Synthèse de l'article *The Small-World Phenomenon: An Algorithmic Perspective* de Jon Kleinberg

Philippe Gambette

25 Janvier 2005

1 Introduction

L'étude des grands graphes a été motivée par la demande de modélisation de divers objets : le web, le réseau internet lui-même, les neurones ou encore les interactions des protéines. La plupart des grands graphes étudiés ont une caractéristique commune : un petit diamètre. Ceci avait été remarqué par Stanley Milgram dès 1960, sur ce qu'on pourrait appeler le graphe des amis : une lettre pouvait être transmise d'ami en ami, entre deux américains quelconques ne se connaissant pas, par une chaîne de 6 amis en moyenne. Des techniques de modélisation de tels graphes à petit diamètre ont donc été trouvées, notamment par Watts et Strogatz.

Mais un autre aspect de l'expérience originale de Milgram est étudié plus particulièrement dans cet article : le fait que la transmission des lettres se fasse de façon décentralisée. En effet, les personnes de la chaîne de l'expérience de Milgram ne connaissaient du destinataire que sa ville et sa profession, or la lettre parvenait à être transmise efficacement.

Jon Kleinberg propose donc un modèle et un algorithme décentralisé pour représenter ce phénomène. Il donne un majorant du nombre de pas d'un de ces algorithmes, et un minorant pour le nombre de pas de tout algorithme décentralisé pour ce modèle.

2 Le diamètre des graphes petit-monde

Ce phénomène de petit diamètre malgré une faible densité du graphe et un grand nombre de noeuds s'explique bien, avec des modèles comme celui de la figure 1 notamment. Ce modèle a permis de représenter par exemple le graphe du web, ou le réseau d'électricité des Etats-Unis.

Le modèle présenté par Kleinberg est dérivé de celui-ci.

3 Le caractère décentralisé du trajet de la lettre et le modèle proposé

Cette question est moins classique que celle du diamètre, et semble plus étonnante : comment chaque personne de l'expérience de Milgram a-t-elle pu savoir que le choix de la connaissance à qui elle passait la lettre serait le bon, ne connaissant que la ville et le métier du destinataire ? Ils ont par exemple dû faire le choix de la donner à un *voisin proche*, ou un *voisin éloigné* géographiquement, ou socialement.

Ce comportement est modélisable par un algorithme décentralisé respectant les points suivants. Chacun des noeuds du réseau connaît :
- la structure de grille du réseau,
- la position sur la grille de la cible t,
- les positions et les voisins éloignés des noeuds qui ont reçu le message.

Cet algorithme est appliqué sur le modèle de Kleinberg, qui diffère de celui de Watts et Strogatz sur plusieurs points :
- les arêtes sont dirigées.
- la structure de grille $n \times n$ au lieu de cercle pour les voisins proches (à distance inférieure à p).
- la probabilité de relier deux voisins éloignés u et v n'est plus la même, on relie à présent chaque noeud à q voisins éloignés avec probabilité proportionnelle à $d(u, v)^r$.

4 Les intérêts du modèle

On retrouve le modèle de Watts et Strogatz pour $r = 0$. Or, dans ce cas, il n'existe pas de bon algorithme décentralisé pour transmettre le message, puisqu'il existe une constante α_0 (dépendante de p et q mais pas de n) telle que l'espérance du nombre de noeuds transmettant le message par l'algorithme est supérieure à $\alpha_0 n^{\frac{2}{3}}$.

Le modèle proposé est intéressant, puisqu'il permet, avec un algorithme simple (choisir toujours le noeud adjacent le plus proche géographiquement de la cible t) d'obtenir une espérance de nombre de pas d'au plus $\alpha_2(log(n))^2$, avec α_2 indépendante de n, $p = q = 1$, et $r = 2$.

La valeur de $r = 2$ est la plus intéressante. Ceci peut se "sentir", puisque c'est avec ces valeurs que les voisins éloignés sont ditribués de façon uniforme sur l'*échelle des surfaces* : tous les ensembles A_i contenant les noeuds situés

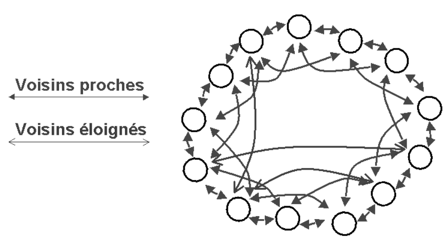

Fig. 1 – Modèle de Watts et Strogatz pour le phénomène petit-monde

entre les distances 2^i et 2^{i+1} sont équiprobables pour accueillir un voisin éloigné de tout noeud. Si $r > 2$, les A_i les plus proches sont défavorisés. Si $r < 2$, les plus éloignés le sont. Cette intuition est confirmé par un théorème qui donne une borne minimale pour l'espérance du nombre de pas de l'algorithme dans les deux cas, comme montré dans la figure 2.

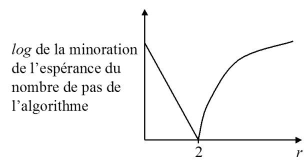

FIG. 2 – Optimalité du cas $r = 2$

5 Conclusion

Le modèle présenté a donc permis d'étudier un aspect peu commun de l'expérience de Milgram : le caractère décentralisé de la transmission du message. Malheureusement, en raison du caractère précis et peu généralisable du modèle, ceci ne peut s'appliquer directement au problème du routage à partir seulement d'informations locales.