Projet 1 - Chaînes de Markov

Eléments de processus stochastiques

Floriane Magera
Romain Mormont
Fabrice Servais
Troisième bachelier en sciences de l'ingénieur

0.1 Question 1.2 : téléportation

1) La formule utilisée pour calculer la matrice de transition Q_t du modèle du surfeur avec téléportation est la suivante :

$$Q_t = (1 - \alpha)Q' + \alpha\tilde{Q}$$

où Q' et $\tilde{Q}$ sont des matrices de transition et α la probabilité de téléportation.

La première est la matrice de transition du graphe initial auquel on a rajouté des arêtes partant des *dangling nodes*. Elle a été calculée en remplaçant tous les éléments de la matrice Q dans des lignes ne contenant que des 0 par $\frac{1}{n}$. Cette valeur $\frac{1}{n}$ a été choisie en considérant une densité de probabilité uniforme entre les différentes arêtes partant des *dangling nodes*.

La seconde est la matrice de transition du graphe complet formé des noeuds du graphe initial. Autrement dit, la matrice de transition représentant la téléportation. Une combinaison linéaire de paramètre α est ensuite appliquée aux deux matrices pour trouver la matrice Q_t.

2) Pour que la distribution stationnaire π_s soit unique, **il faut que la chaîne de Markov soit irréductible**. Autrement dit, il faut que pour tout couple de nœuds (i_1, i_2), il existe une arête les reliant (un probabilité non-nulle de passer de i_1 à i_2). Cette propriété est vérifiée avec le modèle du surfeur modifié puisque la téléportation permet, depuis tout noeud, de se diriger vers un autre noeud tant que $\alpha > 0$.

A partir du moment où $\alpha = 0$, on est plus assuré que chaque paire de nœuds est reliée par une arête et donc que π_s est bien stationnaire.

3) Les sites les plus visités, obtenus à l'aide de la distribution stationnaire, sont les suivants :

1. http ://purl.org/rss/1.0/modules/content
2. http ://www.ulg.ac.be
3. http ://ogp.me/ns♯
4. http ://www.gre-liege.be
5. http ://blog.intelliterwal.net
6. http ://www.jalios.com
7. http ://www.vmfnet.be
8. http ://www.alinoa.be
9. http ://www.ulb.ac.be
10. http ://www.cedia.ulg.ac.be

4)

0.2 Question 1.3 : effet de α

1) Pour prouver que le score PageRank de toute page est au moins $\frac{\alpha}{n}$ (n est le nombre de pages), on peut développer une expression "*explicite*" des éléments de la matrice Q_t en utilisant la formule donnée précédemment :

$$Q_t(i, j) = q_{ij}(1 - \alpha) + \frac{1}{n}\alpha$$

où q_{ij} est un élément de la matrice Q'. Connaissant la relation qui lie $\pi^{(k)}$ et $\pi^{(k-1)}$, on a :

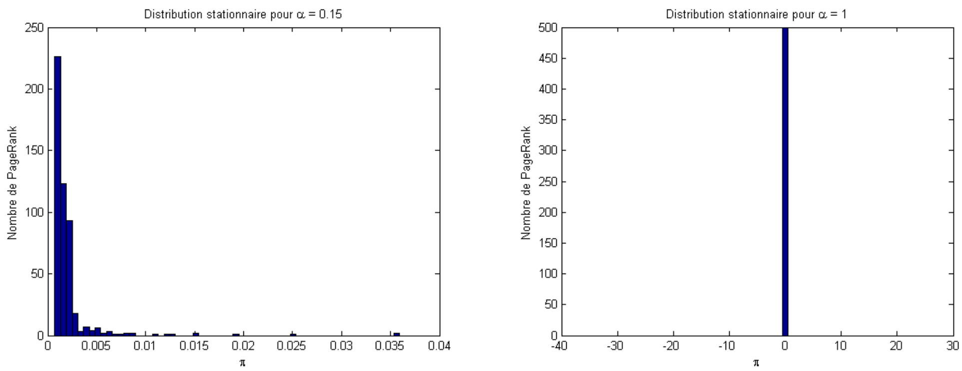
$$\begin{aligned}
\pi_j^{(k)} &= \sum_{i=1}^{n} Q_t(i,j)\pi_j^{(k-1)} \\
&= \sum_{i=1}^{n} \left(q_{ij}(1-\alpha) + \frac{\alpha}{n} \right) \pi_j^{(k-1)} \\
&= \sum_{i=1}^{n} q_{ij}(1-\alpha)\pi_j^{(k-1)} + \sum_{i=1}^{n} \frac{\alpha}{n}\pi_j^{(k-1)} \\
&= \frac{\alpha}{n} + \underbrace{(1-\alpha)\sum_{i=1}^{n} q_{ij}\pi_j^{(k-1)}}_{>0}
\end{aligned}$$

Le deuxième terme est inférieur à 1 (et même inférieur à $(1 - \frac{\alpha}{n})$ afin de respecter le deuxième axiome de Kolmogorov) et surtout, positif. De ce fait, on peut affirmer que :

$$\pi_j^{(k)} \geq \frac{\alpha}{n}$$

Le cas où α tend vers 1 correspond à la situation où le surfeur a majoritairement tendance à se téléporter lorsqu'il change de page. Si on considère qu'en cas de téléportation, la distribution de probabilité est uniforme entre les nœuds de destination, on observera un PageRank uniforme.

Afin de vérifier cette conclusion, nous avons calculé la distribution stationnaire pour un $\alpha = 1$. Le résulat est donné sur la Figure 1 où il est mis en parallèle avec la distribution des pour $\alpha = 0.15$. Le résultat est édifiant, on constate en effet un PageRank uniforme dans le cas où $\alpha = 1$ (écart-type du PageRank : 4.7753×10^{-18}).



(a) $\alpha = 0.15$ (b) $\alpha = 1$

FIGURE 1 – Distribution des PageRank lorsque α tend vers 1