Algorithmique des réseaux sociaux **2013/2014**

Cours 2 — 14/15 Octobre

Enseignant: Marc Lelarge *Scribe: Margot Calbrix*

Pour information

- Page web du cours
  `http://www.di.ens.fr/~lelarge/soc.html`

2.1 Enchères et liens sponsorisés

2.1.1 Historique

À sa création, les fondateurs de Google n'avaient pas prévu de modèle économique. Quand ils ont publié leur algorithme PageRank, à la fin des années 90, ils ont tout d'abord cherché à le vendre à Yahoo !, qui n'étaient pas intéressés. Au même moment, une autre entreprise, GoTo.com, Inc, propose un moteur de recherche basé sur des enchères : les sites enchérissent pour leur place dans le résultat d'une requète. Mais les sites qui paient le plus ne sont pas forcément les sites les plus pertients. Google reprend le principe en 2001, en le restreignant à un petit nombre de publicités affichées avec les résultats de la requète.

2.1.2 Quelques caractéristiques des enchères pour les liens sponsorisés

Ces enchères ont des spécificités propres au contexte des recherches sur Internet. En effet, le bien est ici un emplacement qui peut être sans cesse remis aux enchères. Également, les offres pour un mot-clé étant faite en ligne peuvent changer à tout moment.

Il y a deux façons de concevoir le prix d'un lien sponsorisé. L'entreprise qui fait sa publicité s'intéresse au coût de la campagne de pub, c'est-à-dire le coût d'attirer un client qui fait un achat. Elle veut donc payer uniquement pour les clients que la publicité lui rapporte (modèle qu'utilise Amazon). Le moteur de recherche, a contrario, veut facturer chaque lien affiché, c'est-à-dire à chaque fois qu'un client potentiel a vu la publicité (modèle de la presse écrite). Le modèle utilisé par les moteurs de recherche est un compromis entre ces deux modèles. Le publicitaire paie le moteur de recherche à chaque fois que l'utilisateur clique sur le lien sponsorisé : c'est le pay-per-click.

Une notion importante dans ce modèle est donc le clickthrough rate (CTR). Étant donné un emplacement de publicité, le CTR est le rapport entre le nombre de clics sur une publicité et son nombre d'affichages. Le CTR est évidemment fonction de l'emplacement de la publicité, mais aussi de la pub en elle même (notamment de son rapport avec la requête, par exemple). pour la suite, nous considérerons que le CTR est uniquement fonction de l'emplacement, et donc que sa valeur peut être connue avant l'attribution d'un emplacement à un publicitaire.

Generalized First Price Auction

C'est la forme d'enchères la plus naïve. Les publicitaires enchérissent pour un mot-clé, le gagnant étant celui qui fait la plus grande offre. Il paie le montant de la dernière enchère qu'il a fait pour l'emplacement.

Voici une exemple de ce type d'enchères :

Soient deux emplacement avec des CTR de respectivement 200 clics/h et 100 clics/h, et trois publicitaires Ad1, Ad2 et Ad3 avec respectivement un budget de 10\$, 4\$ et 2\$. Pour être sûr d'avoir un emplacement, il suffit à Ad2 de mettre Ad3 hors course.

première enchère : Ad2 $\rightarrow 2,01\$$

deuxième enchère : Ad1 $\rightarrow 2,02\$$

troisième enchère : Ad2 $\rightarrow 2,03\$$

...

i-ème enchère : Ad1 $\rightarrow 4,01\$$, Ad2 $\rightarrow 2,01\$$

On arrive alors à une situation d'instabilité, sans équilibre possible avec des stratégies pures. Nous allons formaliser ces notions dans la section suivante.

Generalized Second Price Auction

Ce modèle, développé par Google en 2002, vient de l'observation suivante : le publicitaire en position i ne voudra jamais payer plus que le montant de l'offre de son concurrent en position i+1. Le gagnant d'une enchère paie donc le montant de la deuxième plus grande offre.

Dans l'exemple précédent, si tous le monde est honnête, on arrive à la situation suivante : Ad1 paie 4\$ pour l'emplacement 1 et Ad2 paie 2\$ pour l'emplacement 2.

2.1.3 Un modèle pour les enchères de liens sponsorisés

Soient un ensemble d'agents $a = 1, \ldots, A$ et un ensemble d'emplacements $s = 1, \ldots, S$ avec chacun un CTR x_i tels que $x_1 > x_2 > ... > x_S$. L'utilité de l'agent a pour l'emplacement s est $u_{as} = v_a x_s$, où v_a est la valeur intrinsèque de l'agent a par rapport au mot-clé considéré. C'est un modèle simple, composé d'un seul paramètre pour comparer les agents (v_a) et également un seul paramètre pour comparer les emplacements (x_s).

On suppose toujours qu'il y a plus d'agents que d'emplacements à allouer : $A \geq S$.

Chaque agent propose une unique offre b_a, les enchères étant faites selon le modèle GSP Auction présenté précédemment. On classe ensuite les agents selon leur offre. Quitte à réordonner les agents, on notera v_s la valeur de l'agent obtenant la position s. Le prix payé par l'agent en position s est donc le montant de l'offre de l'agent en position $s + 1$: $p_s = b_{s+1}$. Le gain de l'agent s est alors $(v_s - p_s)x_s = (v_s - b_{s+1})x_s$, où on prend $x_s = 0$ pour $s > S$.

Voici un exemple pour clarifier les notations. Noter que par définition, $x_1 > x_2 > \cdots > x_S$ et que le mécanimse d'enchère assure que $b_1 > b_2 > \cdots > b_A$.

position	valeur	enchère	prix	CTR
1	v_1	b_1	$p_1 = b_2$	x_1
2	v_2	b_2	$p_2 = b_3$	x_2
3	v_3	b_3	$p_3 = b_4$	x_3
4	v_4	b_4	$p_4 = b_5$	x_4
5	v_5	b_5	0	0

2.1.4 Équilibre de Nash sur GSP

Un équilibre de Nash est un profil d'actions des joueurs tel qu'aucun joueur n'ait intérêt à dévier unilatéralement de sa stratégie.

On modélise l'enchère par un jeu simultané avec information complète. Chaque agent choisit simultanément une offre b_a. Les offres sont ensuite ordonnées et le prix que chaque agent doit payer est déterminé par l'offre de l'agent le suivant dans le classement. À l'équilibre, chaque agent préfère sa position actuelle à tout autre position. L'action d'un joueur correspond à l'offre qu'il fait dans l'enchère. Une fois ces offres faites, les positions des joueurs, les prix et donc le profit de chaque joueur sont déterminés.

Dans le tableau précédent, si l'agent 3 veut monter dans le classement, il doit enchérir au moins b_2. Dans ce cas, son prix deviendra $b_2 = p_1$ et donc son profit sera $(v_3 - p_1)x_2$. Dans la situation du tableau, le profit de l'agent 3 est $(v_3 - p_3)x_3$ donc l'agent 3 n'a pas intérêt à monter dans le classement si $(v_3 - p_3)x_3 \geq (v_3 - p_1)x_2$.

Inversement, si l'agent 2 veut descendre, il peut n'enchérir que b_4. Dans ce cas, son prix deviendra p_3 et donc son profit $(v_2 - p_3)x_3$. Donc l'agent 2 n'a pas intérêt à descendre dans le classement si $(v_2 - p_2)x_2 \geq (v_2 - p_3)x_3$.

Au vu du raisonement précédent, il est naturel de définir :

Définition 2.1.1 *Un équilibre de Nash (EN) satisfait pour tout $s \in [1, A]$:*

$$(v_s - p_s)x_s \geq (v_s - p_t)x_t, \quad t > s$$
$$(v_s - p_s)x_s \geq (v_s - p_{t-1})x_t, \quad t < s$$

avec $p_t = b_{t+1}$.

Pour l'analyse, nous allons nous restreindre à la classe suivante d'EN :

Définition 2.1.2 *Un équilibre de Nash symétrique (ENS) satisfait :*

$$(v_s - p_s)x_s \geq (v_s - p_t)x_t, \forall s, t$$

Proposition 2.1.1 *EN ⊃ ENS, ie si un ensemble d'enchères est un équilibre de Nash symétrique, alors c'est un équilibre de Nash.*

Démonstration. Comme $p_{t-1} \geq p_t$, on a $(v_s - p_s)x_s \geq (v_s - p_t)x_t \geq (v_s - p_{t-1})x_t$. □

Proposition 2.1.2 *Dans un ENS, $v_s \geq p_s$ ∀s.*

Démonstration.

$$(v_s - p_s)x_s \geq (v_{S+1} - p_{S+1})x_{S+1} = 0$$

□

Proposition 2.1.3 *Dans un ENS, $v_{s-1} \geq v_s$ ∀s, ie l'ENS est une allocation efficace.*

Démonstration.

$$\begin{aligned}
&v_t(x_t - x_s) \geq p_t x_t - p_s x_s \\
&v_s(x_s - x_t) \geq p_s x_s - p_t x_t \\
&\Rightarrow (v_t - v_s)(x_t - x_s) \geq 0
\end{aligned}$$

□

Proposition 2.1.4 *Dans un ENS, $p_{s-1}x_{s-1} > p_s x_s$ et $p_{s-1} \geq p_s$, pour tout s. De plus si $v_s > p_s$, alors $p_{s-1} > p_S$.*

Démonstration. En réarrangeant les termes de la définition d'un ENS, on a

$$p_{s-1}x_{s-1} \geq p_s x_s + v_s(x_{s-1} - x_s) > p_s x_s.$$

De plus comme $v_s \geq p_s$, on a

$$p_{s-1}x_{s-1} \geq p_s x_s + v_s(x_{s-1} - x_s) \geq p_s x_s + p_s(x_{s-1} - x_s) = p_s x_{s-1}.$$

La proposition découle du fait que si $v_s > p_s$ alors la deuxième inégalité ci-dessus est également stricte. □

Proposition 2.1.5 *Si un ensemble d'offres satisfait les inégalités de l'ENS pour $t = s + 1$ et $t = s - 1$, alors il les satisfait toutes.*

Démonstration. Exo ! □

On a donc une caractérisation des offres d'équilibre d'un ENS. Un agent en position s ne veut pas descendre (ie $(v_s - p_s)x_s \geq (v_s - p_{s+1})x_{s+1}$), et un agent en position $s + 1$ ne veut pas monter (ie $(v_{s+1} - p_{s+1})x_{s+1} \geq (v_{s+1} - p_s)x_s$). On a donc :

$$v_s(x_s - x_{s+1}) + p_{s+1}x_{s+1} \geq p_s x_s \geq v_{s+1}(x_s - x_{s+1}) + p_{s+1}x_{s+1}$$

Comme $p_s = b_{s+1}$:

$$v_{s-1}(x_{s-1} - x_s) + b_{s+1}x_s \geq b_s x_{s-1} \geq v_s(x_{s-1} - x_s) + b_{s+1}x_s$$

Donc les cas limites sont :

$$b_s^U x_{s-1} = v_{s-1}(x_{s-1} - x_s) + b_{s+1}^U x_s$$
$$b_s^L x_{s-1} = v_s(x_{s-1} - x_s) + b_{s+1}^L x_s$$

Les solutions sont :

$$b_s^U x_{s-1} = \sum_{t \geq s} v_{t-1}(x_{t-1} - x_t)$$
$$b_s^L x_{s-1} = \sum_{t \geq s} v_t(x_{t-1} - x_t)$$

car $x_s = 0 \forall s > S$.

2.1.5 Revenus générés par l'enchère

Le revenu d'une enchère est la somme des prix payés par les enchérisseurs.

Proposition 2.1.6 *Le revenu maximum sur tous les équilibres de Nash possible est le même que le revenu obtenu par l'équilibre de Nash symétrique avec enchères b_s^U.*

2.1.6 Incitation et honnêteté

Dans les paragraphes précédents, on a supposé que tous les v_i (qui sont des valeurs privées) étaient connus de tout le monde. Il peut être utile de concevoirr un modèle sans apprentissage des valeurs des adversaires, car cela ne bénéficie ni à Google, ni aux utilisateurs. Dans ce paragraphe, nous allons donc nous placer dans la situation où la valeur v_i est privée. L'utilité reste égale à $v_a - p$, ou 0 selon que l'agent obtient l'objet ou pas. On suppose toujours que le nombre d'agents est supérieur au nombre d'objets.

cas d'un seul bien

l'enchère anglaise : l'enchère commence à un prix plancher et elle augmente d'un incrément à chaque enchère. Au fur et à mesure que le prix augmente, des enchérisseurs sortent de l'enchère. Quand il ne reste plus qu'un enchérisseur, il remporte l'objet au prix atteint.

Vickrey Auction : Tous les agents donnent une enveloppe contenant leur offre. La meilleure offre remporte l'objet au prix de la deuxième meilleure offre.

Proposition 2.1.7 *Dans une VA, pour tout joueur i et tout ensemble d'enchères* $\{b_j\}_{j \neq i}$ *des autres joueurs, le joueur i maximise son utilité en enchérissant* $b_i = v_i$.

Démonstration. Si le joueur i enchérit $b'_i > v_i$ et obtient l'objet, alors son utilité est $u = v_i - B$ avec $B = \max_{j \neq i} b_j$.
- si $B < v_i$ alors l'utilité est la même que s'il avait enchéri v_i
- si $B > v_i$ alors $u < 0$.

Si le joueur enchérit $b_i < v_i$, alors :
- si $B < b_i$ alors l'utilité est la même que s'il avait enchéri v_i
- si $B > b_i$ alors $u = 0 \leq u(v_i)$.

□

On dit que VA est 'strategyproof', ou 'truthfull'. Dans ce modèle d'enchères, enchérir de manière honnête est toujours la meilleure stratégie. On note que le résultat d'une enchère anglaise est le même que celle d'une VA.

Enchères combinatoires et VCG

Soit un ensemble S contenant m biens, avec la possibilité d'enchérir sur tout ensemble $T \subseteq S$. On définit la fonction de valuation de l'agent a, $v_a \colon 2^S \rightarrow \mathbb{R}^+$ attribuant à chaque sous-ensemble de S une valeur $v_a(S)$, telle que :

$$v_a(T) \geq 0, \quad v_a(\emptyset) = 0$$
$$v_a(T) \leq v_a(T'), si T \subseteq T'.$$

Si le joueur a reçoit T et paie p, alors son utilité est $v_a(T) - p$.

Algorithme de Vickrey-Clarke-Groves (VCG)

1. Chaque joueur soumet une offre : $(b_a(T), \forall T \subseteq S)$.
2. Choisir une allocation $(T_1^* ... T_n^*)$ qui maximise $\sum_{a=1}^n b_a(T_a)$ parmi toutes les allocations possibles.
3. Chaque joueur a paie un prix p_a (à déterminer).

Proposition 2.1.8 *Si tous les joueurs sont honnêtes, VCG maximise* $\sum_{a=1}^{n} v_a(T_a)$ *sur toutes les allocations possibles.*

Pour qu'une enchère soit réaliste, on fixe le prix à payer p_a = "nuisance causée aux autres joueurs par l'agent a", i.e. :

$$p_a = \left(\max_{\{T_i\}_{i \neq a}} \sum_{i \neq a} b_i(T_i) \right) - \sum_{i \neq a} b_i(T_i^*) \geq 0$$

Proposition 2.1.9 *L'utilité d'un joueur honnête dans VCG est toujours positive.*

Proposition 2.1.10 *VCG est 'strategyproof' : pour tout joueur a, même si a connait les autres offres, il maximise son utilité en étant honnête, i.e.* $b_a(T) = v_a(T)$, $\forall T$.

Retour à GSP

Théorème 2.1.1 *L'équilibre induit par VCG est le même que celui induit par l'ENS(L) ayant revenu minimum parmi les ENS.*

On remarquera qu'il n'y a pas de monotonie entre VCG et GSP : il n'y en a pas un qui fait tout le temps gagner plus que l'autre.