variation *A* étant différents de ceux testés sur la variation *B*, il est nécessaire de définir un moyen de comparer leurs comportements respectifs. Le site étant généralement en production, il est intéressant de garder un compromis entre l'exploration (évaluer les différentes variations) et l'exploitation du test (favoriser la variation avec les meilleurs résultats). Sachant qu'il est impossible de modifier la variation appliquée à un visiteur, il est nécessaire d'affecter au mieux à un nouveau visiteur, la version *A* ou *B* du test.

2 Modélisation et approche multi-contextes

Plusieurs verrous majeurs sont à lever : le modèle décisionnel doit prendre en compte la performance et la pertinence de ses choix. Il devra adapter ses choix en fonction du type de visiteurs et du contexte du test. Le modèle devra être également capable d'apprendre de ses succès/erreurs et de s'auto-corriger.

1. Le terme *A/B* étant un terme générique pour qualifier tout type de tests