

Analyse de réseaux sociaux : vers une pensée intra-communautaire unique ?

ETUDE BASEE SUR LA CAMPAGNE PRESIDENTIELLE DE 2017

Arnaud DIDES, Mathilde BOLTENHAGEN

Tuteurs : Marc PLANTEVIT, Mehdi KAYTOUE

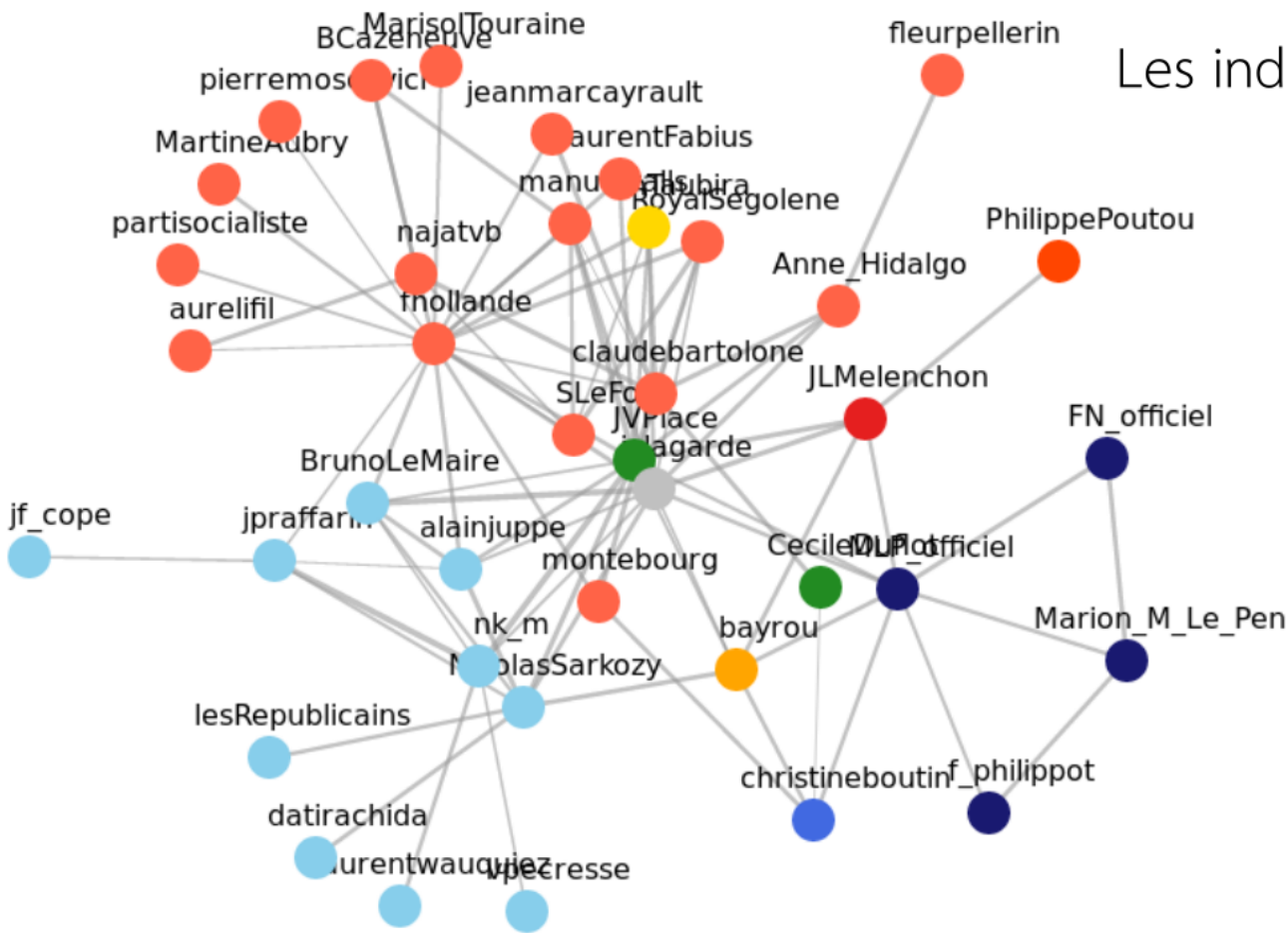
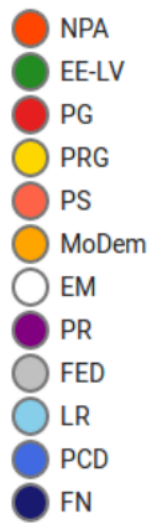
Avril 2017

Motivation :

Etudes des **comptes Twitter** des politiciens français :
- existence d'une **classification des individus** sur les réseaux sociaux
- **distinction d'idées** entre les groupes

Processus suivi :

- création d'un **dataset** [1]
- **préparation** et **analyse** des données
- **visualisations interactives**
- **interprétation** des résultats



Les individus suivent des comptes Twitter similaires :

- arête : au moins 60% de **followers communs** sur la base du compte qui en a le moins
- apparition des partis politiques : les followers adhèrent les idées de personnes d'un même parti

Résultats :

>> distinction de groupes d'individus

Exceptionnal Model Mining/Subgroup Discovery

Hashtags discriminants des cotés politiques :

	Novembre	Décembre	Janvier	1-14 Février	15-28 Février	1-15 Mars	16-31 Mars
Extreme-Gauche	#trump #presidentielle2017 #fidelcastro	#cahuzac #valls	#fillon	#justicepourtheo #fillon			
Droite	#primaireledebat #primaire2016 #laprimaireledebat	#berlin #hollande #noel	#fillonlavillette #fillon2017 #cnlesrepublicains	#macron #fillon2017 #louvre	#macron #fillon2017 #colonisation	#projetfillon #fillon2017 #trocadero	#fillonpresident #fillon2017 #londres
Gauche	#7novembre16h34 #directan #violencesfaitesauxfemmes	#primairescitoyennes #primaire #directan	#primaireledebat #primairescitoyennes #cpa	#hamon2017 #mutualite #hamon	#portugal #hamon2017 #conseilcitoyen	#hamon2017 #journedesdroitsdesfemmes #8mars	#hamon2017 #bercy #hamonbercy
Ecolos	#cop22 #nddl #eelv	#climat #pollution #eelv	#faucheursdechaises #fessenheim #enr	#climat #nucleaire #ceta	#jadot #ceta #eelv	#fukushima #8mars #turquie	#trump #climat #repondonspresent
Centre	#udi #laprimaire #trump	#presidentielle2017 #europe #hollande	#ue #trump #presidentielle2017	#presidentielle2017 #penelopegate	#bourdindirect #bayrou #sia2017	#udi #presidentielle2017 #journedelafemme	#trump #brexit #legislatives2017
Extreme-Droite	#trump #levrafillon #migrants	#noel #fillon #migrants	#trump #marine2017 #migrants	#aunomdupeuple #bobigny #assisesmip	#liban #marine2017 #mip1f1	#macron #aunomdupeuple #marine2017	#aunomdupeuple #marine2017 #debattf1

Technique :

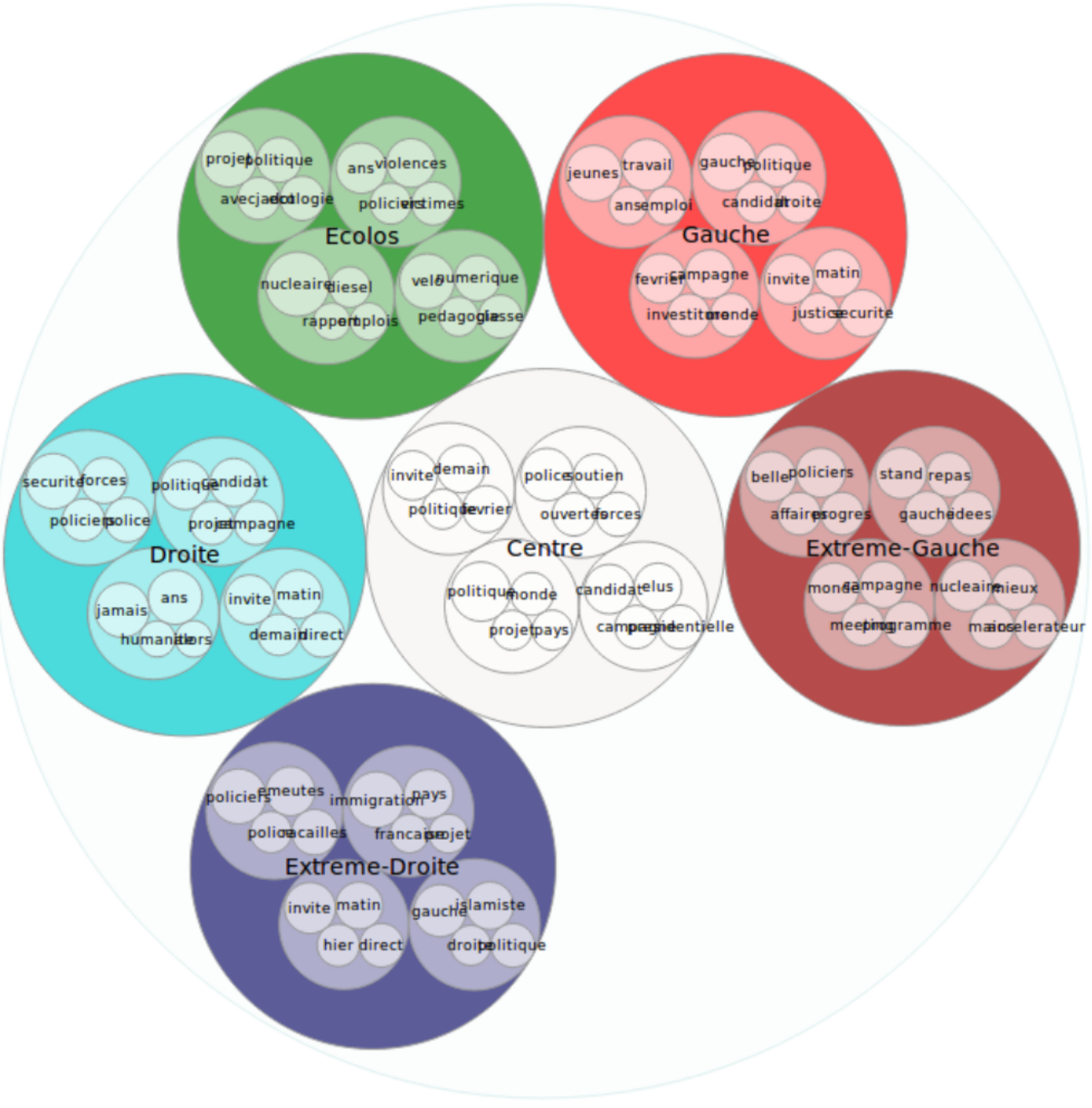
- Exceptionnal Model Mining [3] :
 - Données possédant des attributs cibles et descriptifs
 - Recherche combinatoire sur des attributs descriptifs d'un sous-groupe de données pour lequel le comportement des attributs cibles est intéressant
- Subgroup Discovery : cas particulier de l'EMM où l'on cible un seul attribut
- dans notre analyse : l'**attribut cible** est l'opinion politique, les **attributs de descriptions** sont les fréquences d'utilisation des hashtags

Résultats :

- >> découverte des hashtags qui **caractérisent les cotés politiques**
- >> **évolution** dans le temps

Thématiques des tweets - Latent Dirichlet allocation

Thèmes par côtés politiques en février 2017 :



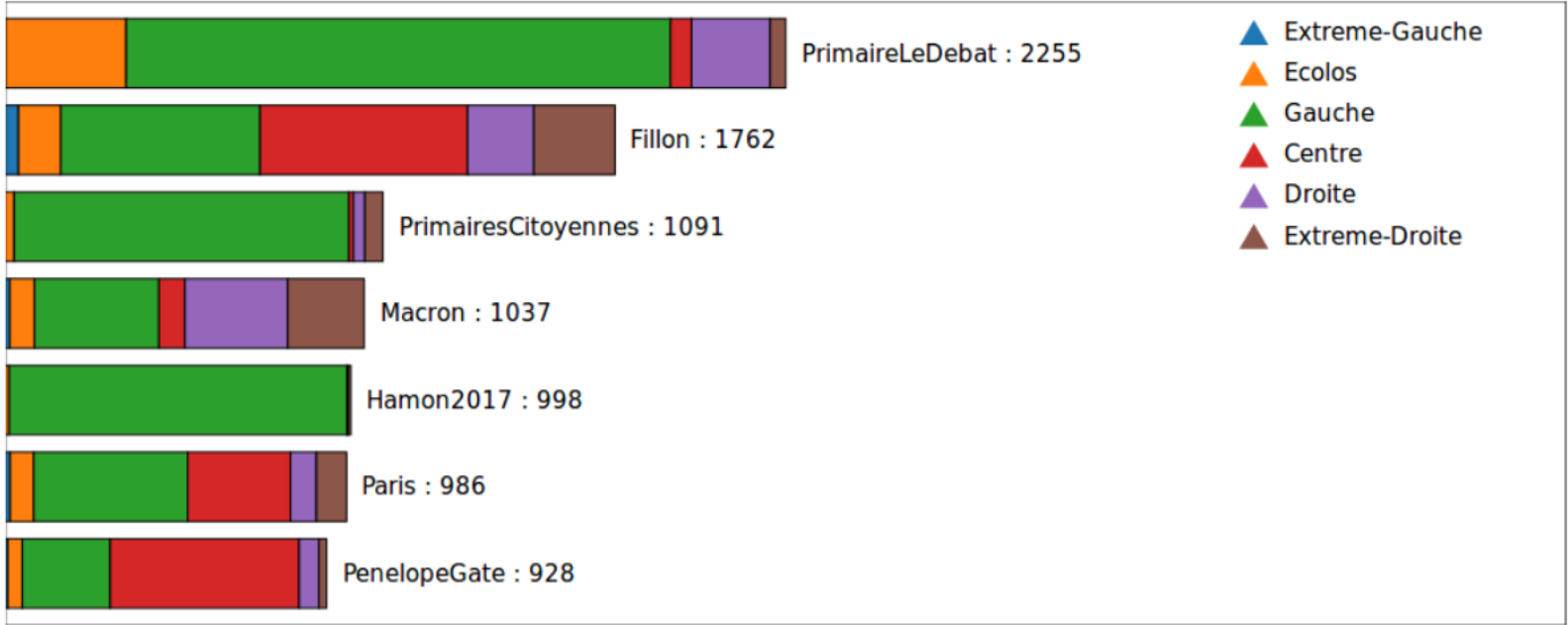
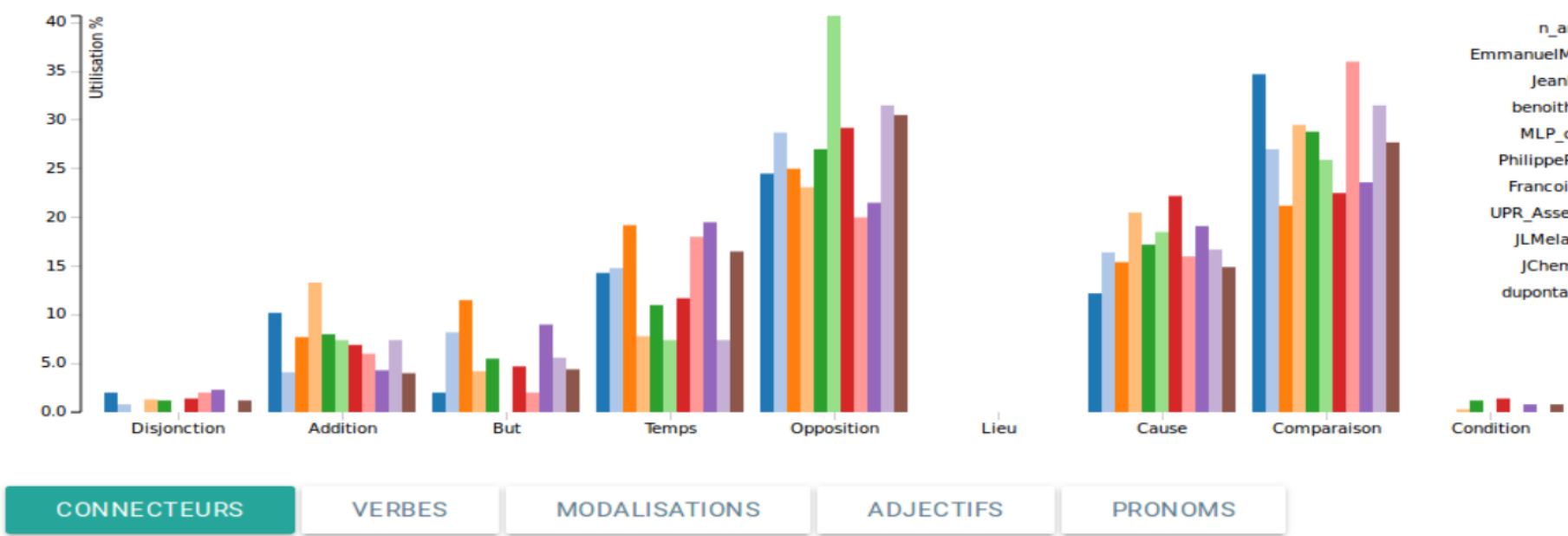
Technique :

- **modèle LDA** [2] : extraire les **sujets** d'un ensemble de textes grâce aux probabilités bayésiennes, similitudes observées attribuent des groupes aux données [1]
- études des thèmes par groupe politique

Résultats :

- >> **réurrences des thèmes** par côté politique dans le temps
- >> **différences** des thèmes selon l'opinion politique

Analyse sémantique et lexicale



- visualisation de l'utilisation des hashtags pour chaque opinion politique

Résultats :

>> **hashtags privatifs** à certains groupes politiques

Outils et références :

Création du dataset :

- API Twitter / scripts Python
- [1a] liste des 50 comptes politiciens populaires : <http://ymobactus.miaouw.net/labo-top-politiques.php?mode=followers&liste=personnalites&tendance=>
- [1b] liste des politiciens français par les décodeurs du projet ANR ContentCheck
- tweets de 385 comptes entre novembre 2016 et avril 2017

Préparation et analyse des données :

- scripts Python
- Knime / Cortana / Tropes
- [2] Explication du LDA : <http://epub.wu.ac.at/3558/1/main.pdf>
- [3] Explication du EMM : <http://hdl.handle.net/1854/LU-7044680>
- [4] inspiré de *Marine Le Pen prise aux mots, décryptage du nouveau discours frontiste* de Cécile Alduy