Affichage de publicités sur des portails *web*

{Victor.Gabillon, Jérémie.Mary, Philippe.Preux} @inria.fr

Le problème

Établir une politique d'affichage de messages publicitaires sur des sites internet dans un **environnement dynamique** et avec des **ressources limitées**.

Modèle « Payement au clic » : Les publicitaires nous payent à chaque clic produit par leurs pubs. Les affichages qui ne produisent pas de clics sont perdus. Il faut donc bien cibler les affichages.

Buts :

– Déterminer à quel type de visiteur plaît chaque pub. → Bandits

– Gérer les contraintes de budget.
→ Programmation linéaire

Notations

- $\mathbf{K^t}$ publicités $\mathbf{Ad_1, Ad_2, \ldots, Ad_{K^t}}$.
- Chaque pub Ad_j dispose d'un budget restant $\mathbf{B_j^t}$ de clics pendant sa durée de vie restante $\mathbf{L_j^t}$.
- Les visites au portail *web* sont segmentées en $\mathbf{N}$ profils notés $Profile_i, i \in 1 \mathinner{..} N$.
- Fréquentations par les profils : $\mathbf{R_1, \ldots, R_N}$.
- Les inconnues sont les $\mathbf{p_{i,j}}$ probabilités de clic du profile i sur Ad_j.

Hybridations Bandits+PL

Les informations contextuelles disponibles sur les visites permettent d'estimer des probabilités de clic pour chaque catégorie de visiteurs. Cependant, cet apprentissage étant réalisé en ligne, il faut combiner l'estimation de ces paramètres inconnus avec l'exploitation de leur estimation courante. Ce problème peut-être formulé dans le cadre des bandits multi-bras. Néanmoins, le fait que les budgets de clics soient finis rend le problème combinatoire, dans un contexte incertain, stochastique et évoluant dans le temps. Ainsi, simplement afficher la publicité qui a la plus grande probabilité de clic n'est pas la stratégie optimale. Une technique envisageable est la programmation linéaire (PL) basée sur l'estimation courante des probabilités de clics. Aussi, nous proposons de mixer la programmation linéaire aux bandits.

Environnement Dynamique

De nouvelles pubs apparaissent créées par un modèle génératif.

La programmation linéaire ?

PL planifie avec un nombre fini de visiteurs. Ce nombre devient un paramètre de l'algorithme, l' Horizon.

Dans la suite…

PL va nous permettre de planifier l'allocation de nos pubs actuelles Sans essayer de prédire le futur.

Conclusions

– Amélioration quelque soit l'horizon choisi.

– Reste à fixer l'horizon optimal. Découverte des dépendances aux paramètres du problème dans un cadre réaliste.

– Demande une puissance de calcul raisonnable.

Programme linéaire

Maximiser le nombre de clics $\sum_{\substack{1\le i\le N\\ 1\le j\le K^t}} x_{i,j}\, p_{i,j}$

Sous contraintes

Budget de clics	$\sum_{1\le i\le N} x_{i,j}\, p_{i,j}$	$\le B_j^t$	$\forall j \in \{1 \mathinner{..} K^t\}$
Répartition des profiles	$\sum_{1\le j\le K^t} x_{i,j}$	$\le R_i * H$	$\forall i \in \{1 \mathinner{..} N\}$
Limites d'affichage	$\sum_{1\le k\le j} x_{i,k}$	$\le L_j^t * R_i$	$\forall i \in \{1 \mathinner{..} N\}$, $\forall j \in \{1 \mathinner{..} K^t\}$
	$x_{i,j}$	≥ 0	$\forall i \in \{1 \mathinner{..} N\}$, $\forall j \in \{1 \mathinner{..} K^t\}$

Principe de l'algorithme

Itération au temps t:

Politique d'allocation:

Estimateurs courants: $\widehat{p}_{i,j}^t$

Résoudre le programme linéaire: $x^t = PL(\widehat{p}^t, B^t, L^t, R, H)$

Calculer les probabilités d'allocation: $s_{i,j}^t = \frac{x_{i,j}^t}{\|x_i^t\|}\ \forall i \in \{1 \ldots N\}, \forall j \in \{1 \ldots K^t\}$

avec $\|x_i^t\| = \sum_{j=1}^{K^t} x_{i,j}^t$

Visite: Pour notre t^e visite, un $Profile_i$ apparaît.

Affichage: Une publicité Ad_l est choisie selon la densité $s_{i,\cdot}^t$.

Clic: Un clic se produit ou non.

Mise à jour: $\widehat{p}_{i,l}^t$ est mis à jour.

$L_k^{t+1} \leftarrow L_k^t - 1 \qquad \forall\, k \in \{1, \ldots, K^t\}$

$B_l^{t+1} \leftarrow B_l^t - 1 \qquad$ si un clic s'est produit.

Une nouvelle publicité apparaît avec probabilité u (Mettre alors à jour K^t).

Nécessité d'introduire de l'exploration dans cette routine : Soit en deviant légèrement de la planification, soit en modifiant directement les estimateurs (voir article)

Rôle de l'horizon

	Ad 1	Ad 2
Profile 1	0,8	0,1
Profile 2	0,8	0,5

À gauche, le tableau représente les probabilités de clic (augmentées mille fois) des profiles de visite sur chaque publicité. En dessous les politiques d'allocation pour différent choix d'horizon H sont comparés.

H = **20**, R_i : $1/2 \rightarrow$ Profile 1, $1/2 \rightarrow$ Profile 2

Budget	100	100
	Ad 1	Ad 2
Profile 1	**10**	0
Profile 2	**10**	0

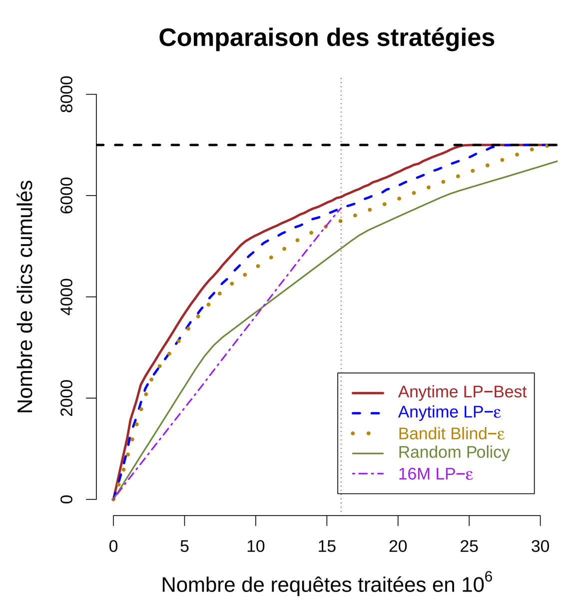
Planification avec $H = 20 \rightarrow$ **Gourmand**

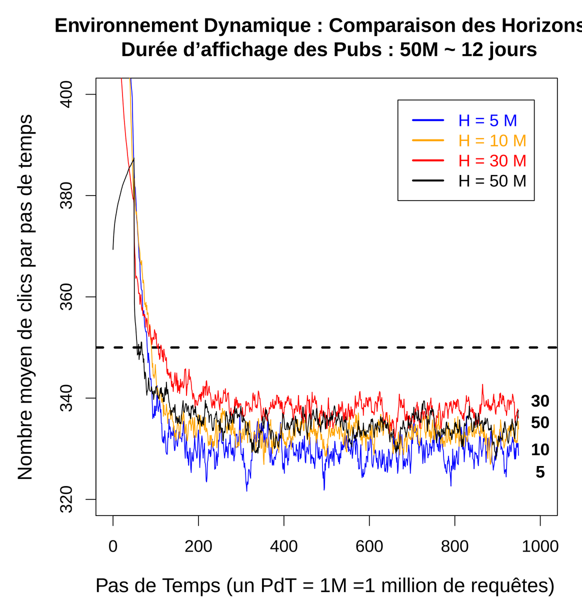
H = **300**, R_i : $1/2 \rightarrow$ Profile 1, $1/2 \rightarrow$ Profile 2

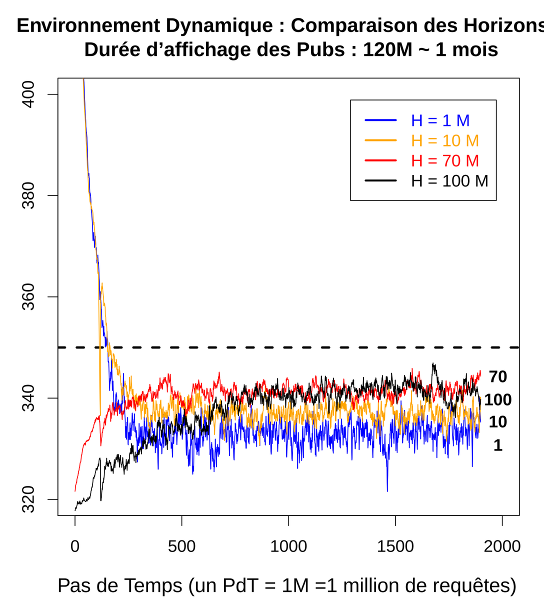
Budget	100	100
	Ad 1	Ad 2
Profile 1	**125**	**25**
Profile 2	0	**150**

Planification avec $H = 300 \rightarrow$ **Prévoyant**

Expériences







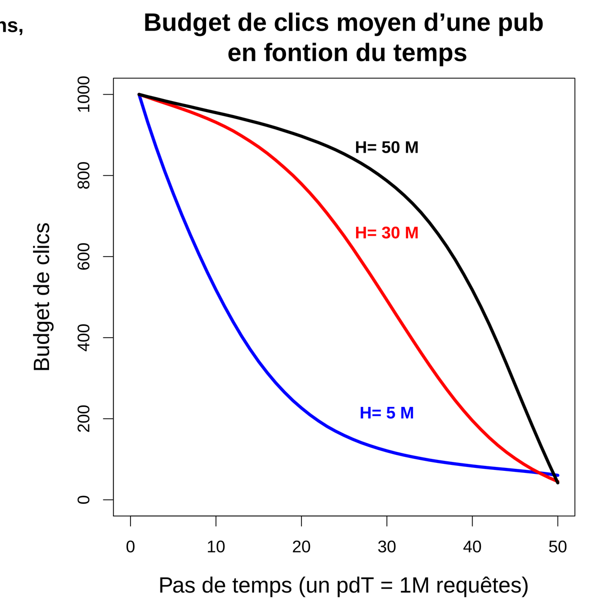


Figure 1 : Le bandit-simplex est proche de l'optimum.
Figure 2 & 3 : Quelque soit l'horizon les performances sont meilleures que les bandits parallèles. L'horizon optimal dépend de la fréquence d'arrivée des pubs et de leur durée de vie. Plus la fréquence de renouvellement est grande moins il est nécessaire d'agrandir l'horizon. Plus la durée d'affichage est grande, plus H doit l'être. Un horizon très faible est peu couteux en temps de calcul (simplex e plus rapide résoudre).
Figure 4 : L'horizon optimal correspond aussi à l'écoulement le plus uniforme.

Remerciements

Ce travail a été partiellement financé par les Orange Labs *via* le contrat de recherche externalisé numéro 46 146 063 - 8360.

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE
centre de recherche LILLE - NORD EUROPE