

Analyse de Transactions Financières en Bitcoins

Matthieu Latapy

stages@complexnetworks.fr

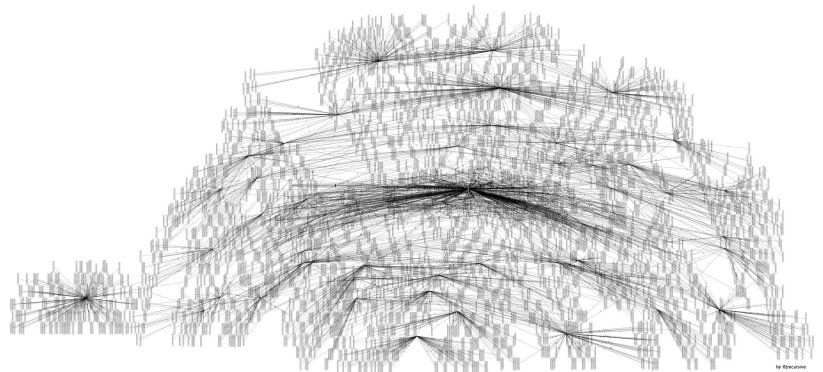
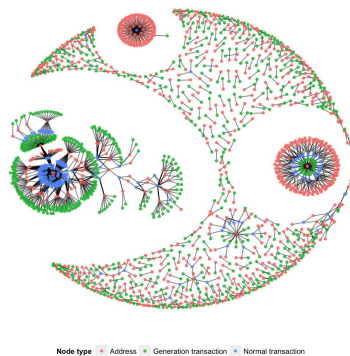
<http://complexnetworks.fr>

LIP6 – CNRS et UPMC – Paris

Bitcoin est la première et la plus répandue des monnaies numériques décentralisées. Plus de 13 millions de bitcoins sont actuellement en circulation, ce qui correspond à plusieurs milliards d'Euros (suivant les cours), et une centaine de milliers de commerçants l'acceptent dans le monde. Elle repose sur une base de données pair-à-pair et publique issue d'une technologie appelée BlockChain. Cette technologie a de nombreuses propriétés importantes, dont la transparence des transactions, l'absence d'autorité centrale de régulation, un nombre maximal de bitcoins, etc.

Comme l'ensemble des transactions Bitcoin est disponible publiquement, cette plateforme offre une **opportunité unique pour l'étude de flux financiers**. Elle contient plus de 100 millions de transactions, et s'enrichit typiquement de plus de 100 000 transactions par jour (avec une forte variabilité). Représenter cet ensemble de transactions par un graphe (comptes reliés par des transactions entre eux) est naturel et permet d'explorer diverses structures, voir la figure ci-dessous.

Unknown generation transactions and addresses since the start of 2014



Deux exemples de représentations d'ensembles de transactions bitcoins comme graphes.

Ces représentations ne capturent toutefois pas la dynamique temporelle des transactions, qui est essentielle pour comprendre l'activité sous-jacente. Nous proposons ici de représenter chaque transaction par un **flot de liens**, c'est-à-dire une suite de triplets (t, u, v) indiquant qu'à l'instant t une transaction a eu lieu entre u et v . Cette représentation a l'avantage de capturer à la fois la structure des transactions (comme les graphes) et leur dynamique temporelle.

Afin de **décrire ces flots de liens d'un point de vue à la fois structurel et temporel**, il s'agira en un premier temps de travailler à la généralisation aux flots de liens de plusieurs notions classiques sur les graphes : degrés, densité, chemins, communautés, centralités, etc. Il s'agira ensuite de calculer ces propriétés sur la masse des transactions bitcoins publiquement disponible, ce qui pose des questions d'efficacité difficiles, et d'en tirer une description fine. On s'attachera en particulier à **identifier des événements** dans cette dynamique, qu'on rapprochera d'événements connus dans la sphère Bitcoin (comme par exemple des fraudes, des détournements, ou des événements boursiers).

Ce projet utilisera des notions de programmation, de manipulation de données, de graphes, de statistiques, ... On ne s'attend pas à ce que les candidats possèdent toutes ces compétences mais ils s'y formeront au cours du projet.