Comment optimiser l'architecture d'un réseau pour résister aux attaques ?
Rapport final de TIPE

Matthias Goffette

June 11, 2017

Abstract

Networks need to be reliable, so that they can resist accidents and attacks. We focus on two types of networks: homogenous and scale-free. We model networks as graphs and perform agent-based simulations. On these, attacks are performed : attackers falsify the information before spreading it. The simulations show that for the same mean degree of nodes, homogenous networks are less resiliant than scale-free networks. In scale-free networks, resistance depends on sum of degrees of attackers.

1 Préambule

Mon objectif a peu changé depuis la MCOT. J'ai poursuivi l'étude de l'impact d'une attaque sur les noeuds du réseau, selon sa topologie. C'est là que réside la différence avec le travail de Dimitri Granger, l'autre membre du groupe, qui s'est intéressé à des attaques sur les liens. Je ne me suis finalement pas concentré sur la vitesse de transmission des informations. En effet, cette partie m'a semblé moins intéressante.

2 Introduction

Sécuriser les réseaux informatique est aujourd'hui un point primordial. En effet, ils sont le support de communications de plus en plus nombreuses entre les ordinateurs. Dans ce cadre, je me suis intéressé à l'effet de l'architecture d'un réseau sur sa vulnérabilité aux attaques. Pour apporter des éléments de réponse à cette question, j'ai utilisé une modélisation multi-agents en langage Python. J'ai effectué des simulations sur deux types de réseaux, *scale-free* et *homogène*. J'ai fait varier deux paramètres de la taille du réseau : le nombre de nœuds et le nombre d'arêtes par nœud.

3 Corps Principal

3.1 Modalités d'action

Il a fallu tout d'abord définir la manière dont une attaque allait opérer. Un noeud peut être soit normal, soit attaquant. Un noeud normal transmettra à ses voisins les informations qu'il reçoit sans les modifier. Mais un attaquant, avant de transmettre des informations, les faussera.

J'ai ensuite choisi la représentation des objets utiles pour la modélisation. J'ai utilisé des objets Pythons : les agents qui représentent les noeuds du réseau, les tunnels qui représentent les arêtes, et un objet qui rassemble les deux premiers, le réseau. Un dernier objet, l'information, est utilisé pour observer la propagation d'informations faussées par les attaquants. Une information peut prendre deux valeurs, vrai ou faux.

J'ai généré des réseaux *scale-free*, avec l'algorithme de Barabási–Albert. Il consiste à prendre un graphe initial, et à ajouter des nœuds. A chaque ajout d'un nœud i, on le lie à un nœud j avec une probabilité proportionnelle à le degré de j. Cela crée un réseau dans lequel la probabilité qu'un noeud ait k voisins est $\alpha k^{-\gamma}$. Ici, $\gamma = 3$ et $\alpha \approx 0.83$. Ainsi, quelques nœuds ont une forte connectivité, et une majorité en ont une faible. Pour générer des réseaux homogènes, j'ai utilisé la bibliothèque Python NetworkX.

J'ai défini le plan d'expériences et conçu l'architecture du code me permettant de le réaliser.

3.2 Restitution des résultats

Dans le cas des réseaux homogènes, j'ai considéré des réseaux ayant 10, 50 puis 100 nœuds, avec dans chaque cas trois arêtes par nœud. J'ai également fait varier le nombre d'arêtes par nœud, pour un réseau de 50 agents, avec des degrés de 3 et 20 successivement.

Pour chaque jeu de paramètres, j'ai effectué 100 simulations. Une simulation consiste en la génération d'un réseau. Pour chacun de ces réseaux, le nombre d'attaquants varie entre 0 et le nombre de nœuds n du réseau. Pour chacun de ces cas, l'algorithme a effectué une diffusion de l'information dans le réseau. Il en résulte une proportion d'informations fausses en fonction du nombre d'attaquants. La comparaison de la vulnérabilité des différents réseaux en fonction de leur architecture se fonde sur la moyenne des 100 simulations.

Sur les réseaux homogènes, j'observe que le nombre d'informations fausses dans le réseau croît avec le nombre d'attaquants (voir figure 1). La courbe est concave : plus il y a d'attaquants, moins l'action d'en ajouter un nouveau a un effet important. Cela s'explique par une redondance d'informations fausses.

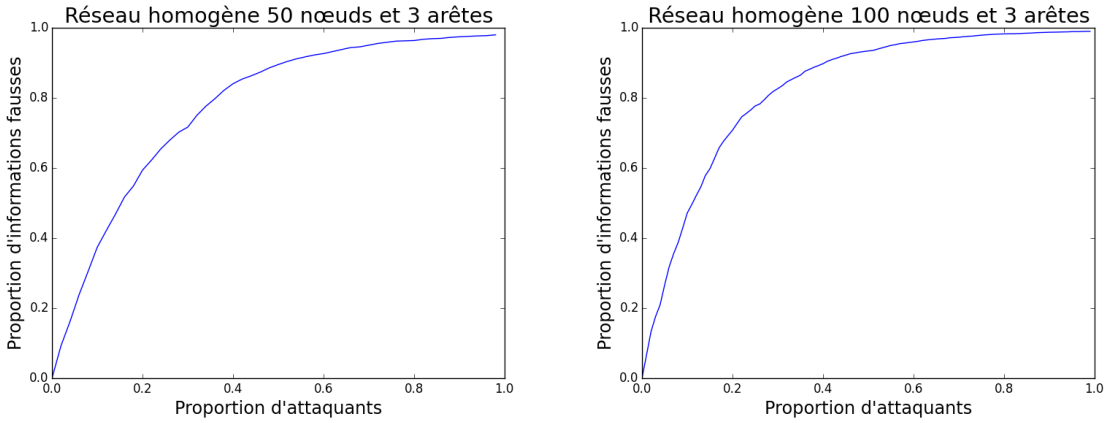

Figure 1: Réseau homogène - Variation du nombre de nœuds

Pour un nombre de nœuds constant, l'augmentation du nombre d'arêtes par nœuds se traduit, au-delà d'un certain seuil, par l'apparition d'une partie affine de la courbe, qui correspond à un stade où les seuls nœuds non attaquants sont voisins de l'émetteur (voir figure 2).

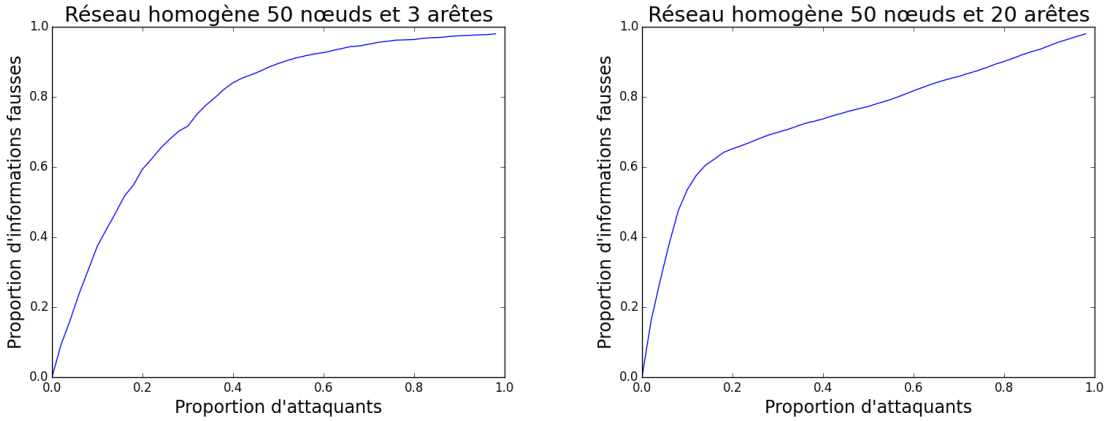

Figure 2: Réseau homogène - Variation du nombre d'arêtes par nœud

Dans le cas des réseaux invariants d'échelle, on observe en moyenne le même type de courbe concave que pour un réseau homogène. Cependant, pour une simulation donnée, la courbe présente des paliers (voir figure 3). Lorsqu'un nœud ayant une forte connectivité devient attaquant, il a un fort impact sur le reste du réseau.

Pour une même somme de degrés des nœuds attaquants, le nombre d'attaquants varie beaucoup (voir figure 4). Or, la somme des degrés des nœuds attaquants est une mesure de l'influence des attaquants. Sur certains réseaux, les nœuds ayant des degrés élevés seront contaminés assez tôt, alors que sur d'autres, ils seront contaminés au contraire assez tardivement.

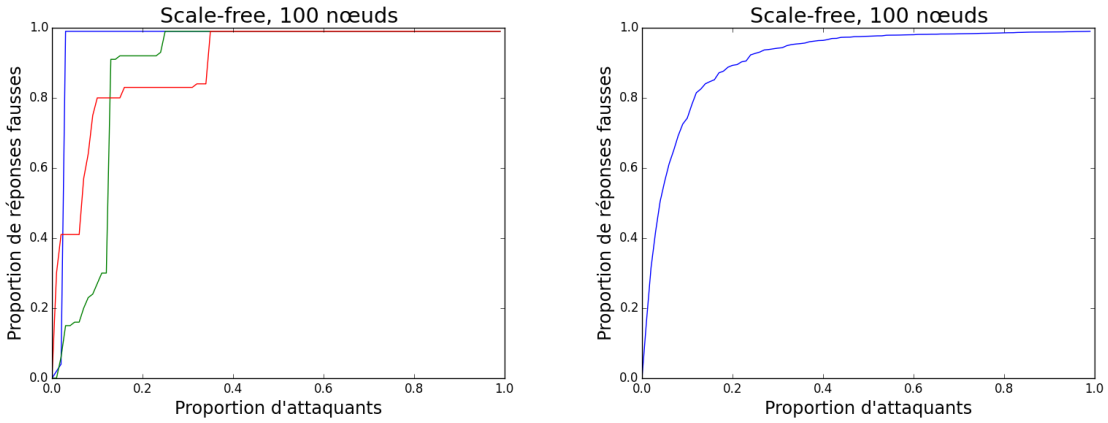

Figure 3: Réseaux scale-free - Courbes de trois simulations et moyenne sur cent simulations

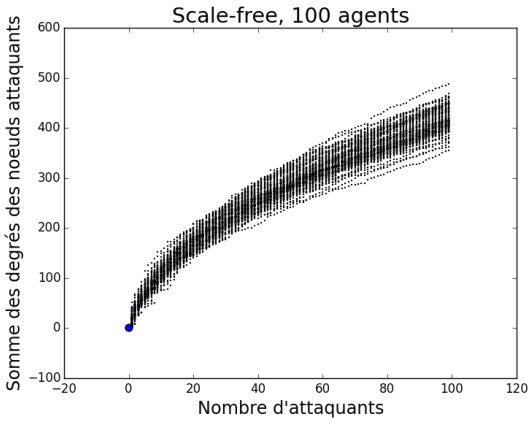

Figure 4: Somme des degrés attaquants selon le nombre d'attaquants

Pour comparer les deux types de réseaux, j'ai généré plusieurs réseaux invariants d'échelle ayant 100 nœuds et calculé le degré moyen par nœud, grâce à un script Python. Ce dernier vaut deux pour un réseau généré avec l'algorithme de Barabási–Albert. Or, on constate que les réseaux homogènes de degré deux ont en moyenne une plus grande proportion d'informations fausses (voir figure 5).

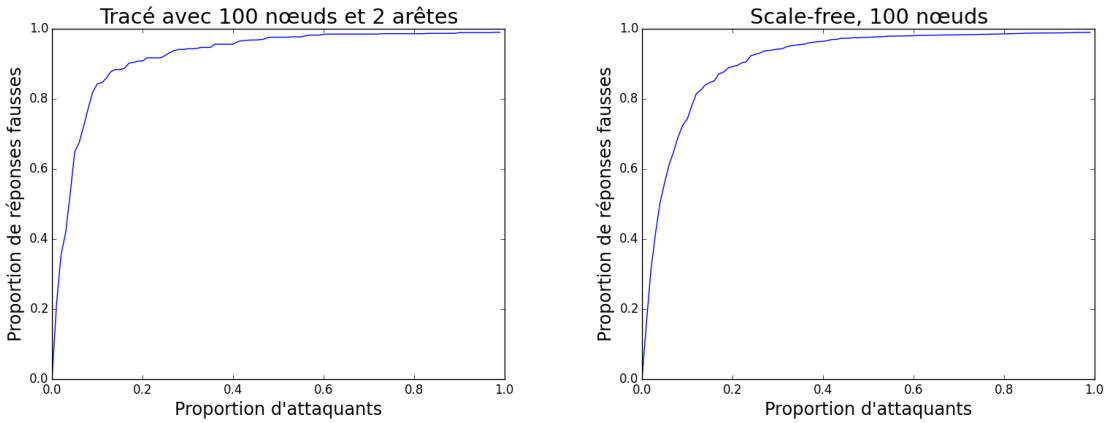

Figure 5: Comparaison des réseaux homogène et invariant d'échelle

3.3 Analyse - Exploitation - Discussion

Pour les réseaux homogènes et invariants d'échelle, l'attaque modélisée contamine le réseau très rapidement. D'après mes simulations, il suffit de 20% d'attaquants pour contaminer plus de la moitié du réseau. De plus, l'augmentation du nombre de nœuds dans un réseau homogène ne va pas diluer les informations fausses, mais au contraire renforcer l'effet des attaquants. On peut appréhender ce phénomène en considérant que les modélisations effectuées sont équivalentes à créer un arbre dont la racine est le nœud responsable de l'émission de l'information, et dont les branches sont les chemins suivis par l'information. Ajouter des nœuds revient à augmenter la taille de l'arbre. Les nœuds attaquants sont distribués dans tout l'arbre, mais plus ils sont près de la racine, plus ils ont d'influence.

Les résultats sur les réseaux homogènes peuvent être intéressants, notamment dans le cas où chaque nœud à un degré élevé. Mais ce type de réseau est peu facile à mettre en pratique. C'est donc un objet principalement théorique.

Une limite du modèle concerne la probabilité pour un nœud d'être attaquant. Je l'ai considérée comme uniforme, or en pratique, les serveurs très connectés sont de gros serveurs, généralement plus protégés que les autres. Ils devraient donc être moins souvent infectés, ce qui augmente les performance des réseaux invariants d'échelle.

4 Conclusion générale

Les réseaux homogènes semblent être moins robustes que les réseaux invariants d'échelle. Cependant, augmenter le degré moyen des réseaux homogènes permet d'accroître leur résistance aux attaques.

L'étude de l'autre membre du groupe considère des attaques sur les arêtes pour des graphes de Harary et en étoile. L'architecture optimale dépend alors du coût de protection.

Les réseaux scale-free sont des groupements de réseaux en étoile, et les réseaux de Harary sont des réseaux homogènes.