



---

# Twitter et les sondages automatiques

## RAPPORT DU PROJET DAC

---

Kim-Anh Laura NGUYEN  
Arij RIABI  
Promo DAC 2018-2019

*Encadrants :* Vincent GUIGUE  
Nicolas BASKIOTIS

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Analyse de sentiments et transfert</b>	<b>3</b>
<b>3</b>	<b>Analyse de communautés</b>	<b>4</b>
<b>4</b>	<b>Fusion des deux approches</b>	<b>6</b>
<b>5</b>	<b>Conclusion</b>	<b>7</b>

# 1 Introduction

## 2 Analyse de sentiments et transfert

### 3 Analyse de communautés

Dans l'objectif d'extraire des communautés d'utilisateurs favorables à un candidat en particulier, nous considérons désormais une approche axée, non pas sur le contenu des tweets, mais sur les liens entre utilisateurs. Nous souhaitons donc construire un graphe de communautés basé sur les interactions entre utilisateurs, i.e., un graphe où chaque nœud correspond à un utilisateur et chaque arc désigne un retweet, une citation ou une réponse. Nous utilisons la bibliothèque Python `networkx` [2], spécialisée dans l'étude des graphes et des réseaux.

L'algorithme de partitionnement que nous utilisons est la méthode de Louvain [1], qui permet d'extraire les communautés de réseaux en optimisant la modularité, i.e., la densité d'arêtes à l'intérieur des communautés comparée à celle des arêtes reliant les communautés entre elles.

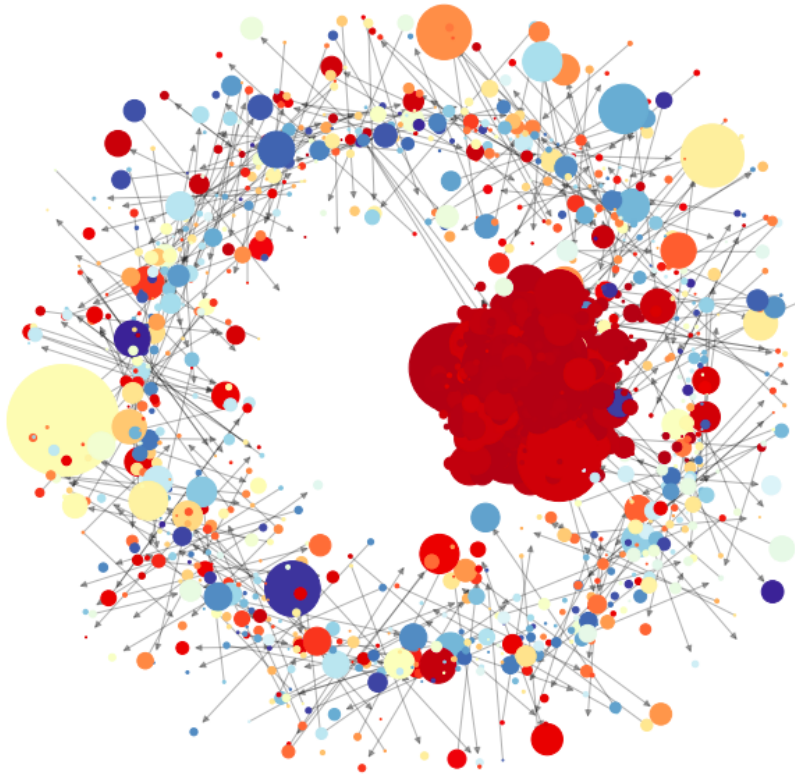


FIGURE 1 – Graphe de communautés construit sur 10000 tweets

La figure 1 contient le graphe de communautés construit à partir d'un échantillon de 10000 tweets. Cette partition contient plus de 700 communautés, et nombreuses sont celles ne contenant très peu d'utilisateurs (moins de 100) pour être significatives.

Nous décidons donc de garder uniquement les communautés de plus de 100 nœuds. La figure 2 contient les 12 plus grandes communautés de la partition.

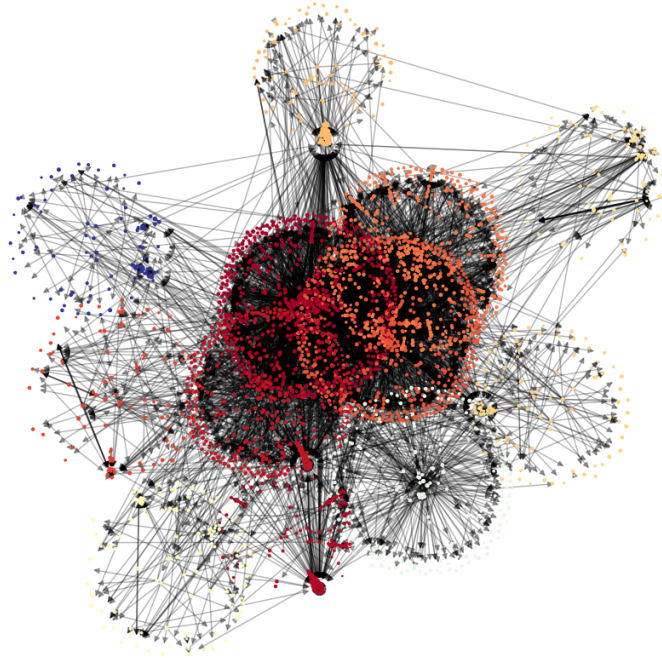


FIGURE 2 – Graphe contenant uniquement les 12 plus grandes communautés

Nous enlevons les nœuds correspondant aux utilisateurs ne faisant que retweeter - et qui sont donc possiblement des bots - ainsi que ceux ne faisant pas suffisamment partie de la communautés.

## 4 Fusion des deux approches

## 5 Conclusion

### Références

- [1] V. D. BLONDEL, J.-L. GUILLAUME, R. LAMBIOTTE et E. LEFEBVRE. “Fast unfolding of communities in large networks”. In : *Journal of statistical mechanics : theory and experiment* 2008.10 (2008), P10008.
- [2] A. HAGBERG, P. SWART et D. S CHULT. *Exploring network structure, dynamics, and function using NetworkX*. Rapp. tech. Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.