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

Matthias Goffette

June 11, 2017

## 1 Abstract

Networks need to be efficient, in terms of communication speed, and reliable, so that they can resist to accidents and attacks. We focus on two types of networks : homogenous and scale-free. We model the networks as a graph, each node representing an agent. On these, attacks are performed. The methodology is the following : an agent has an initial true information. Then, it passes it to its neighbours. If they are normal agents, they repeat the process, but if they are attackers, they falsify the information before spreading them. The simulation show that for the same mean degree of nodes, homogenous networks are more resiliant than scale-free networks.

## 2 Préambule

Mon objectif a peu changé depuis la MCOT. J'ai donc poursuivi l'étude de l'impact d'une attaque sur les noeuds du réseau, selon sa topologie. Cependant, je ne me suis finalement pas concentré sur la vitesse de transmission des informations. En effet, cette partie m'a semblé moins intéressante. J'ai préféré me concentrer sur l'étude des attaques.

## 3 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.

# 4 Corps Principal

## 4.1 Modalités d'action

Il a fallu tout d'abord définir la manière dont une attaque allait opérer. Je me suis penché sur une attaque sur les noeuds. 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. Ainsi, 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 d'attaques sur un réseau. Pour cela, 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 concentré mon étude sur les réseaux homogènes, et *scale-free*. Pour générer des réseaux *scale-free*, j'ai utilisé 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 la connectivité 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 ayant une faible. Pour générer des réseaux homogènes, j'ai utilisé la bibliothèque Python NetworkX.

Ensuite, j'ai défini les expériences que je souhaitais réaliser. Puis j'ai conçu l'architecture du code me permettant de réaliser ce plan d'expérience.

## 4.2 Restitution des résultats

Dans le cas des réseaux homogènes, j'ai fait varier le nombre de nœuds, 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, j'ai fait varier le nombre d'attaquants 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, conformément à mes attentes, 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. Pour un degré constant, la diffusion d'informations fausses est plus importante pour un réseau de grande taille.

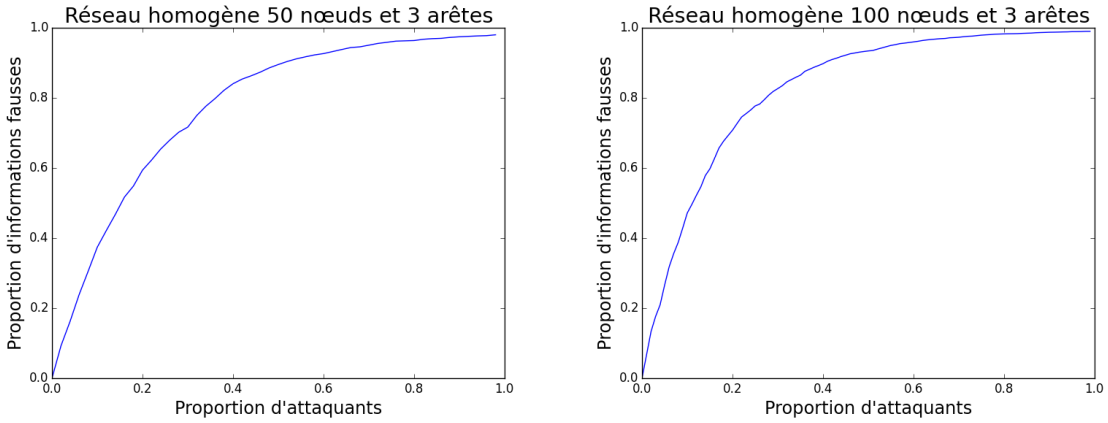


Figure 1: Réseau homogène - Variation du nombre de nœuds

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).

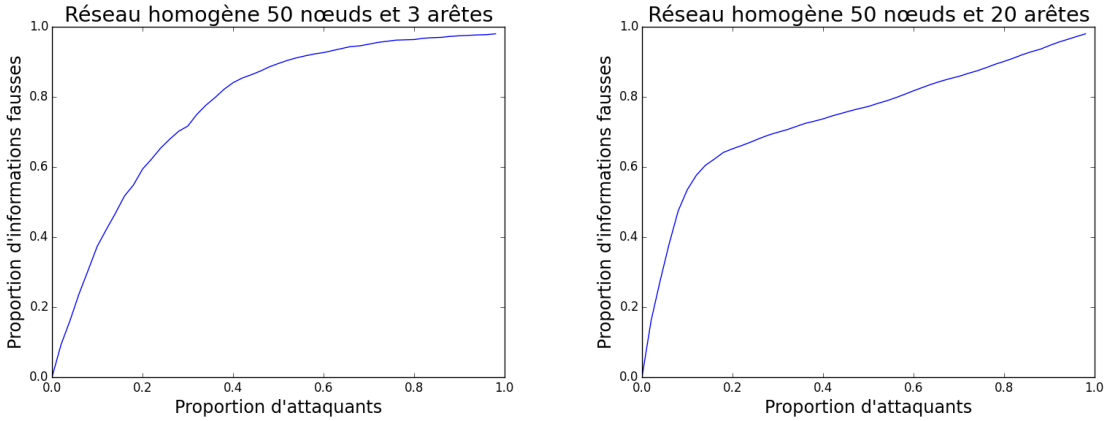


Figure 2: Réseau homogène - Variation du nombre d'arêtes par nœud

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, on voit qu'il s'agit d'une courbe à paliers (voir figure 3). En effet, lorsqu'un nœud ayant une forte connectivité devient attaquant, il a un fort impact sur le reste du réseau, d'où l'apparition de paliers de diffusion.

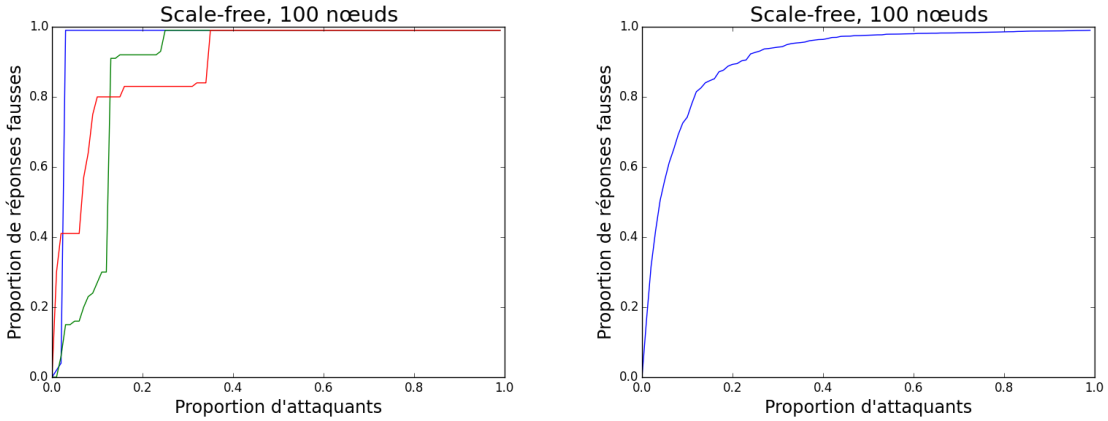

Figure 3: Réseaux scale-free - Courbes de trois simulations et moyenne sur cent simulations

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. Par conséquent, il va y avoir de grands écarts selon les réseaux. Sur certains, les nœuds ayant des degrés élevés seront contaminés assez tôt, alors que sur d'autres, ils seront contaminés au contraire assez tardivement. C'est ce qui explique les écarts.

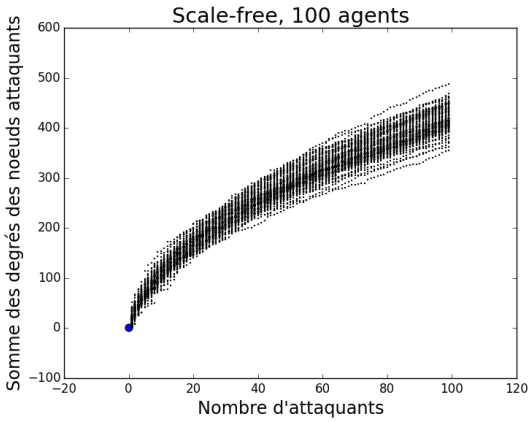


Figure 4: Somme des degrés attaquants selon le nombre d'attaquants

Pour comparer les deux types de réseaux, j'ai entrepris d'établir le degré moyen d'un nœud dans un réseau invariant d'échelle. Ce degré moyen vaut deux. Alors, pour un même degré moyen, on constate que les réseaux homogènes ont en moyenne une plus grande proportion d'informations fausses (voir figure 5).

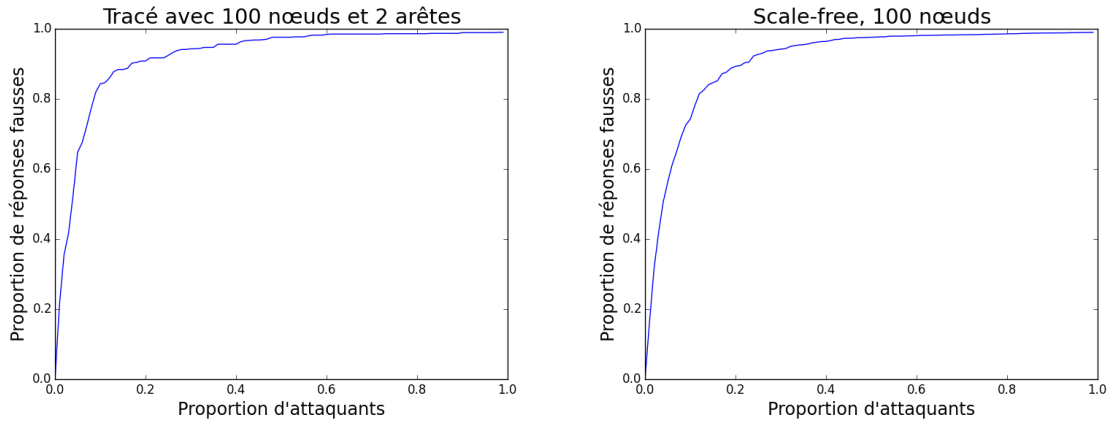


Figure 5: Comparaison des réseaux homogène et invariant d'échelle

## 4.3 Analyse - Exploitation - Discussion

On remarque que pour les réseaux homogènes et invariants d'échelle, l'attaque modélisée contamine le réseau très rapidement. En effet, d'après les graphes ci-dessus, il suffit de 20% d'attaquants pour contaminer plus de la moitié du réseau. De plus, on constate que 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 de racine le nœud responsable de l'émission de l'information, et dont les branches sont les chemins suivis par l'information. Alors, rajouter des nœuds revient à augmenter la taille de l'arbre. Les nœuds attaquants sont distribués dans tous l'arbre, mais le plus près de la racine ils sont, le plus d'influence ils auront.

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 considéré comme uniforme, or en pratique, les serveurs très connectés sont de gros serveurs, généralement plus protégé que les autres. Ils devraient donc être moins souvent infecté, ce qui augmente les performance des réseaux invariants d'échelle.

# 5 Conclusion générale

Les résultats valident les hypothèses que nous avions formulées. Mais on voit que les réseaux homogènes, bien que plus difficiles à mettre en place, semblent être moins vulnérables que les réseaux scale-free. Le modèle des réseaux *scale-free* a une importance pratique, puisque beaucoup de réseaux réels, comme Internet, prennent cette forme. Les réseaux homogènes sont plus difficiles à mettre en pratique, et donc peu utilisés. Mais il peremettent une base intéressante de comparaison.