THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Ophélie FRAISIER

Le 7 décembre 2018

Détection de points de vue sur les médias sociaux numériques

Ecole doctorale : **EDMITT - Ecole Doctorale Mathématiques, Informatique et Télécommunications de Toulouse**

Spécialité : **Informatique et Télécommunications**

Unité de recherche :
IRIT : Institut de Recherche en Informatique de Toulouse

Thèse dirigée par
Mohand BOUGHANEM et Romaric BESANçON

Jury

M. Julien VELCIN, Rapporteur
M. Gaël DIAS, Rapporteur
M. Guillaume CABANAC, Examinateur
M. Yoann PITARCH, Examinateur
Mme Sihem AMER-YAHIA, Examinateur
M. Guy MELANÇON, Examinateur
M. Mohand BOUGHANEM, Directeur de thèse
M. Romaric BESANÇON, Co-directeur de thèse

PUBLICATIONS

PUBLICATIONS D'AUDIENCE INTERNATIONALE

▷ *Articles longs présentés en conférences avec sélection par comité de programme*

Fraisier, Ophélie, Guillaume Cabanac, Yoann Pitarch, Romaric Besancon et Mohand Boughanem (2018b). « Stance Classification through Proximity-based Community Detection ». In : *ACM Conference on Hypertext & Social Media (HT)*. Baltimore, Maryland, USA, 09/07/2018–12/07/2018 : ACM, p. 220-228. DOI : 10.1145/3209542.3209549.

Fraisier, Ophélie, Guillaume Cabanac, Yoann Pitarch, Romaric Besancon et Mohand Boughanem (2018d). « #Élysée2017fr : The 2017 French Presidential Campaign on Twitter ». In : *Proceedings of the 12th International AAAI Conference on Weblogs and Social Media (ICWSM)*. Stanford, California, USA, 25/06/2018–28/06/2018 : AAAI Press, p. 501-510. URL : https://aaai.org/ocs/index.php/ICWSM/ICWSM18/paper/view/17821.

▷ *Articles courts et posters présentés en conférences avec sélection par comité de programme*

Fraisier, Ophélie, Guillaume Cabanac, Yoann Pitarch, Romaric Besancon et Mohand Boughanem (2017c). « Uncovering Like-minded Political Communities on Twitter ». In : *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval (ICTIR)*. Amsterdam, The Netherlands, 01/10/2017–04/10/2017 : ACM, p. 261-264. DOI : 10.1145/3121050.3121091.

Fraisier, Ophélie, Guillaume Cabanac, Yoann Pitarch, Romaric Besancon et Mohand Boughanem (2018c). « The 2017 French Presidential Campaign on Twitter (poster) ». In : *EuroScience Open Forum (ESOF)*. Toulouse, France, 09/07/2018–14/07/2018. URL : https://www.irit.fr/publis/IRIS/2018_ESOF_FCPBB_poster.pdf.

▷ *Conférencière invitée dans un atelier avec sélection par comité de programme*

Fraisier, Ophélie (2018). « Political Communities on Twitter ». In : *RevOpiD-2018 Workshop on Opinion Mining, Summarization and Diversification co-located with ACM Conference on Hypertext & Social Media (HT)*. Baltimore, Maryland, USA, 09/07/2018 : ACM. URL : https://www.irit.fr/publis/IRIS/2018_HT_RevOpiD_F.pdf.

PUBLICATIONS D'AUDIENCE NATIONALE

▷ *Revue avec sélection par comité de rédaction*

FAVRE, Cecile, Chloé ARTAUD, Clément DUFFAU, Ophélie FRAISIER et Roland KOTTO KOMBI (2017). « Forum Jeunes Chercheurs à Inforsid 2016 ». In : *Ingénierie des Systèmes d'Information* 22.2, p. 121-147. DOI : 10.3166/isi.22.2.121-147. Article regroupant les quatre meilleures contributions du Forum Jeunes Chercheurs INFORSID 2016.

▷ *Article long présenté en conférences avec sélection par comité de programme*

FRAISIER, Ophélie, Guillaume CABANAC, Yoann PITARCH, Romaric BESANCON et Mohand BOUGHANEM (2017a). « Détection de points de vue à l'aide des proximités inter-profils ». In : *Conférence sur les Modèles et l'Analyse des Réseaux : Approches Mathématiques et Informatique (MARAMI)*. La Rochelle, 18/10/2017–20/10/2017, (en ligne). URL : https://www.irit.fr/publis/IRIS/2017_MARAMI_FCPBB.pdf.

▷ *Article Jeunes Chercheurs·ses et poster présenté en conférence avec sélection par comité de programme*

FRAISIER, Ophélie (2016). « Intégration du contexte spatio-temporel et social pour l'analyse de sentiments sur Twitter ». In : *INFormatique des Organisations et Systemes d'Information et de Decision (INFORSID)*. Grenoble, France, 31/05/2016–03/06/2016 : INFORSID (actes électroniques), (en ligne). URL : https://www.irit.fr/publis/IRIS/2016_INFORSID_FJC_F.pdf.

▷ *Autres communications*

BENBOUZID, Bilel, Nikos SMYRNAIOS, Julien FIGEAC, Ophélie FRAISIER, Anne L'HÔTE, Benjamin LOVELUCK, Alexis PERRIER, Pierre RATINAUD et Tristan SALORD (2017). « Twitter vs. Facebook : réseaux et discours de l'extrême droite ». In : *Participation au sein du projet Listic au sprint Datapol organisé par Sciences Po*. Paris, 29/11/17–2/12/2017. URL : https://www.irit.fr/publis/IRIS/2017_Datapol_BSFFLLPRS.pdf. Présenté par Nikos Smyrnaios et Bilel Benbouzid.

FRAISIER, Ophélie, Guillaume CABANAC, Yoann PITARCH, Romaric BESANCON et Mohand BOUGHANEM (2017b). « Expression politique sur Twitter : les communautés comme indicateurs de point de vue ». In : *6ème édition des journées « Big Data Mining and Visualization » de l'association EGC*. Lille, 26/06/2017–27/06/2017. URL : https://www.irit.fr/publis/IRIS/2017_JEEGC_FCPBB.pdf. Présenté par Ophélie Fraisier.

Fraisier, Ophélie, Guillaume Cabanac, Yoann Pitarch, Romaric Besancon et Mohand Boughanem (2018a). « Découverte de points de vue sur les médias sociaux ». In : *Séminaire « Savoirs, réseaux, médiations » du Laboratoire Interdisciplinaire Solidarités, Sociétés, Territoires (LISST)*. Toulouse, France, 13/04/2018. URL : https://www.irit.fr/publis/IRIS/2018_SRM_FCPBB.pdf. Présenté par Ophélie Fraisier.

Marchand, Pascal, Pierre Ratinaud, Julien Figeac, Guillaume Cabanac, Ophélie Fraisier, Gilles Hubert, Xavier Milliner, Yoann Pitarch, Tristan Salord, Nikos Smyrnaios et Thibaut Thonet (2018). « La campagne présidentielle 2017 sur les réseaux socionumériques ». In : *Colloque du Labex Structuration des Mondes Sociaux (SMS)*. Toulouse, 27/03/2018–29/03/2018. Présenté par Pascal Marchand et Pierre Ratinaud.

REMERCIEMENTS

I may not have gone where I intended to go,
but I think I have ended up where I needed to be.

— Douglas Adams

J'aimerais prendre ici le temps pour remercier les personnes qui m'ont aidée tout au long de cette aventure et ont toutes, à leur façon, contribué à la création de ce manuscrit.

Je tiens tout d'abord à exprimer ma reconnaissance envers mes directeurs et encadrants pour leurs conseils durant ces années de thèse : Messieurs Mohand Boughanem, Romaric Besançon, Yoann Pitarch et Guillaume Cabanac. J'ai eu la chance d'avoir quatre personnes complémentaires en termes d'intérêts et de personnalités pour me guider et vers qui me tourner en cas de besoin. Je suis également extrêmement reconnaissante d'avoir eu la liberté d'orienter mes travaux vers des problématiques qui m'étaient chères et que je n'aurais pas nécessairement pu poursuivre avec d'autres encadrants. J'aimerais remercier tout particulièrement Guillaume pour avoir été un mentor pour moi depuis bien avant le début de ma thèse. En effet, son côté mélomane nous a rapproché dès ma formation à l'IUT, et c'est également à ses cotés que j'ai découvert à la fois la recherche en informatique, la scientométrie et les études de genre lors de mon stage de L3. C'est enfin lui qui, 2 ans plus tard, m'a très professionnellement proposé cette thèse par SMS, proposition à laquelle j'avais répondu de façon toute aussi professionnelle qu'« avec toute mon affection, [il] m'emmerd[ait] ». Je le remercie donc de m'avoir emmerdé il y a 3 ans, avec toute mon affection. Je suis également très reconnaissante envers Messieurs Gaël Dias et Julien Velcin pour avoir accepté d'examiner mon travail en tant que rapporteurs, et pour leurs remarques stimulantes et enrichissantes, ainsi qu'envers Madame Sihem Amer-Yahia et Messieurs Guy Melançon et Julien Figeac pour avoir accepté de figurer dans mon jury.

Mes collègues ont bien évidemment participé à rendre plaisantes ces trois années de thèse. Tout d'abord mes collègues du CEA Tech Occitanie, que je remercie pour nos échanges et leur support, tout particulièrement Hubert Dubois qui, bien qu'il ne fasse pas parti de mes encadrants officiels, était toujours disponible pour discuter, répondre à mes interrogations et me conseiller. Ensuite, mes collègues de l'équipe IRIS, pour la bonne humeur générale de l'équipe et nos discussions de couloir. Je suis particulièrement reconnaissante envers les occupant·e·s du bureau 406, actuel·le·s et passé·e·s, qui n'hésitaient jamais à proposer leur aide. Enfin, je remercie mes collègues du projet Listic, qui m'ont

offert un bol d'air interdisciplinaire qui a profondément nourrit ma thèse, ainsi que mes collègues de déjeuner et fournisseurs officieux de café.
J'ai eu l'occasion au cours de ces années de thèse de m'impliquer dans des projets annexes, et j'aimerais donc aussi remercier les personnes que j'ai rencontré grâce à l'organisation d'ESOF ainsi que mes collègues de la commission des doctorant·e·s, deux projets pour lesquels j'ai été ravie de représenter mes camarades. Les enseignements que j'ai eu la chance de donner durant cette thèse ont aussi été extrêmement enrichissants pour moi, et je suis extrêmement reconnaissante envers toute l'équipe pédagogique et administrative de la licence GTIDM de l'IUT Informatique ainsi que les personnels de l'URFIST et de Mediad'Oc Toulouse pour m'avoir fait confiance.

Je dois bien évidemment aussi remercier ma famille et mes ami·e·s pour leur présence. Tout d'abord ma mère qui m'a toujours inspiré par sa détermination et son énergie, ainsi que sa propension à préparer des festins à chacune de nos visites. Ensuite mes beaux-parents, pour leur compréhension alors que je passais les fêtes de Noël à travailler, et mon beau-frère pour ses expérimentations culinaires. Enfin la famille de cœur que représentent mes ami·e·s : Livy et Chloë pour nos incessants babillages (je regrette de ne pas pouvoir insérer un gif de Jeff Goldblum ou de Chris Pratt), Justine pour être toujours la même après 10 ans (et avoir insisté pour que je regarde Lalaland), Antoine et Soum pour nos débats cinéma et LaTeX (je regrette de ne pas pouvoir insérer un gif de Jeff Goldblum ou de Nathan Fillion), Jérémie pour nos discussions chaotiques (et pour qu'il n'oublie pas qu'il me doit encore pas mal de bières) et toutes les autres personnes qui me sont chères et que pour la plupart je n'ai malheureusement pas assez vu ces dernières années.

Pour finir, la personne que je dois le plus remercier est l'homme qui me supporte, et me *supporte*, depuis huit ans. Il a toujours été présent pour partager mes joies, faire taire mes peurs et consoler mes peines, et je suis malheureusement incapable d'exprimer correctement ma reconnaissance en quelques lignes. Je peux par contre dire que le fait de pouvoir le présenter en tant que Clément Fraisier-Vannier le jour de ma soutenance me remplit de joie et de fierté, et que je trouve étonnamment approprié que cette année représente à la fois pour moi la fin de cette étrange et belle aventure qu'a été la thèse et le début de cette étrange et belle aventure que va très certainement être notre mariage.

Table des matières

1 INTRODUCTION 1
1.1 Contexte . . . 1
1.2 Problématique . . . 3
1.3 Contributions . . . 5
1.4 Organisation du mémoire . . . 6

I ÉTAT DE L'ART ET JEUX DE DONNÉES POUR LA DÉTECTION DE POINTS DE VUE 9

2 ÉTAT DE L'ART 11
2.1 Définitions . . . 11
2.1.1 Du réseau social au média social numérique . . . 11
2.1.2 Comment définir un point de vue ? . . . 14
2.1.3 Points de vue et médias sociaux numériques . . . 17
2.2 Détection de points de vue statiques . . . 18
2.2.1 Méthodes fondées sur le contenu textuel . . . 18
2.2.2 Méthodes fondées sur les interactions sociales . . . 24
2.2.3 Méthodes mixtes . . . 26
2.2.4 Synthèse . . . 29
2.3 Détection de points de vue dynamiques . . . 31
2.4 Jeux de données Twitter portant sur les points de vue politiques 33

3 CADRE THÉORIQUE 35
3.1 Définitions . . . 35
3.1.1 Graphe . . . 35
3.1.2 Matrice d'adjacence . . . 35
3.1.3 Degré . . . 37
3.1.4 Clique . . . 37
3.1.5 Communauté . . . 37
3.1.6 Multiplex . . . 37
3.2 Détection de communautés . . . 38
3.2.1 Communautés non recouvrantes . . . 38
3.2.2 Communautés recouvrantes . . . 40
3.3 Conclusion . . . 41

4 CONSTITUTION DU CORPUS #ÉLYSÉE2017FR 43
4.1 Présentation du contexte : la campagne présidentielle française de 2017 . . . 44

4.2 Méthodologie de collecte . . . 44
4.2.1 Présentation rapide de l'API Streaming de Twitter . . . 44
4.2.2 Sélection des mots-clés . . . 45
4.3 Sélection des profils à annoter . . . 45
4.4 Protocole d'annotation de l'idéologie politique des profils . . . 46
4.4.1 Pourquoi choisir une annotation par experts ? . . . 46
4.4.2 Annotation d'un profil . . . 47
4.4.3 Mesure de l'accord inter-annotateurs . . . 50
4.5 Présentation du dataset . . . 52
4.5.1 Statistiques globales . . . 52
4.5.2 Aspects temporels . . . 55
4.5.3 Aspects géographiques . . . 57
4.5.4 Interactions sociales . . . 60
4.6 Considérations éthiques . . . 62
4.7 Conclusion . . . 63

II MODÈLES SEMI-SUPERVISÉS DE DÉTECTION DE POINTS DE VUE SUR LES MÉDIAS SOCIAUX 65

5 FONDATIONS DES MODÈLES 67
5.1 Présentation des hypothèses . . . 67
5.1.1 (H1) Utilisation de l'homophilie sur les médias sociaux . 67
5.1.2 (H2) Importance des faisceaux d'indices multiples . . . 68
5.2 Formalisation . . . 68
5.3 Jeux de données utilisés . . . 70
5.4 Proximités . . . 71
5.4.1 Proximités fondées sur le contenu textuel . . . 72
5.4.2 Proximités fondées sur les interactions sociales . . . 72
5.4.3 Proximités fondées sur le contexte géographique . . . 75
5.5 Validation des hypothèses . . . 75
5.5.1 Structure communautaire . . . 75
5.5.2 Homogénéité des communautés . . . 76
5.5.3 Information partagée entre proximités . . . 80
5.6 Implications pour la détection de points de vue . . . 81

6 MODÈLE COMMUNAUTAIRE SÉQUENTIEL DE DÉTECTION DE POINTS DE VUE 83
6.1 Principes de base . . . 83
6.2 Fonctionnement du modèle . . . 84
6.2.1 Ordonnancement des proximités . . . 85
6.2.2 Sélection des profils-graines . . . 88
6.3 Expérimentations . . . 90
6.3.1 Influence de la fonction de sélection des profils-graines φ 90
6.3.2 Contribution de chaque proximité . . . 92
6.3.3 Influence de la fonction d'ordonnancement ω . . . 94
6.3.4 Influence du nombre de profils-graines considérés . . . 95

6.3.5 Comparaison avec les modèles de base . . . 96
6.4 Conclusion . . . 98

7 ÉVOLUTION TEMPORELLE DES POINTS DE VUE 101
7.1 Introduction au modèle . . . 101
7.2 Présentation formelle . . . 102
7.2.1 Construction des matrices de voisinage . . . 102
7.2.2 Mesure des similarités entre profils et points de vue . . . 102
7.2.3 Prédiction des points de vue des profils . . . 104
7.2.4 Gestion de l'absence d'activité . . . 105
7.3 Expérimentations . . . 107
7.3.1 Cadre expérimental . . . 107
7.3.2 Validation de l'hypothèse du voisinage . . . 108
7.3.3 Changements de points de vue . . . 112
7.4 Conclusion . . . 117

8 CONCLUSION 119
8.1 Synthèse des contributions . . . 119
8.2 Perspectives . . . 121
8.2.1 Utilisations possibles d'#Élysée2017fr . . . 121
8.2.2 Améliorations des modèles de détection de points de vue 121
8.3 Limites d'utilisation de ces modèles . . . 123
8.3.1 Polarisation . . . 123
8.3.2 Biais des données collectées sur les médias sociaux . . . 124

Annexes i

A PANORAMA DES MÉDIAS SOCIAUX iii
A.1 The Conversation Prism . . . iii
A.2 Audience des médias sociaux . . . v

B CORPUS #ÉLYSÉE2017FR xi
B.1 Structure du jeu de données . . . xi
B.2 Données détaillées . . . xii

LISTE DES FIGURES xxiii

LISTE DES TABLEAUX xxv

LISTE DES ALGORITHMES xxvii

BIBLIOGRAPHIE xxix

1 INTRODUCTION

1.1 CONTEXTE

Depuis le lancement de *SixDegrees* en 1997, les médias sociaux numériques se sont profondément ancrés dans la vie de leurs utilisateurs et utilisatrices (BOYD et ELLISON, 2007). Cela comprend bien évidemment les plateformes incontournables, telles que Facebook ou Twitter, mais également tous les sites permettant aux utilisateurs de réagir sur leur contenu. En effet, avec l'évolution des technologies liées au *Web 2.0*, de nombreux sites ont intégré des fonctionnalités sociales, quel que soit leur domaine. Peu à peu, les utilisateurs·trices ont eu la possibilité de *noter* une recette de cuisine sur Marmiton, de *partager* un livre depuis Amazon ou de *commenter* un article du Monde. Avec la croissance rapide des contenus générés par les utilisateurs·trices, de nombreux domaines ont évolué pour tenter de tirer avantage de cette surabondance de données. Le nombre de travaux traitant de médias sociaux ou réseaux sociaux a presque décuplé tous les 10 ans depuis 1950 d'après la bibliothèque numérique de l'*Association for Computing Machinery* (voir Figure 1.1).

Dans l'ensemble, les médias sociaux numériques sont devenus, ces dernières années, un matériau incontournable de recherches pour les chercheurs·ses – en informatique comme en sciences humaines, comme l'illustre la présentation de Twitter par GOLDER et MACY (2015) :

> « Twitter has emerged as the single most powerful "*socioscope*" available [...] for collecting fine-grained time-stamped records of human behavior and social interaction. »

Le marketing a tenté de créer des indicateurs pour le marché boursier (BARRELET, KUZULUGIL et BENER, 2016), de prédire le cours des actions (HASANUZZAMAN et al., 2016a ; NGUYEN, SHIRAI et VELCIN, 2015), d'estimer la valeur d'organisations caritatives (PHETHEAN, TIROPANIS et HARRIS, 2015) ou de mesurer l'impact de campagnes médiatiques (SAPRYKIN, KURCHEEVA et BAKAEV, 2016). D'autres applications se sont révélées utiles pour la cybersécurité, telles que la découverte de connections entre utilisateurs·trices en temps réel (CHEUNG, LI et SHE, 2017) ou encore les mesures de fiabilité (MEO et al., 2017). Les médias sociaux ont également été particulièrement utilisés pour étudier des élections ou référendums (BRIGADIR, GREENE et CUNNINGHAM, 2015 ; GAYO-AVELLO, 2012), des campagnes de dénigrement et de « fake news » (RATKIEWICZ et al., 2011 ; SAEZ-TRUMPER, 2014), des campagnes de harcè-

lement comme le *Gamergate*[1] (Burgess et Matamoros-Fernández, 2016 ; Trice, 2015) ou encore les propos racistes en ligne (Hasanuzzaman, Dias et Way, 2017).

De nombreuses applications mentionnées ci-dessus tirent parti de la fouille d'opinions, qui permet d'automatiquement détecter les états subjectifs partagés par les utilisateurs·trices (Liu, 2012). Une grande partie de la littérature sur ce sujet se concentre sur les critiques laissées par les utilisateurs·trices sur des plateformes numériques, par exemple des critiques de produits sur Amazon (Gindl, Weichselbraun et Scharl, 2013), de films sur IMDb (Koumpouri, Mporas et Megalooikonomou, 2015), de restaurants sur Yelp (Proios, Eirinaki et Varlamis, 2015) ou encore d'hôtels sur TripAdvisor (Yang et al., 2015). Les opinions sur lesquelles nous allons nous concentrer dans ce manuscrit sont quelque peu différentes de ce type de critiques. Nous allons travailler sur des positionnements sociaux plus complexes, que nous appellerons *points de vue* pour éviter toute confusion. Nous expliquons en détail dans le Chapitre 2 la différence que nous opérons entre *opinions* et *points de vue*, mais nous pouvons résumer ici notre définition d'un point de vue : il s'agit du positionnement social d'une personne, réfléchi et justifié par un ensemble de valeurs et de croyances. Nous considérons donc la fouille de points de vue comme un sous-domaine de la fouille d'opinions.

1. Le Gamergate est une série de campagnes de harcèlement ayant eu lieu dans le monde vidéoludique en 2014 en réaction à la diversification du domaine.

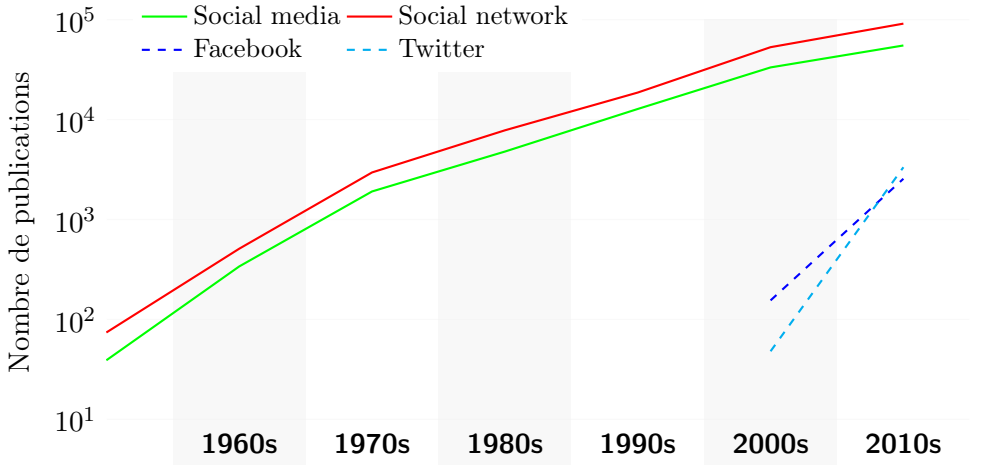


Figure 1.1 – Nombre de publications mentionnant les termes « Social media », « Social network », « Facebook » ou « Twitter » référencés dans la bibliothèque numérique de l'*Association for Computing Machinery* (ACM)

.

Dans la fouille de points de vue utilisant des modèles d'apprentissage automatique, nous pouvons distinguer deux grandes familles :

1. Les modèles supervisés : ces modèles reposent sur des données annotées. Dans un premier temps, le modèle est entraîné sur ces données d'apprentissage, avant d'être évalué ou utilisé sur les données à catégoriser. Les performances de ces modèles sont, sans surprise, directement liées à la quantité et à la qualité des données annotées fournies en entrée, selon le principe bien connu par les informaticien·ne·s du *GIGO* : Garbage In, Garbage Out [2] ;
2. Les modèles non supervisés : ces modèles tentent de directement découvrir des liens entre les données fournies, sans l'apport d'annotations. Ils regroupent habituellement les données en plusieurs catégories que l'utilisateur·trice final·e doit ensuite interpréter.

La plupart des modèles existants de fouille de points de vue sont des modèles supervisés. Ces derniers obtiennent généralement des performances supérieures aux modèles non supervisés dû à la difficulté de catégorisation. En effet, contrairement à la détection d'opinions ou de sentiments, pour laquelle nous pouvons nous aider de dictionnaires de polarité, la diversité des points de vue possibles implique que nous n'avons le plus souvent pas de lexiques ou de phrases catégorisables a priori.

Malheureusement, les jeux de données annotées de haute qualité sont une denrée rare, bien qu'ils soient essentiels pour améliorer et mesurer de manière fiable les performances des modèles. En effet, bien que la collecte des données soit assez aisée de nos jours, annoter un ensemble de données est une tâche difficile, ce qui explique pourquoi les jeux de données existants sont souvent de petite taille (Kratzke, 2017) ou se concentrent sur des situations binaires, telles que « Démocrates » / « Républicains » pour les jeux de données portant sur le paysage politique des États-Unis ou encore les positions « Non » / « Oui » lors du référendum sur l'indépendance écossaise de 2014 (Brigadir, Greene et Cunningham, 2015).

Dans ces conditions, comment pouvons-nous proposer des outils pour déterminer efficacement les points de vue d'utilisateurs·trices sur les médias sociaux numériques ?

1.2 PROBLÉMATIQUE

Comme évoqué dans la section précédente, la plupart des modèles de détection de points de vue sont supervisés et requièrent donc des jeux de données annotées qui sont compliqués à obtenir. Les modèles non supervisés présentent, eux, l'inconvénient de nécessiter une étape supplémentaire d'interprétation par l'utilisateur·trice final·e après la classification. Nous avons donc décidé de nous

2. https://en.wikipedia.org/wiki/Garbage_in,_garbage_out

orienter vers une solution intermédiaire : les modèles semi-supervisés. Ces modèles nécessitent également des données annotées en entrée, mais en quantité largement inférieure aux modèles supervisés traditionnels, sans que cela n'affecte leurs performances. La tâche que nous explorons ici est donc, en utilisant un petit échantillon de profils ayant un point de vue connu, nommés profils-graines, de catégoriser les autres profils en leur attribuant une position, par exemple « Démocrate » ou « Républicain ».

Afin de propager de façon fiable les points de vue à partir de nos profils-graines, nous allons nous appuyer sur le fait que, sur les médias sociaux, les utilisateurs·trices tendent à se rapprocher de personnes similaires, et donc souvent de personnes partageant leurs convictions. Ce phénomène se nomme l'*homophilie* et est défini formellement comme le principe selon lequel un contact entre personnes similaires a une probabilité plus élevée de se produire qu'entre des personnes peu semblables. L'impact de l'homophilie, et de l'influence sociale en général, sur la construction des opinions est connu des sociologues depuis des décennies :

> « Dans les situations où se constitue l'opinion, en particulier les situations de crise, les gens sont devant des opinions constituées, des opinions soutenues par des groupes, en sorte que choisir entre des opinions, c'est très évidemment choisir entre des groupes. » (Bourdieu, 1973)

Des études ultérieures ont montré que ces observations étaient également valables sur les réseaux sociaux numérique (Iyengar et Westwood, 2015 ; McPherson, Smith-Lovin et Cook, 2001), sur lesquels les profils pouvaient aller jusqu'à « s'enfermer » virtuellement dans des espaces totalement uniformes en matière d'idéologie, phénomène popularisé par Sunstein (2009) sous le terme de *chambres d'écho*.

Les profils sur les réseaux sociaux sont liés par diverses informations : éléments de langage, interactions sociales diverses, localisations géographiques, etc. Si nous essayons d'inférer leur point de vue sur un sujet, nous comprenons intuitivement que l'utilisation de plusieurs faisceaux d'indices peut être bénéfique : un profil peut largement partager les publications d'un candidat politique, alors qu'un autre peut être plus discret quant à ses publications mais utiliser la même rhétorique dans ses rares messages, ou peut être complètement passif mais vivre dans une région connue pour être en faveur d'un point de vue spécifique. Au vu de l'homophilie présente sur les médias sociaux, nous supposons que si plusieurs proximités différentes évoquent une forte similarité entre deux profils, il est probable qu'ils partagent un point de vue identique.

Compte-tenu du cadre de réflexion énoncé ici, nous nous efforçons dans ce manuscrit de répondre aux questions suivantes :

1. Comment définir et exploiter efficacement une proximité inter-profils ?
   - Avec quels éléments la construire ?
   - Comment mesurer sa force en matière de diffusion de points de vue ?
   - Toutes les proximités se valent-elles pour diffuser des points de vue ?
   - Comment déterminer, à partir d'une proximité, si deux profils partagent le même point de vue ?
   - Existe-t-il un intérêt à considérer plusieurs proximités ?
2. Comment gérer le nombre et la nature des profils ayant un point de vue connu au préalable ?
   - Quel est le nombre minimum de profils à utiliser pour initialiser le processus de découverte ?
   - Y a-t-il une méthode idéale de sélection de ces profils ?
   - Quelle variation dans les performances entraîne une modification de ces profils ?
3. Les utilisateurs·trices expriment-ils·elles plusieurs points de vue sur les médias sociaux ?
   - Quelle est la proportion de profils évoluant à la lisière de plusieurs points de vue distincts ?
   - Au vu de la forte homophilie présente sur les médias sociaux, les profils peuvent-ils ouvertement changer d'avis ?

Nos contributions pour répondre à ces questions sont présentées dans la section suivante.

1.3 CONTRIBUTIONS

Comme indiqué dans la Section 1.1 et la Section 1.2, l'une des grandes difficultés de la fouille de points de vue est l'accès à des données annotées de qualité. Nos contributions au domaine de la détection de points de vue sur les médias sociaux tentent de pallier cette limite de plusieurs façons.

Tout d'abord, pour notre première contribution, nous proposons un nouveau jeu de données Twitter permettant d'évaluer et d'améliorer les modèles de détection de points de vue. Ce jeu de données a été conçu pour proposer un cas d'étude plus proche des situations complexes de la vie réelle que les collections existantes. Il porte sur la campagne présidentielle française de 2017 et comporte cinq points de vue politiques (correspondant aux cinq partis politiques principaux de la campagne) ainsi qu'un point de vue *Indéterminé*. Nous offrons également certaines informations supplémentaires pour mieux caractériser les profils. Il s'agit, à notre connaissance, du premier jeu de données de cette envergure (plus de 20 000 profils) ainsi que du premier jeu de données proposant des

points de vue multiples pour un même profil permettant de vraiment prendre en compte les profils oscillant entre deux partis.

Notre seconde contribution regroupe plusieurs modèles de détection de points de vue semi-supervisés génériques, exploitant l'homophilie à travers diverses proximités (par exemple textuelles, sociales ou encore géographiques). Ces modèles sont adaptables à n'importe quel medium à condition qu'il existe des éléments permettant de relier les profils. Notre premier modèle est un modèle communautaire séquentiel de détection de points de vue (Sequential Community-based Stance Detection model ou `SCSD`). Il s'agit d'un modèle conçu pour réduire les coûts d'annotations. Il fonctionne donc avec un très petit nombre de profils-graines et repose sur la complémentarité entre liens forts et liens faibles pour propager itérativement les points de vue de communauté en communauté via différentes proximités. Nos évaluations, réalisées sur cinq jeux de données provenant de deux plateformes différentes, confirment l'importance de l'influence communautaire sur le point de vue des profils sur les médias sociaux. Nous obtenons en effet une efficacité supérieure aux modèles de référence, avec un score F1 allant jusqu'à 95 % avec seulement 1 % de données annotées. Nous offrons également une comparaison de plusieurs proximités permettant d'étudier lesquelles sont les plus utiles pour détecter le point de vue. Pour notre second modèle nous avons souhaité nous concentrer sur l'aspect dynamique des points de vue et observer leur évolution au cours d'un événement. Nous nous appuyons pour cela sur des similarités fondées sur le voisinage des profils. Ce procédé nous permet d'obtenir des résultats pertinents avec seulement un profil-graine par point de vue.

1.4 ORGANISATION DU MÉMOIRE

Ce mémoire est organisé en deux parties : la première présente le cadre de ce travail alors que la seconde se concentre sur nos modèles semi-supervisés de détection de points de vue.

Plus spécifiquement, la Partie I regroupe les chapitres 2 à 4. Le Chapitre 2 pose les définitions des concepts de base de ce manuscrit et présente les différents types de modèles de détection de points de vue sur les médias sociaux existants, ainsi que les jeux de données employés dans ces travaux, avec un accent sur les caractéristiques principales des jeux de données Twitter portant sur des sujets politiques. Le Chapitre 3 introduit le cadre théorique nécessaire pour une lecture aisée du reste du manuscrit. Il introduit les notions de base de la théorie des graphes et les principales familles de détection de communautés. Notre première contribution est présentée dans le Chapitre 4 : il s'agit du jeu de données original #Élysée2017fr, évoqué dans la Section 1.3. Il décrit le contexte de celui-ci, détaille les processus de collecte et d'annotation des données, puis offre quelques analyses préliminaires permettant de dresser un panorama du jeu de données et des possibilités qu'il offre.

La Partie II inclut les chapitres 5 à 7 contenant nos contributions aux modèles de détection des points de vue. Le Chapitre 5 présente les bases théoriques de nos modèles semi-supervisés de détection de points de vue et les expérimentations réalisées pour valider nos deux hypothèses principales : l'homogénéité des communautés détectées en termes de points de vue et la complémentarité des différentes proximités présentes sur les médias sociaux. Le premier modèle construit sur ces hypothèses, le modèle `SCSD`, est introduit dans le Chapitre 6. Nous présentons les principes de son fonctionnement, détaillons les modules le constituant puis analysons ses résultats dans diverses conditions. Notre second modèle, destiné à l'observation de l'évolution des points de vue au cours d'un événement, est décrit dans le Chapitre 7, accompagné des expérimentations permettant de valider ses performances.

Enfin, nous concluons ce manuscrit dans le Chapitre 8 par une synthèse de nos contributions, une présentation des perspectives futures de ces travaux, mais également une section revenant sur les limites d'utilisation des modèles conçus pour les médias sociaux.

Première partie

ÉTAT DE L'ART ET JEUX DE DONNÉES POUR LA DÉTECTION DE POINTS DE VUE

2

ÉTAT DE L'ART

2.1 DÉFINITIONS

Nous présentons dans cette section les définitions incontournables pour notre travail, à savoir les notions de *média social* et de *point de vue*. Nous justifions également la validité de leur utilisation conjointe.

2.1.1 *Du réseau social au média social numérique*

L'analyse des réseaux sociaux est une discipline ancienne, puisant notamment ses racines dans la *sociométrie* de Moreno (1935), qui considère chaque individu comme un *atome social*, source d'un réseau de relations. Ses *sociogrammes* sont parmi les premières « représentation[s] graphique[s] des interrelations qui unissent les membres et les groupes d'une même collectivité » et offrent une « géographie psychologique » d'un groupe social (Moreno et Maucorps, 1955, p. 238). Avec l'apparition et la démocratisation d'internet, de nombreuses personnes ont tiré parti de l'analyse des réseaux sociaux traditionnels pour analyser les nouveaux types de connexions sociales développés grâce aux technologies numériques. boyd et Ellison (2007) définissent un réseau social numérique (RSN) comme un service web permettant à des individus de (a) construire un profil au moins partiellement public, (b) gérer une liste d'autres utilisateurs·trices avec lesquels ils partagent des connexions (dont la nature et la nomenclature varient en fonction du site) et (c) voir les connexions faites par d'autres dans le système.

Cependant, cette définition paraît trop restrictive pour un *média social numérique*. En effet, comme ils le notent eux-mêmes, la démocratisation du *Web 2.0* – notamment caractérisé par les contenus générés par les utilisateurs·trices – a entraîné une « socialisation » des sites. Nombre d'entre eux ont intégré, en plus de leur contenu principal, des fonctionnalités annexes permettant d'interagir avec le contenu et les autres utilisateurs·trices : abonnements, commentaires, réactions (« J'aime », « Favori », ...), etc. Nous allons donc opter pour la définition plus générique de l'Office québécois de la langue française [1] :

> Média numérique basé sur les technologies du *Web 2.0*, qui vise à faciliter la création et le partage de contenu généré par les utilisateurs, la collaboration et l'interaction sociale.

1. http://www.granddictionnaire.com/ficheOqlf.aspx?Id_Fiche=26502881

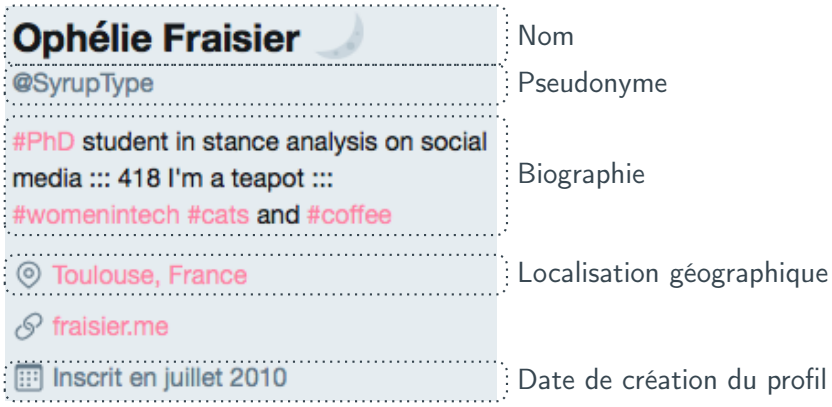


FIGURE 2.1 – Exemple d'un profil Twitter.

Cette définition permet d'englober les RSN et les sites web dont la fonction principale n'est pas la construction d'un réseau mais qui permettent néanmoins les interactions sociales. Elle est également proche de la définition employée par certains économistes tels KAPLAN et HAENLEIN (2010), qui définissent les médias sociaux numériques comme « un groupe d'applications en ligne qui se fondent sur l'idéologie et la technologie du *Web 2.0* et permettent la création et l'échange du contenu généré par les utilisateurs ». Par simplicité, nous utiliserons par la suite la formule *médias sociaux* pour désigner les *médias sociaux numériques*. La Figure A.1.1 présentée en Annexe A (p. iii) offre un aperçu de la multitude et de la variété des médias sociaux existants.

2.1.1.1 *Pourquoi parler de « profil » ?*

Nous allons privilégier dans le reste du manuscrit l'appellation « *profil* » à « *utilisateur* » ou « *utilisatrice* ». En effet, il est important de se souvenir qu'il n'existe pas toujours un lien bijectif entre personne et profil et qu'il est souvent impossible pour nous de savoir comment sont administrés les profils étudiés. Certains profils sont gérés par des groupes d'utilisateurs·trices qui peuvent se relayer ou avoir accès au profil en parallèle mais se charger de différents aspects de celui-ci (par exemple des profils de groupes de militant·e·s ou de membres d'associations). À l'inverse, certaines personnes vont avoir plusieurs profils sur la même plateforme pour pouvoir garder séparés des intérêts ou des mondes sociaux (l'exemple type est l'utilisation d'un profil professionnel et d'un profil personnel).

2.1.1.2 *Présentation de la plateforme Twitter*

Twitter est aujourd'hui devenu un matériau de recherche incontournable, autant pour les chercheurs en informatique que pour les chercheurs en sciences sociales. Afin de faciliter la lecture des sections suivantes, qui présentent une revue de l'état de l'art dont de nombreux modèles sont établis sur cette plateforme, nous présentons brièvement ici son fonctionnement. Twitter est une

FIGURE 2.2 – Exemple d'un extrait de fil Twitter.

plateforme de microblogs, permettant à ses utilisateurs·trices de partager de courtes publications appelées *« tweets »* via leur *profil*. La longueur maximale d'un tweet est passée de 140 caractères au lancement de la plateforme, le 21 mars 2006, à 280 caractères le 7 novembre 2017[2]. Un tweet peut contenir des mots-clés appelés *hashtags*, débutant avec le caractère « # ». Chaque profil est caractérisé par un identifiant alphanumérique, que nous appellerons *pseudonyme*, unique sur la plateforme et reconnaissable grâce au préfixe « @ ». Un profil peut également contenir : (a) un *nom*, qui n'est pas soumis aux limitations des pseudonymes, (b) un court paragraphe permettant d'indiquer des informations biographiques ou des intérêts, appelé la *biographie*, et (c) une localisation géographique, librement renseignée par l'utilisateur·trice. La Figure 2.1 présente un exemple de profil Twitter contenant ces éléments. Après s'être connecté à Twitter, un profil accède à son *fil*, c'est-à-dire aux tweets

2. Il est techniquement possible de dépasser la limite des 280 caractères étant donné que Twitter ne décompte pas le nombre de caractères utilisés lors de l'insertion de certains contenus.

postés par ses *amis*. Les amis sont les profils auxquels le profil considéré est abonné[3] (respectivement, les *abonnés* d'un profil sont les profils suivant les tweets du-dit profil). Il est possible d'agir sur les tweets des autres profils de plusieurs façons : (a) un *« retweet »* permet de diffuser le tweet d'un profil auprès de ses propres abonnés, c'est-à-dire de faire apparaître ce tweet dans leur fil, (b) un profil peut être *mentionné* dans un tweet grâce à son pseudonyme. La Figure 2.2 illustre les concepts présentés jusqu'alors.

2.1.2 *Comment définir un point de vue ?*

À l'heure actuelle, la terminologie du domaine de la fouille d'opinions et de sentiments, dont fait partie la fouille de points de vue, est encore loin d'être stabilisée. Un même terme peut donc recouvrir plusieurs notions proches mais néanmoins différentes. Nous présentons ici la définition de *point de vue* que nous employons dans le reste du manuscrit, ainsi que les principaux termes utilisés dans la littérature pour parler d'états subjectifs, leurs similitudes et leurs différences, afin que les lecteurs·trices puissent se familiariser avec ceux-ci.

2.1.2.1 *Le* point de vue *dans ce manuscrit*

Le *point de vue* est une notion couramment utilisée mais, de ce fait, rarement définie formellement. Afin d'éviter toute confusion, nous allons donc prendre le temps dans cette section d'expliciter ce que nous entendons par ce terme. Le point de vue peut être présenté comme un type d'opinion particulier. En effet, comme le souligne Thonet (2017, p. 10), l'une des définitions courante d'opinion est :

> Ensemble des idées d'un groupe social sur les problèmes politiques, économiques, moraux, etc. *[Exemple] : L'opinion française.*[4]

Cette définition est également fréquemment utilisée dans les travaux de sciences humaines, notamment les études et critiques de l'*opinion publique* (Allport, 1937 ; Bourdieu, 1973 ; Gayo-Avello, 2015). Un point de vue traduit donc le positionnement social d'une personne, un positionnement réfléchi, justifié par un ensemble de valeurs et de croyances, et à mettre en relation avec les autres points de vue existant sur le sujet donné. Dans ce manuscrit, nous considérons qu'un *point de vue* est un positionnement complexe construit à partir d'*opinions* de plus bas niveau. La fouille de points de vue consiste donc à détecter ce positionnement en s'aidant des divers faisceaux d'indices disponibles.

3. La notion d'*ami·e* implique normalement une idée de réciprocité dans la relation, mais nous employons ici la terminologie officielle de Twitter : https://web.archive.org/web/20180814053739/https://developer.twitter.com/en/docs/accounts-and-users/follow-search-get-users/overview.

4. https://www.larousse.fr/dictionnaires/francais/opinion/56197

2.1.2.2 *Autres états subjectifs dans la littérature*

La fouille de points de vue, telle que nous l'avons définie dans la section précédente, est parfois aussi appelée « *fouille d'opinions* » dans la littérature, mais nous préférons nous éloigner de cette appellation car elle est également utilisée pour étudier des objets bien moins complexes. Dans ces situations, les opinions se définissent par un sentiment exprimé sur une échelle de valeurs allant de positif à négatif[5] à l'encontre d'un aspect particulier d'une entité (Liu, 2012). L'exemple classique de ce type d'opinions est la critique de biens ou de services sur des plateformes telles qu'Amazon ou Yelp, sur lesquelles les utilisateurs·trices peuvent noter leurs expériences ou produits (Sungsri et Ua-apisitwong, 2017 ; Vu et al., 2016). Il n'y a dans ce cas là aucune réflexion particulière liée au positionnement social de l'auteur·e de la critique. Les approches supervisées sont les plus répandues pour cette tâche, tout particulièrement avec le regain en popularité ces dernières années des plongements lexicaux. Ces derniers sont utilisés pour tenter de mieux appréhender la sémantique des opinions et donc détecter avec plus d'exactitude les sentiments des auteur·e·s (Gunes, 2016 ; Zhi et al., 2017).

Dû à la présence de sentiments dans les opinions telles que définies par Liu, cette tâche est parfois également appelée « *analyse de sentiments* » (Giachanou et Crestani, 2016 ; Liu, 2012 ; Musto et al., 2017). Cependant, ce terme est lui aussi ambigu car il peut parfois désigner la simple détection de tonalités positives ou négatives dans les textes sans cible spécifique (parfois aussi appelée « *analyse de polarité* ») (Blaz et Becker, 2016 ; Ohbe, Ozono et Shintani, 2017 ; Sluban et al., 2014), ou être synonyme d'« *analyse d'émotions* », c'est-à-dire la détection des six sentiments fondamentaux tels que défini par Ekman et al. (1987) : la colère, le dégoût, la peur, la joie, la tristesse et la surprise (Gkontzis et al., 2017 ; Tang et al., 2018). Ces tâches s'appuient également souvent sur des modèles supervisés ainsi que sur des lexiques de mots auxquels sont associés respectivement une polarité ou une émotion en fonction de la tâche. Abdaoui et al. (2017) présente un état de l'art des ressources lexicales utilisées pour ces tâches, en français (Tableau 1) et en anglais (Tableau 2), dont les classiques *General Inquirer* (Stone et al., 1966) et *WordNet Affect* (Strapparava et Mihalcea, 2008), qui offrent des annotations de polarités et de sentiments pour certains termes anglosaxons. Nous pouvons également citer *SentiWordNet* (Baccianella, Esuli et Sebastiani, 2010), qui associe à environ 100 000 mots anglais des scores de polarité. Le Tableau 2.1 résume les chevauchements lexicaux entre les termes présentés ici.

5. Dans de nombreux cas, cette échelle est simplement binaire (positif / négatif) ou ternaire (positif / neutre / négatif).

FOUILLE DE POINTS DE VUE	FOUILLE D'OPINIONS	ANALYSE DE SENTIMENTS	ANALYSE D'ÉMOTIONS
Détection de positionnements sociaux bâtis sur un ensemble de valeurs et de croyances (ex : idéologies politiques).	Détection de sentiments positifs ou négatifs, souvent exprimés envers une cible précise (ex : critiques de livres ou de films).	Détection des six sentiments fondamentaux d'Ekman et al. (1987) : la colère, le dégoût, la peur, la joie, la tristesse et la surprise.	

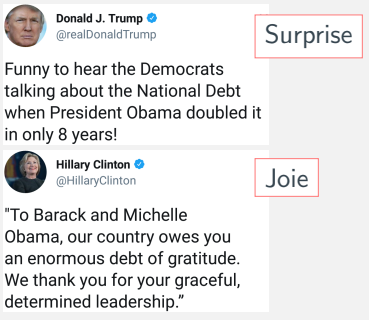


TABLEAU 2.1 – Recouvrements lexicaux entre les termes « *fouille de points de vue* », « *fouille d'opinions* », « *analyse de sentiments* » et « *analyse d'émotions* ».

2.1.3 *Points de vue et médias sociaux numériques*

Nous avons défini précédemment les notions de *point de vue* et de *média social* mais nous n'avons pas justifié leur utilisation conjointe. D'un point de vue pratique, cette utilisation s'explique simplement par le foisonnement de points de vue aisément accessibles sur les médias sociaux. Néanmoins, les données collectées sur les médias sociaux ont des limites connues : leur qualité est difficilement mesurable et elles manquent généralement de profondeur en matière d'argumentation (GAYO-AVELLO, 2011 ; MADLBERGER et ALMANSOUR, 2014). Il est donc légitime de se poser la question de la *pertinence* de l'utilisation de ces données et des points de vue détectés sur ce type de plateforme. Or, si nous les comparons aux caractéristiques données par ALLPORT (1937, p. 13) dans son travail reconnu sur l'opinion publique, les similarités sont non négligeables :

— nous retrouvons sur les médias sociaux des verbalisations produites par de nombreuses personnes, portant sur des sujets ancrés dans la culture commune et importants pour leurs auteur·e·s, représentant souvent une forme d'approbation ou de désapprobation de l'objet commun discuté ;

— les utilisateurs·trices de médias sociaux sont conscient·e·s que d'autres peuvent réagir à propos de ces sujets et de leurs publications sans être nécessairement en contact direct avec ces personnes ;

— ils / elles savent également que leur comportement peut leur permettre d'atteindre un but (partager le plus largement possible une information, convaincre quelqu'un, etc.) ;

— ces comportements suscitent souvent des conflits interpersonnels lorsque les utilisateurs·trices ne partagent pas le même point de vue.

Malgré leurs limites, les médias sociaux semblent donc être des vecteurs d'information intéressants pour étudier les points de vue, leur diffusion et leur évolution.

2.2 DÉTECTION DE POINTS DE VUE STATIQUES

Nous allons présenter dans cette section les travaux de détection des points de vue sur les médias sociaux. Le Tableau 2.2 synthétise les plateformes utilisées par les modèles introduits ici et permet de constater que Twitter est largement majoritaire pour cette tâche. En effet, bien que certains travaux utilisent d'autres médias sociaux grand public, tels que Facebook ou Youtube, ou des plateformes moins connues, la grande majorité des modèles sont évalués exclusivement ou en partie sur Twitter. Le Tableau 2.3 se concentre, lui, sur les thèmes servant à l'évaluation de ces modèles, que nous avons répartis en cinq catégories principales : points de vue politiques globaux (par exemple « Gauche » vs « Droite »), points de vue sur un thème politique particulier (par exemple concernant le changement climatique), points de vue religieux, autres points de vue et points de vue multiples (cette dernière catégorie regroupant simplement les deux travaux couvrant plusieurs thèmes éloignés). Comme le suggère déjà notre catégorisation, les points de vue politiques sont très largement prédominants. Ce tableau présente également la localisation géographique des thèmes et permet de constater que ceux ne concernant pas les États-Unis sont anecdotiques.

Nous nous concentrerons ici sur la détection de points de vue statiques, c'est-à-dire des points de vue n'évoluant pas dans le temps. Nous présenterons les modèles de détection en trois temps, en fonction des informations utilisées par ceux-ci. Nous nous attarderons tout d'abord sur les méthodes fondées exclusivement sur le contenu textuel, puis sur les méthodes se distinguant par leur usage des interactions sociales présentes sur les médias sociaux et, pour finir, sur les méthodes mixtes exploitant à la fois le texte et l'aspect social.

2.2.1 *Méthodes fondées sur le contenu textuel*

Le contenu textuel a longtemps été la seule source d'information disponible et de nombreux modèles ont été développés pour détecter les points de vue en s'appuyant exclusivement sur celui-ci. Nous pouvons par exemple citer le modèle thématique de Thonet et al. (2016), différenciant mots thématiques et mots d'opinion afin d'identifier la tendance idéologique des auteur·e·s en fonction de leurs choix lexicaux. En revanche, sur les médias sociaux, le contenu textuel est rarement utilisé seul, principalement à cause de l'importance des interactions sociales sur ces plateformes, mais également car le langage utilisé sur les médias sociaux évolue extrêmement rapidement. Eisenstein (2013) illustre notamment le fait que la proportion de non-concordance lexicale d'un système de traitement automatique de la langue augmente avec le temps, le rendant rapidement obsolète s'il n'est pas mis à jour pour prendre en compte les nouvelles formes apparues sur la plateforme. Certains travaux sont d'ailleurs consacrés à la découverte et l'actualisation du vocabulaire utilisé pour discuter de certains sujets (Mahendiran et al., 2014).

TABLEAU 2.2 – Médias sociaux utilisés pour évaluer les modèles de détection de points de vue.

PLATEFORME	TRAVAUX
Twitter [a]	BARBERÁ (2015), BOIREAU (2014), CERON (2017), CHEREPNALKOSKI et MOZETIC (2015), CONOVER et al. (2011b), DAVID et al. (2016), FANG et al. (2015), GUERRERO-SOLÉ (2017), MAGDY et al. (2016), MAKAZHANOV, RAFIEI et WAQAR (2014), MOHAMMAD, SOBHANI et KIRITCHENKO (2017), PENNACCHIOTTI et POPESCU (2011), RABELO, PRUDENCIO et BARROS (2012), RAJADESINGAN et LIU (2014), RIZOS, PAPADOPOULOS et KOMPATSIARIS (2017), THONET et al. (2017), VOLKOVA, BACHRACH et DURME (2016), WONG et al. (2013) et ZUBIAGA et al. (2018)
Facebook [b]	ABBASI et al. (2014), CERON (2017)
CreateDebate [c]	HASAN et NG (2013) et TRABELSI et ZAIANE (2018)
Youtube [d]	RIZOS, PAPADOPOULOS et KOMPATSIARIS (2017)
Flickr [e]	RIZOS, PAPADOPOULOS et KOMPATSIARIS (2017)
CNN News [f]	DONG et al. (2017)
Forum	DONG et al. (2017) [g], TRABELSI et ZAIANE (2018) [g], ZHANG et al. (2017) [h]
Blog	CERON (2017)
Données synthétiques	AKOGLU (2014)

a. Réseau social de microblogage, https://twitter.com
b. Réseau social, https://www.facebook.com
c. Plateforme de débats, http://www.createdebate.com
d. Plateforme de partage de vidéos, https://www.youtube.com
e. Plateforme de partage de photographies, https://www.flickr.com
f. Site web de la chaîne de télévision d'information en continu américaine CNN, https://edition.cnn.com
g. Forum consacré aux débats politiques, http://www.4forums.com/political/
h. Forum consacré au cancer du sein, http://www.breastcancer.org

TABLEAU 2.3 – Thèmes utilisés pour évaluer les modèles de détection de points de vue.
Le modèle d'AKOGLU (2014) ayant été évalué sur des données synthétiques, il ne figure pas dans ce tableau.

		RÉGION GÉOGRAPHIQUE		
	CATÉGORIE	États-Unis	Europe	Autre
Points de vue politiques globaux	Partis politiques	BARBERÁ et RIVERO (2015), PENNACCHIOTTI et POPESCU (2011) et RABELO, PRUDENCIO et BARROS (2012)	BARBERÁ et RIVERO (2015), BOIREAU (2014), CERON (2017) et CHEREPNALKOSKI et MOZETIC (2015)	MAKAZHANOV, RAFIEI et WAQAR (2014)
	Conservateurs vs Libéraux	ABBASI et al. (2014) et WONG et al. (2013, 2016)		VOLKOVA, BACHRACH et DURME (2016)
	Gauche vs Droite	CONOVER et al. (2011a)		DAVID et al. (2016)
	Séparatistes vs Unionistes		FANG et al. (2015) et GUERRERO-SOLÉ (2017)	
Points de vue sur un thème politique particulier	Figure politique	HASAN et NG (2013), MOHAMMAD, SOBHANI et KIRITCHENKO (2017) et TRABELSI et ZAIANE (2018)		
	Avortement	DONG et al. (2017), HASAN et NG (2013), MOHAMMAD, SOBHANI et KIRITCHENKO (2017) et TRABELSI et ZAIANE (2018)		
	Changement climatique	MOHAMMAD, SOBHANI et KIRITCHENKO (2017)		
	Santé	HASAN et NG (2013) et ZHANG et al. (2017)		
	Féminisme	MOHAMMAD, SOBHANI et KIRITCHENKO (2017)		
	Droits LGBT	DONG et al. (2017), HASAN et NG (2013) et TRABELSI et ZAIANE (2018)		
	Conflit israëlo-palestinien	DONG et al. (2017)		
	Contrôle des armes à feu	DONG et al. (2017), RAJADESINGAN et LIU (2014) et TRABELSI et ZAIANE (2018)		
	Immigration	DONG et al. (2017) et MAGDY et al. (2016)		
Points de vue religieux		MOHAMMAD, SOBHANI et KIRITCHENKO (2017)		VOLKOVA, BACHRACH et DURME (2016)
Autres points de vue		DONG et al. (2017) et PENNACCHIOTTI et POPESCU (2011)		
Points de vue multiples				RIZOS, PAPADOPOULOS et KOMPATSIARIS (2017) et ZUBIAGA et al. (2018)

Néanmoins, il existe des modèles de détection de points de vue qui restent concentrés sur ce type d'information, par choix ou en raison de contraintes liées à la tâche considérée. Ceci est particulièrement vrai pour les travaux s'intéressant à cette tâche dans le cadre de plateformes sur lesquelles les interactions sociales sont limitées ou difficiles à extraire, telles que les sites de débats ou les forums. Nous présenterons donc ici des travaux fondés exclusivement sur le texte. Nous commencerons par décrire les modèles tentant de prédire les points de vue des profils en modélisant les spécificités de leur discours, supposant que celles-ci varient en fonction du point de vue, puis ceux s'appuyant sur l'aspect séquentiel d'une discussion pour déduire les points de vue en fonction de l'enchaînement des arguments et, enfin, les modèles exploitant les opinions et sentiments des profils sur des problématiques de plus bas niveau, qui tentent donc d'agréger ces éléments pour déterminer les points de vue correspondants. Nous présenterons également certains modèles dépendant d'une tâche légèrement différente de la nôtre mais néanmoins proche, à savoir la prédiction d'élections.

2.2.1.1 *Modélisation du discours*

Ces modèles sont construits selon une hypothèse de spécificité du discours en fonction du point de vue de l'émetteur : les profils tendent à discuter de façons différentes de thèmes spécifiques qui peuvent refléter leur positionnement politique. Afin de capter ces spécificités, Fang et al. (2015) ont proposé un système de classification fondé sur un classifieur naïf bayésien (en anglais *Naive Bayes* ou NB) enrichi d'un modèle thématique non supervisé sous-jacent. En l'occurrence, ils ont montré que lors du référendum sur l'indépendance écossaise, les profils ne discutaient pas de la même façon des différentes problématiques selon qu'ils étaient unionistes ou séparatistes mais parlaient de thèmes différents ou utilisaient des termes différents pour parler de thèmes semblables. Boireau (2014) a lui choisi de modéliser le discours présent dans les publications Twitter des parlementaires belges à l'aide de l'algorithme *Wordfish* (Lo, Proksch et Slapin, 2016 ; Slapin et Proksch, 2008) – représentant le discours à l'aide d'une distribution de Poisson – afin de les positionner sur un axe idéologique. Ses observations prouvaient que l'opposition classique gauche-droite du paysage politique belge pouvait être retrouvée sur Twitter. *Wordfish* a également été utilisé par Ceron (2017) pour l'évaluation de positions politiques intra-parti à partir de profils Twitter et Facebook et de billets de blogs. Il a montré que ce modèle pouvait être utile pour distinguer plusieurs factions appartenant au même mouvement politique. David et al. (2016) ont pour leur part tenté de déterminer le point de vue de profils Facebook en utilisant une classification par corpus croisés. Ils se sont pour cela appuyés sur le discours présent sur les pages Facebook officielles des partis, modélisé par un séparateur à vaste marge (SVM) entraîné sur les unigrammes et bigrammes présents. La classification par corpus croisés désigne le fait d'entraîner un classifieur sur un corpus et de l'utiliser pour catégoriser les données d'un autre corpus. Il s'agit d'une méthode utile pour diminuer la quantité d'annotations nécessaires au bon fonctionne-

ment du modèle, mais elle ne donne pas toujours de bons résultats. Dans ce cas précis, David et al. ont montré que, malgré une baisse de performance, leur modèle restait efficace sur des profils d'utilisateurs·trices classiques, avec une exactitude (*accuracy*) de 82 %.

2.2.1.2 *Aspect séquentiel de la discussion*

Certains travaux ont tenté de prendre en compte l'aspect séquentiel d'une discussion se déroulant sur les médias sociaux pour améliorer la détection des points de vue. Hasan et Ng (2013) ont pour cela utilisé plusieurs modèles supervisés pour catégoriser des publications postées sur le site de débat Create-Debate, sur les thèmes de l'avortement, du cannabis, d'Obama et des droits des personnes LGBT[6], notamment des modèles de Markov cachés (*Hidden Markov model* ou HMM en anglais) et des champs aléatoires conditionnels (ou CRF pour *Conditional Random Fields*). Leurs résultats suggéraient l'intérêt de cette démarche étant donné que les modèles HMM et CRF obtenaient de meilleurs scores que les modèles SVM et NB utilisant les même caractéristiques, à savoir les n-grammes, les statistiques du document, la ponctuation et les dépendances syntaxiques. Ils ont tout de même noté le fait que le meilleur moyen pour eux d'améliorer les scores consistait simplement à augmenter la quantité d'annotations en entrée. De façon similaire, Zhang et al. (2017) ont tenté d'identifier les points de vue des profils d'un forum consacré au cancer du sein sur le sujet des médecines alternatives. Ils ont pour cela utilisé un réseau de neurones convolutif (*Convolutional Neural Network* ou CNN), ainsi qu'une régression linéaire, tous deux fondés sur des caractéristiques lexicales, la ponctuation, les mots exprimant des sentiments positifs ou négatifs et la présence de noms d'autres profils ou de mots spécifiques sélectionnés par Zhang et al. Sur ce type de plateforme, en revanche, les résultats ne montraient pas de différence significative entre le CNN et la régression linéaire.

2.2.1.3 *Inférences fondées sur des opinions et des sentiments*

Comme présenté dans la Section 2.1.2, un point de vue peut être représenté comme le résultat d'une réflexion fondée sur un ensemble d'opinions de plus bas niveaux. Sur cette base, Akoglu (2014) a tenté de déterminer le point de vue politique global de profils en fonction de leurs opinions sur des problématiques spécifiques. Le problème était modélisé sous forme d'un graphe bipartite – c'est-à-dire un graphe contenant des nœuds de deux natures différentes : les profils et les sujets – reliés par des liens positifs (resp. négatifs) si le profil considéré avait une bonne (resp. mauvaise) opinion du sujet. Une fois le graphe initial créé, un algorithme de propagation de labels permettait de prédire les valeurs des opinions inconnues, et les points de vue globaux étaient déterminés pour chaque profil par vraisemblance maximale vis-à-vis des opinions détec-

6. Lesbiennes, Gays, Bis et Trans.

tées précédemment. MOHAMMAD, SOBHANI et KIRITCHENKO (2017) se sont eux concentrés sur une tâche légèrement différente : détecter avec un tweet unique le point de vue de son auteur·e vis-à-vis de diverses problématiques politiques : l'athéisme, la légalisation de l'avortement, le mouvement féministe, Hillary Clinton et l'existence du réchauffement climatique. Ils ont eu recours pour cela à un SVM reposant sur diverses caractéristiques lexicales classiques extraites des tweets – n-grammes, ponctuation et présence de majuscules – ainsi que sur la présence de mots ou d'émoticônes exprimant des sentiments positifs ou négatifs. Les résultats de leur modèle soulignaient la difficulté de cette tâche ainsi que le fait que bien que les sentiments puissent être utiles, ils sont loin d'être suffisants pour obtenir un résultat convenable.

2.2.1.4 *Prédiction d'élections*

Bien qu'il s'agisse d'une tâche annexe à la prédiction de points de vue à proprement parler, nous nous devons d'aborder également dans cet état de l'art les modèles de prédiction d'élections. Dans ce cas de figure, le but n'est pas de prédire avec exactitude le point de vue de chaque profil mais les résultats obtenus par les différents partis lors des élections considérées. Ces modèles peuvent être divisés en trois catégories en fonction de leur méthode de prédiction.

VOLUME DE TWEETS PUBLIÉS. Ces modèles reposent sur l'hypothèse que plus un parti est populaire dans les urnes, plus le nombre de tweets le concernant sera important (BERMINGHAM et SMEATON, 2011 ; JUNGHERR, JÜRGENS et SCHOEN, 2012 ; METAXAS, MUSTAFARAJ et GAYO-AVELLO, 2011 ; SANG et BOS, 2012 ; SKORIC et al., 2012 ; TUMASJAN et al., 2010).

ANALYSE DE SENTIMENTS. Ces modèles reposent cette fois-ci sur l'hypothèse qu'un parti populaire en matière de votes génèrera un grand nombre de tweets *positifs*, et tentent donc de détecter la polarité des tweets avant d'effectuer une prédiction (BERMINGHAM et SMEATON, 2011 ; GAYO-AVELLO, 2011 ; METAXAS, MUSTAFARAJ et GAYO-AVELLO, 2011 ; O'CONNOR et al., 2010 ; RAZZAQ, QAMAR et HAFIZ SYED MUHAMMAD BILAL, 2014 ; SANG et BOS, 2012).

AUTRE PROCÉDÉ. Une minorité de modèle a tenté d'appliquer des modèles un peu plus complexes, tels que des régressions linéaires (LIVNE et al., 2011) ou des modèles thématiques (CASTRO, KUFFO et VACA, 2017 ; CASTRO et VACA, 2017).

Cependant, malgré la variété des élections couvertes, la majorité des modèles proposés obtiennent des prédictions incorrectes ou ne donnent pas de meilleurs résultats que des sondages traditionnels. De plus, les résultats sont extrêmement sensibles aux variations dans la collecte des données et prennent rarement en compte les biais liés à l'utilisation de Twitter : il y a de nombreuses différences entre les profils Twitter et les électeurs·trices étudié·e·s, notamment pour ce qui est de la démographie et de l'éligibilité à voter (GAYO-AVELLO, 2013 ; JUNGHERR, JÜRGENS et SCHOEN, 2012).

2.2.2 *Méthodes fondées sur les interactions sociales*

Avec l'essor des médias sociaux, de nombreux modèles ont tenté d'exploiter la nouvelle catégorie d'information que représentaient les interactions sociales entre profils, en partant du principe que l'entourage social d'un profil permettait de le caractériser. Ces travaux présentent également l'avantage de ne pas être sensible à l'ironie, très fréquemment utilisée sur les médias sociaux, et qui peut grandement handicaper les modèles fondés sur le contenu textuel. L'utilisation des interactions sociales a été de plus justifiée par les travaux portant sur l'homophilie et les *chambres d'écho* sur les médias sociaux, particulièrement pour les points de vue en politique.

Le phénomène d'homophilie désigne le fait que la probabilité que deux personnes soient en contact est plus élevée si celles-ci sont similaires. Il existe deux sortes d'homophilie : l'homophilie de statut – établie sur le statut formel, informel ou déduit, et l'homophilie de valeur – établie sur les valeurs, les croyances et les comportements (McPherson, Smith-Lovin et Cook, 2001). Sur les médias sociaux numériques il peut être compliqué de connaître avec certitude le statut des membres donc nous pouvons supposer que l'homophilie de valeur est plus présente que l'homophilie de statut. Ce phénomène d'homophilie peut donner naissance à des *chambres d'écho* (ou *chambres de résonance*), concept popularisé par Sunstein (2009) qui les définit comme des espaces de discussion en ligne où les profils ne s'exposent qu'à des idées renforçant leurs convictions initiales. La présence des chambres d'échos a notamment été observée sur Twitter par plusieurs travaux analysant la polarisation politique des profils. L'analyse la plus connue est celle de Conover et al. (2011a) qui ont démontré que, lors de la campagne pour les élections de mi-mandat 2010 aux États-Unis, les retweets étaient une interaction hautement partisane alors que le réseau des mentions était beaucoup moins divisé. Cette observation a été confirmée plusieurs fois par la suite (Barberá et al., 2015 ; Merry, 2016 ; Smith et al., 2013), y compris sur des réseaux ne portant pas sur le paysage politique américain (Weber, Garimella et Batayneh, 2013). Cet aspect partisan des retweets a été confirmé par les profils eux-mêmes dans les travaux de boyd, Golder et Lotan (2010), dans lesquels ils citent comme raisons de retweeter le fait de publiquement montrer son accord avec le profil retweeté, ou encore le fait de valider les pensées de l'auteur·e du tweet retweeté.

Nous présentons en trois temps les travaux tentant de prédire les points de vue des profils fondés sur les interactions sociales de ces derniers. Les premiers modèles présentés propagent les points de vue le long des liens inter-profils. Les modèles utilisant des communautés pour déterminer les points de vue sont ensuite introduits, suivis par ceux exploitant des représentations alternatives de la similarité inter-profils.

2.2.2.1 *Propagation de l'information*

Certains modèles tentent de propager les points de vue en fonction des liens existant entre profils. Par exemple, partant des observations énoncées ci-dessus, Conover et al. (2011b) ont utilisé un algorithme classique de propagation de labels sur le réseau des retweets pour déterminer les points de vue des profils. Sur leur jeu de données, composé de tweets discutant de la vie politique des États-Unis, ils ont observé que cette méthode était plus efficace qu'un SVM entraîné sur le contenu des tweets ou qu'un SVM entraîné sur les hashtags présents, ces modèles obtenant respectivement 80 % et 91 % d'exactitude (*accuracy*) contre une exactitude de 95 % pour la méthode fondée exclusivement sur les retweets. De la même façon, Abbasi et al. (2014) se sont appuyés sur les liens d'abonnements pour déterminer le point de vue politique de profils Facebook, en analysant plus précisément si la prédiction du point de vue était plus efficace en considérant les abonnements *du* profil considéré – i.e. les profils qu'il suit – ou les abonnements *au* profil considéré – i.e. les profils qui le suivent. Leurs résultats suggèrent que les deux sources d'informations sont efficaces mais que les meilleurs scores sont obtenus en utilisant les abonnements *aux* profils.

2.2.2.2 *Détection de communautés*

Les modèles présentés ici tentent, pour leur part, de prédire les points de vue en s'appuyant sur des outils de détection de communautés des profils. Cherepnalkoski et Mozetic (2015) et Guerrero-Solé (2017) ont notamment utilisé un algorithme de détection de communautés sur un réseau de retweets pour détecter respectivement le groupe politique d'appartenance des parlementaires européens et la position – séparatiste ou unioniste – des profils Twitter sur le sujet de l'indépendance catalane. Rizos, Papadopoulos et Kompatsiaris (2017) se sont eux appuyés sur un algorithme de détection de communautés centré sur les profils, établi sur le voisinage social de ceux-ci. Leur méthode n'est pas conçue pour détecter spécifiquement les points de vue, mais simplement des communautés de profils semblables. Les points de vue peuvent toutefois être considérés comme un attribut caractérisant cette similarité. Leur modèle a été évalué sur des données issues de plusieurs médias sociaux, à savoir Twitter en utilisant les mentions et les retweets, ainsi que Youtube et Flickr avec les liens d'abonnement.

2.2.2.3 *Autres représentations de la similarité entre profils*

Il est bien entendu possible d'utiliser d'autres représentations des similarités inter-profils. Barberá (2015) s'est par exemple appuyé sur une modélisation bayésienne des liens d'abonnement pour estimer l'idéologie politique latente de profils Twitter sur une échelle gauche–droite. Ce modèle permet de retrouver de manière relativement fidèle les paysages politiques des États-Unis, du Royaume-Uni, de l'Espagne, de l'Italie et des Pays-Bas à partir des profils officiels des politicien·ne·s et de leur liste d'abonné·e·s. Il confirme également que,

si les politicien·ne·s sont aisément positionné·e·s au sein de leur parti, les profils « ordinaires » sont dans l'ensemble bien moins polarisés. Une autre représentation encore a été utilisé par Volkova, Bachrach et Durme (2016). Ils ont tenté de prédire, à partir des amis de profils Twitter, les traits psychodémographiques de ceux-ci, tels que le genre, l'âge, ou des positionnements sociaux similaires aux points de vue que nous avons définis précédemment, comme le fait d'être conservateur ou libéral, ou encore religieux ou athée. Ces traits étaient prédits à partir des *intérêts Twitter* : Twitter catégorise les profils très populaires en 26 classes d'intérêts (« actualités », « musique », « sport », « politique », . . .). Volkova, Bachrach et Durme utilisaient cette catégorisation pour déterminer les centres d'intérêts des profils, qui permettaient d'entraîner un modèle de régression linéaire prédisant les traits psychodémographiques. Leurs résultats indiquaient par exemple que les intérêts « business » et « golf » étaient plus liés aux profils conservateurs alors que les profils libéraux présentaient les intérêts « musique », « jeux » et « gouvernement » (le président des États-Unis était Obama au moment de la collecte des données). Il est important de noter que bien que l'utilisation des intérêts Twitter soit une idée intéressante, leurs données présentaient de sérieuses limites : les profils utilisés pour l'expérimentation étaient annotés sur leur traits psychodémographiques *perçus*, et non leurs traits *déclarés*, et les accords inter-annotateurs·trices portant sur les traits « orientation politique » et « religion » étaient particulièrement faibles (Volkova, Bachrach et Durme, 2016, Tableau 1). Nous pouvons aussi noter que les auteur·e·s déclaraient n'avoir mesuré l'accord inter-annotateurs·trices que sur *certains* des profils, sans aucune indication quantitative supplémentaire.

2.2.3 *Méthodes mixtes*

La majorité des modèles actuels tentent de tirer partie à la fois du contenu textuel et des interactions sociales. En effet, comme nous l'avons vu dans les sections précédents, chaque type d'information apporte des éléments de caractérisation pertinents pour détecter le point de vue d'un profil. De nombreuses combinaisons texte / interactions sociales sont possibles. Nous avons ici catégorisé les modèles en deux familles, selon leur fonctionnement : ceux utilisant les informations textuelles et sociales de façon conjointe et ceux ayant deux modules distincts qui se renforcent mutuellement afin de produire des prédictions plus fiables.

2.2.3.1 *Utilisation conjointe*

Les modèles présentés ici exploitent simultanément le contenu textuel publié par les profils et leurs interactions sociales. Makazhanov, Rafiei et Waqar (2014) ont ainsi construit un système de détection des points de vue politiques en se basant sur des *profils d'interaction*, construits pour chaque parti à partir des tweets de ses candidat·e·s. Un *profil d'interaction* est une liste ordonnée de termes caractéristiques du parti considéré. Chaque profil Twitter est ensuite

catégorisé en fonction des termes utilisés dans ses publications, en comparaison avec les *profils d'interaction* définis précédemment, et des interactions sociales ayant eu lieu entre le profil Twitter et les candidat·e·s (qu'il s'agisse de retweets, de mentions ou d'abonnements). MAGDY et al. (2016) ont tenté de déterminer s'il était possible de prédire l'attitude future de profils Twitter vis-à-vis d'un sujet sur lequel ils ne s'étaient pas nécessairement exprimés par le passé, en fonction de leur présence sur la plateforme. Ils ont collecté des données avant et après un événement majeur et polarisant, en l'occurrence les attentats de Paris du 13 novembre 2015, et prédit à l'aide d'un SVM entraîné sur le contenu textuel et le réseau d'interactions sociales pré-attentat les positions des profils sur l'Islam post-attentat. Leur principale observation est que l'élément déterminant pour cette tâche est le réseau social des profils, y compris pour ceux n'ayant jamais mentionné l'Islam dans leur tweets avant les attentats.

THONET et al. (2017) ont, pour leur part, exploité le contenu textuel et les interactions sociales pour déterminer de façon non supervisée les points de vue sur les réseaux sociaux. Le point fort de leur approche est de compenser la nature éparse des interactions sociales entre profils en essayant de tirer parti des liens faibles. Ces liens ne sont pas directement présents dans le graphe d'interactions, mais peuvent être conceptuellement définis comme une similarité probable entre *connaissances de connaissances*. Ils ont pour cela utilisé des urnes de Pólya généralisées, leur permettant de modéliser les liens entre profils en fonction de leurs interactions mais aussi de leur distance les uns par rapport aux autres. DONG et al. (2017) ont également proposé un modèle générique latent de détection de points de vue fondé sur les interactions sociales des utilisateurs dans une discussion / un débat et le contenu de leurs publications. Leur modèle a été conçu pour estimer à partir d'une petite quantité de données annotées les points de vue des profils et les distributions de mots des deux points de vue opposés sur la question étudiée. Il est intéressant de noter que leurs interactions sociales pouvaient être *négatives*, par exemple dans le cas d'une réponse étant en désaccord avec la publication initiale. De façon intéressante, TRABELSI et ZAIANE (2018) ont pris dans leur modèle latent de détection de points de vue l'hypothèse inverse de la plupart des modèles présentés ici, en se concentrant sur *l'hétérophilie*. Ils ont pour cela considéré que les différences de points de vue encourageaient les profils à interagir entre eux, permettant ainsi de détecter leur positionnement. Il est intéressant de noter qui si leur modèle obtient de bons résultats sur des sites explicitement dédiés aux débats, et donc où les utilisateurs·trices se rendent dans le but de défendre leurs idées et de se confronter à des opinions adverses, il n'est pas évalué sur les sites de débats « informels » que sont les réseaux sociaux grand public tels que Twitter et Facebook, sur lesquels la présence d'homophilie n'est plus à démontrer.

2.2.3.2 *Renforcement mutuel*

Les modèles fondés sur un renforcement mutuel entre contenu textuel et interactions sociales sont extrêmement variés, les combinaisons possibles étant nombreuses. Pour une meilleure lisibilité nous les avons séparés ici en trois

familles : les modèles utilisant en premier le contenu textuel, ceux utilisant d'abord le graphe social et ceux utilisant les deux alternativement jusqu'à l'obtention d'un équilibre.

TEXTE SUIVI DES INTERACTIONS SOCIALES. Ces méthodes de détection de points de vue sont fondées sur un modèle supervisé entraîné sur le contenu textuel des publications utilisé en prélude à l'utilisation d'un graphe représentant les interactions sociales des profils. Par exemple, PENNACCHIOTTI et POPESCU (2011) se sont intéressés au profilage des profils Twitter au sens large : détection de caractéristiques socio-économiques, sujets d'intérêts, points de vue, ... Ils ont pour cela utilisé un modèle de classification supervisée permettant une classification initiale des nouveaux profils suivi d'une phase de correction de prédiction s'appuyant sur le graphe des relations sociales. Le modèle supervisé se basait sur un boosting par gradient d'arbres de décision (GBDT pour *Gradient Boosted Decision Trees*) – un classifieur ensembliste qui consiste à créer plusieurs arbres de décision et à les agréger de façon à obtenir de meilleurs performances, de façon similaire aux forêts aléatoires. Les caractéristiques utilisées pour l'entraînement du modèle étaient extraites de la présentation des profils, du comportement sur la plateforme, des spécificités linguistiques du contenu textuel et enfin des interactions sociales. Le graphe des relations sociales était, lui, construit à partir des amis et des abonnés des profils et était présent pour corriger d'éventuelles erreurs de classification faites par le modèle GBDT : un profil catégorisé comme *Démocrate* mais ayant une majorité d'amis et d'abonnés *Républicain* était par exemple considéré comme ayant été mal catégorisé à l'étape précédente et mis à jour comme étant lui aussi *Républicain*. RABELO, PRUDENCIO et BARROS (2012) ont pour leur part tenté d'exploiter les graphes d'amis Twitter. La première étape de leur modèle consiste en un classifieur textuel qui catégorise les points de vue des profils sur un sujet donné. Les catégorisations considérées comme étant suffisamment fiables sont ensuite propagées sur le reste du graphe, de façon à affecter un point de vue au reste des profils, y compris les profils ne s'étant pas directement exprimés sur le sujet mais étant proches de profils l'ayant fait.

INTERACTIONS SOCIALES SUIVIES DU TEXTE. Ce modèle, proposé par RAJADESINGAN et LIU (2014), permet d'identifier des paires de profils ayant des opinions opposées à l'aide du réseau de retweets. À partir d'une poignée de profils dont les tweets sont sélectionnés pour être les graînes des deux points de vue opposés, un algorithme de propagation de labels catégorise une partie des tweets du réseau, puis un classifieur Naive Bayes détermine les labels des tweets restants en utilisant les tweets catégorisés à l'étape précédente comme données d'entraînement. Il est à noter que ce modèle a été conçu et évalué au niveau tweet, et non au niveau profil, mais il serait aisé de rajouter une phase finale d'agrégation pour déterminer le point de vue de chaque profil à partir des labels de ses tweets (en utilisant par exemple une simple règle de majorité).

ÉQUILIBRE ENTRE TEXTE ET INTERACTIONS. Ces modèles se fondent sur l'hypothèse que les points de vue déduits à partir du contenu textuel et ceux tirés du graphe social doivent être cohérents pour un même profil, et recherchent donc l'équilibre entre ces deux éléments. WONG et al. (2013) se sont concentrés sur l'estimation de l'orientation politique des médias d'actualités et des experts présents sur Twitter afin de mettre en évidence d'éventuels biais. Leur hypothèse est formalisée sous forme d'un *problème linéaire inverse mal posé* dont le but est de minimiser l'écart entre les orientations données par les tweets et les retweets. L'estimation de l'idéologie dans leur travail est principalement fondée sur les sentiments extraits des tweets. Leur second modèle prend en compte le fait que des profils similaires tendent à être retweetés par des audiences similaires afin d'estimer l'orientation politique de profils « standards » ne correspondant pas à des personnalités publiques (WONG et al., 2016). Ils ont pour cela posé un *problème d'optimisation convexe* qui permet d'intégrer à la fonction à maximiser la similarité des profils en plus de l'accord tweet-retweet.

2.2.4 *Synthèse*

Le Tableau 2.4 présente une synthèse des travaux présentés, en résumant quelles informations sont utilisées, le nombre de plateformes et de points de vue gérés et la quantité d'annotations nécessaire à leur bon fonctionnement. Il permet de mettre en évidence les limites principales des modèles existants :

1. Ils sont rarement génériques : la plupart des modèles sont conçus pour une plateforme spécifique et les auteur·e·s ne donnent pas d'indications sur un fonctionnement éventuel sur un autre média social. Ces modèles sont établis sur un élément spécifique d'information et ignorent une large partie de l'information disponible sur les profils.
2. Certains modèles ont besoin d'une grande quantité de données annotées afin de fonctionner correctement. Or, les jeux de données portant sur les médias sociaux étant généralement volumineux, cela implique une phase d'annotation au coût non négligeable, aussi bien en temps qu'en argent si les annotateurs·trices sont rémunérés. Une solution possible est d'utiliser les données annotées d'un autre jeu de données mais, chaque jeu de données ayant des caractéristiques particulières, les résultats sont habituellement dégradés et il est difficile d'évaluer cette dégradation.
3. La majorité des modèles ne gèrent que deux points de vue maximum. Il s'agit d'une limite majeure lorsque la situation à étudier est plus complexe qu'une simple opposition binaire, mais comprend plusieurs acteurs ayant chacun des intérêts distincts. Si nous voulions, par exemple, analyser la problématique du libre accès des articles scientifiques, nous pouvons lister au moins six acteurs : les auteur·e·s, les autres chercheurs·se·e, les universités, le grand public, les bibliothèques et les éditeurs·trices. Étant donné les motivations et réserves de chaque acteur, il serait difficile de simplement tous les placer sur un même axe « Pour / Contre ».

Nos contributions présentées en Partie II tentent de pallier ces problèmes.

TABLEAU 2.4 – Synthèse de la littérature existante sur la détection des points de vue statiques.
Les colonnes PLUS D'UNE PLATEFORME GÉRÉE et PLUS DE DEUX POINTS DE VUE GÉRÉS indiquent si le modèle proposé est défini pour *théoriquement* prendre en compte ces paramètres mais n'impliquent pas que ces cas de figure aient été évalués par les auteur·e·s. Le symbole ☑ indique les éléments présentant des limites.

	INFORMATIONS UTILISÉES			PEU D'ANNOTATIONS NÉCESSAIRES	PLUS D'UNE PLATEFORME GÉRÉE	PLUS DE DEUX POINTS DE VUE GÉRÉS
	Contenu textuel	Interactions sociales	Autres informations			
ABBASI et al. (2014)	✗	✓	✗	✗	✓	✓
AKOGLU (2014)	✓	✗	✗	✗	✓	✗
BARBERÁ (2015)	✗	✓	✗	✓	✗	☑ [a]
BOIREAU (2014)	✓	✗	✗	✓	✓	☑ [a]
CERON (2017)	✓	✗	✗	✓	✓	☑ [a]
CONOVER et al. (2011b)	✗	✓	✗	✓	✗	✓
DAVID et al. (2016)	✓	✗	✗	✓	✗	☑ [a]
DONG et al. (2017)	✓	✓	✗	✓	✓	✗
FANG et al. (2015)	✓	✗	✗	✓	✓	✓
HASAN et NG (2013)	✓	✗	✗	✓	✓	✗
MAGDY et al. (2016)	✓	✓	✓	✗	✗	✗
MAKAZHANOV, RAFIEI et WAQAR (2014)	✓	✓	✗	✗	✗	✓
MOHAMMAD, SOBHANI et KIRITCHENKO (2017)	✓	✗	✓	✗	✗	✗
PENNACCHIOTTI et POPESCU (2011)	✓	✓	✓	✗	✗	✗
RABELO, PRUDENCIO et BARROS (2012)	✓	✓	✗	✓	✗	✓
RAJADESINGAN et LIU (2014)	✓	✓	✗	✓	✗	✗
RIZOS, PAPADOPOULOS et KOMPATSIARIS (2017)	✗	✓	✗	✓	✓	✓
THONET et al. (2017)	✓	✓	✗	✓	✓	✓
VOLKOVA, BACHRACH et DURME (2016)	✗	✓	✗	✗	✗	✗
WONG et al. (2013)	✓	✓	✗	✓	✗	✗
WONG et al. (2016)	✓	✓	✗	✓	✗	✗
ZHANG et al. (2017)	✓	✗	✗	✗	✗	✗
ZUBIAGA et al. (2018)	✓	✓	✓	✗	✗	✓
Contributions (Partie II)	✓	✓	✓	✓	✓	✓

a. Même si l'utilisation d'une échelle de valeurs permet une catégorisation plus fine des points de vue, la logique sous-jacente reste binaire étant donné qu'il s'agit d'une opposition classique entre 2 valeurs.

2.3 DÉTECTION DE POINTS DE VUE DYNAMIQUES

Comme présenté dans la section précédente, la détection des points de vue statiques a été largement couverte par la littérature. L'aspect dynamique des points de vue et de leur évolution est, en revanche, en grande partie ignoré. L'enchaînement des messages ou des interactions d'une discussion est occasionnellement pris en compte pour déterminer le point de vue d'un profil à un moment précis (Zubiaga et al., 2018), mais peu de travaux ont étudié l'évolution temporelle des points de vue d'un même profil et la reconfiguration des communautés de points de vue au fil du temps sur des événements longs.

Smith et al. (2013) se sont intéressés à cette problématique en analysant les points de vue de profils Twitter sur plusieurs *propositions* politiques sur lesquelles les électeurs et électrices de Californie ont eu à se prononcer en 2012. Les thèmes des propositions apparaissant dans ce scrutin étaient variés, certaines portant sur les taxes, les OGM ou encore l'abolition de la peine de mort. Leur constatation principale était que l'immense majorité des profils semblait ne pas changer de points de vue, leur positionnement juste avant l'élection étant fortement corrélée avec leur positionnement du mois précédent – le fait d'évaluer l'évolution sur une période d'un mois seulement étant dans ce cas pertinente car les propositions sont habituellement publiquement débattues peu de temps avant l'élection. Ils avancent deux hypothèses principales pour l'absence de changements : (a) la nouveauté de la plateforme, qui impliquait qu'à l'époque les gens n'étaient pas encore habitués à l'utiliser pour exprimer leurs opinions, et (b) la possibilité que seules les personnes étant déjà fortement convaincues par un point de vue se sentiraient suffisamment à l'aise pour le partager publiquement via leur profil Twitter.

Gaumont, Panahi et Chavalarias (2017) ont tenté de capter les reconfigurations successives des communautés politiques des différent·e·s candidat·e·s de l'élection présidentielle française de 2017. Ils ont pour cela monitoré Twitter de juin 2016 à avril 2017 et effectué toutes les semaines une détection de communauté sur le graphe de retweets. Ils considéraient que les communautés contenant les profils officiels des candidat·e·s représentaient leur communauté politique proche et surveillaient donc les évolutions de celles-ci. Le graphe alluvial obtenu (voir la Figure 2.3) permet de voir les ruptures et les alliances ayant eu lieu au cours des mois, mais la limite de cette méthode est l'absence d'évaluation quantitative quand à l'assignation des points de vue aux profils.

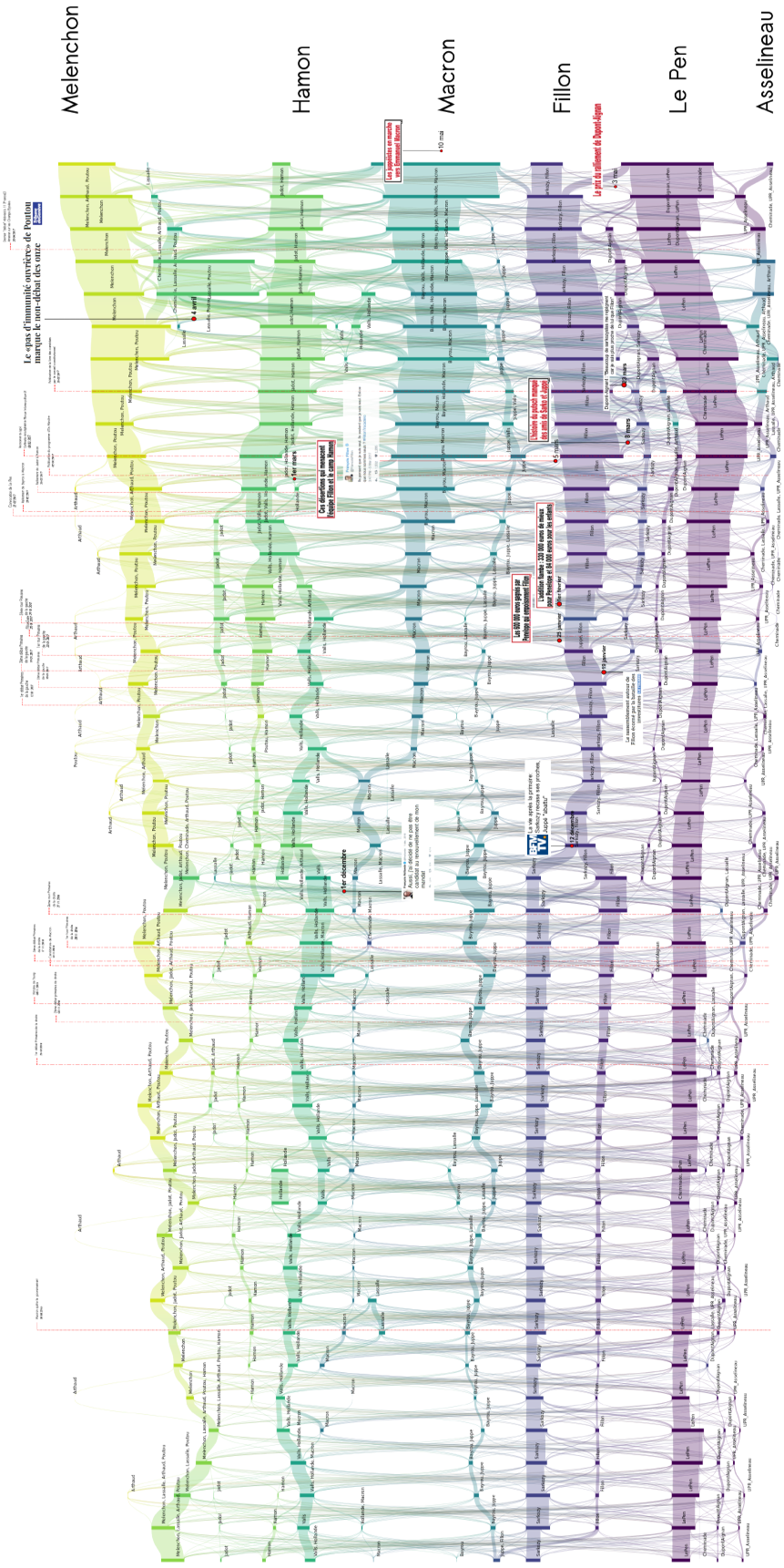


FIGURE 2.3 – Évolution des communautés politiques durant la campagne présidentielle française de 2017 (GAUMONT, PANAHI et CHAVALARIAS, 2017).

2.4 JEUX DE DONNÉES TWITTER PORTANT SUR LES POINTS DE VUE POLITIQUES

En raison de son usage répandu et de la facilité de collecte des données, plusieurs jeux de données ont été construits sur le thème des points de vue politiques des profils sur Twitter (voir Tableau 2.5)[7]. Ces jeux de données abordent des thèmes variés, mais malgré leur apport à l'analyse de discours politique en ligne, une quantité non négligeable d'entre eux sont construits à partir de méthodes automatiques – habituellement des méthodes Bayésiennes ou à base de graphes, s'appuyant sur les liens de retweets ou d'abonnements. L'absence de vérification manuelle rend ces derniers impossibles à utiliser pour évaluer et améliorer les méthodes de détection des points de vue étant donné que les erreurs ne peuvent pas être mesurées avec certitude.

Certains jeux de données annotés ou validés manuellement figurent cependant dans la littérature. Néanmoins, le Tableau 2.5 permet de constater que ces jeux de données capturent difficilement la complexité et le désordre des situations réelles. Ils sont habituellement centrés sur un petit nombre d'annotations, sur des points de vue binaires[8] (opposant par exemple « Démocrates » et « Républicains », ou « Séparatistes » et « Unionistes »), ou sur une catégorie spécifique de profils Twitter facilement identifiables et peu représentative des utilisateurs·trices standard, comme des profils de politicien·ne·s.

Malgré leurs limites, ces jeux de données ont été utilisés pour mettre au point des modèles de détection de points de vue, présentés dans la section précédente (Conover et al., 2011b ; Gayo-Avello, 2012 ; Jungherr, Jürgens et Schoen, 2012 ; Makazhanov, Rafiei et Waqar, 2014 ; Pennacchiotti et Popescu, 2011 ; Sang et Bos, 2012 ; Tumasjan et al., 2010 ; Weber, Garimella et Teka, 2013), mais également pour étudier des problématiques annexes comme les campagne de dénigrement et les « fake news » (Ratkiewicz et al., 2011 ; Saez-Trumper, 2014 ; Vosoughi, Mohsenvand et Roy, 2017) ou encore la polarisation sur Twitter (Conover et al., 2011a ; Fraisier et al., 2017 ; Weber, Garimella et Batayneh, 2013).

7. Nous nous concentrons ici sur les jeux de données donnant un point de vue au niveau « profil » et non sur les jeux de données annotant individuellement chaque publication.

8. Certains jeux de données fournissent des points de vue placés sur une échelle pour les catégoriser plus finement mais comme indiqué dans le Tableau 2.4, la logique sous-jacente reste binaire.

TABLEAU 2.5 – Jeux de données existant sur le thème des points de vue politiques au niveau profil sur Twitter.
Le symbole ✅ indique les points forts des jeux de données, le symbole ❌ leurs points faibles et ☑ indique les éléments avantageux mais présentant des limites. La colonne VÉRACITÉ indique si le jeu de données a été annoté ou validé manuellement, la colonne POINTS DE VUE si ces derniers sont binaires ou pas, et la colonne PROFILS si le jeu de données contient plus de 10 000 profils.

THÈME	DATE	VÉRACITÉ	POINTS DE VUE	PROFILS		SOURCE
Élection présidentielle aux États-Unis	2012	❌	☑ Échelle continue : libéral à conservateur	✅	50 000	BARBERÁ ET RIVERO (2015)
	2016	❌	❌ 2 points de vue opposés non explicités	✅	80 544	GARIMELLA ET AL. (2017)
Élection de mi-mandat aux États-Unis	2014	✅	❌ Démocrate, Républicain	❌	1 939 [a]	BRIGADIR, GREENE ET CUNNINGHAM (2015)
Positionnement politique aux États-Unis	2017	✅	☑ Échelle à 7 points : « très conservateur » à « très libéral »	❌	3 938 [b]	PREOŢIUC-PIETRO ET AL. (2017)
ObamaCare	2015	✅	☑ Échelle à 5 points : « fort support » à « forte opposition »	❌	504	LU, CAVERLEE ET NIU (2015)
	2016	❌	❌ 2 points de vue opposés non explicités	✅	334 617	GARIMELLA ET AL. (2017)
Référendum sur l'indépendance écossaise	2014	✅	❌ Oui, Non	❌	1 218 [a]	BRIGADIR, GREENE ET CUNNINGHAM (2015)
Brexit	2016	❌	❌ 2 points de vue opposés non explicités	✅	22 745	GARIMELLA ET AL. (2017)
Élection législatives en Espagne	2011	❌	☑ Échelle continue : libéral à conservateur	✅	12 000	BARBERÁ ET RIVERO (2015)
Élection sénatoriale aux Pays-Bas	2011	❌	✅ PVV, VVD, CDA, PvdA, SP, GL, D66, CU, PvdD, SGP, 50+, OSF	✅	$\sim 10^6$	SANG ET BOS (2012)
Élection fédérale en Allemagne	2009	❌	✅ CDU/CSU, SPD, FDP, Grüne, Linke, Piraten	✅	11 212	JUNGHERR (2013)
	2009	❌	✅ CDU/CSU, SPD, FDP, Grüne, Linke	✅	14 056	TUMASJAN ET AL. (2010)
	2017	✅	✅ CDU/CSU, SPD, FDP, Grüne, Linke, AfD	❌	364 [a]	KRATZKE (2017)
Avortement	2015	✅	☑ Échelle à 5 points : « fort support » à « forte opposition »	❌	504	LU, CAVERLEE ET NIU (2015)
	2016	❌	❌ 2 points de vue opposés non explicités	✅	279 505	GARIMELLA ET AL. (2017)
Contrôle des armes à feu	2015	✅	☑ Échelle à 5 points : « fort support » à « forte opposition »	❌	504	LU, CAVERLEE ET NIU (2015)
Fracturation hydrolique	2016	❌	❌ 2 points de vue opposés non explicités	✅	374 403	GARIMELLA ET AL. (2017)

a. ❌ Profils sélectionnés par des expert·e·s mais limités à des politicien·ne·s ou des activistes connu·e·s : pour le référendum écossais, des membres de la commission électorale ou des profils indiquant leur point de vue de façon claire dans leur biographie, pour les élections de mi-mandat aux États-Unis, des militant·e·s ou de candidat·e·s et pour les éléctions fédérales allemandes, des candidat·e·s.

b. ✅ Positionnement déclaré par les participant·e·s via des questionnaires.

3

CADRE THÉORIQUE

Afin de permettre une lecture aisée des chapitres développant nos contributions, nous présentons ci-après le cadre théorique utilisé. Nous introduisons dans la première section de ce chapitre les définitions des concepts de base de la théorie des graphes ainsi que les définitions des concepts incontournables pour comprendre notre travail. La seconde section présente, elle, un aperçu des algorithmes de détection de communautés recouvrantes et non recouvrantes.

3.1 DÉFINITIONS

3.1.1 Graphe

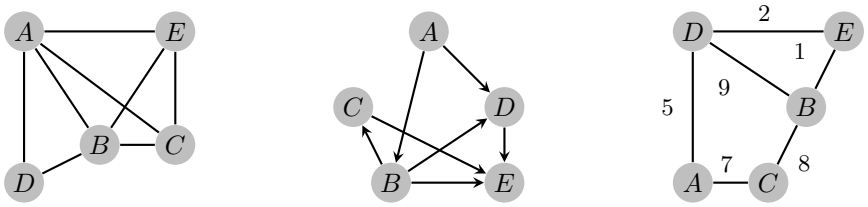
Un graphe, usuellement noté $G = (V, E)$, est un ensemble de nœuds ou vertex V, reliés entre eux par un ensemble d'arêtes E, chaque arête étant une paire d'éléments contenus dans V. Deux nœuds sont dit adjacents s'ils sont reliés par une arête. Les arêtes peuvent être orientées : elles sont alors appelées des arcs et le graphe est dit dirigé. Il peut également être pondéré si un poids est affecté à chacune de ses arêtes. Nous considérons ici uniquement le cas standard d'un graphe simple, c'est-à-dire un graphe qui ne peut posséder qu'une seule arête au plus entre deux même nœuds. La Figure 3.1 présente des exemples de graphes illustrant ces concepts.

3.1.2 Matrice d'adjacence

Soit un graphe $G = (V, E)$. Sa représentation matricielle est la matrice A de dimension $|V| \times |V|$, avec

$$a_{ij} = \begin{cases} w_{ij} & \text{si } (v_i, v_j) \in E \\ 0 & \textit{sinon.} \end{cases}$$

avec w_{ij} le poids de l'arête allant de v_i à v_j. Si G est non pondéré, $a_{ij} = 1$. Par construction, A est symétrique si G est non dirigé car $a_{ij} = a_{ji}$. La Figure 3.2 présente les matrices d'adjacence des graphes présentés dans la Figure 3.1.



(a) Graphe non dirigé (b) Graphe dirigé (c) Graphe non dirigé pondéré

FIGURE 3.1 – Exemples de graphes.

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix} \end{array}$$

(a) Graphe 3.1a

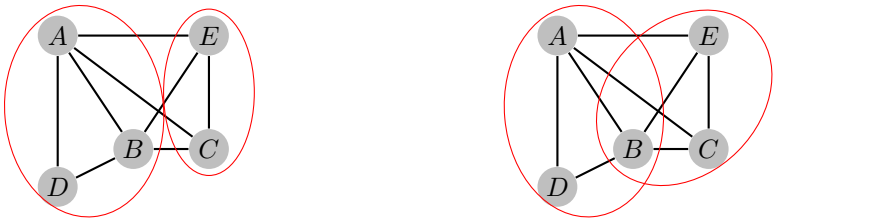
$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ P \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

(b) Graphe 3.1b

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 0 & 0 & 7 & 5 & 0 \\ 0 & 0 & 8 & 9 & 1 \\ 7 & 8 & 0 & 0 & 0 \\ 5 & 9 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 & 0 \end{pmatrix} \end{array}$$

(c) Graphe 3.1c

FIGURE 3.2 – Exemples de matrices d'adjacences.



Communautés : $((A, B, D), (C, E))$ Communautés : $((A, B, D), (B, C, E))$

(a) Communautés non recouvrantes (b) Communautés recouvrantes

FIGURE 3.3 – Exemples de communautés.

3.1.3 Degré

Le degré d'un nœud v_i est le nombre d'arêtes connectant v_i au reste du graphe, aussi nommées les arêtes adjacentes à v_i. Si le graphe est pondéré, il s'agit de la somme des poids des arêtes.

$$\text{Degré}(v_i) = \sum_{j=1}^{|V|} a_{ij}$$

avec a_{ij} les éléments de la matrice d'adjacence A.

Dans un graphe dirigé, le degré entrant est le nombre d'arcs entrant sur v_i et respectivement le degré sortant est le nombre d'arcs partant de v_i, ou la somme de leur poids dans le cas d'un graphe dirigé pondéré.

3.1.4 Clique

Dans un graphe non orienté, une clique est un sous-ensemble de nœuds formant un sous-graphe complet, c'est-à-dire dont les nœuds sont tous adjacents deux à deux. Une p-clique désigne une clique contenant p nœuds. Deux p-cliques sont dites adjacentes si elles partagent $p-1$ nœuds.

Le graphe 3.1a contient par exemple six cliques :

4-clique : (A, B, C, E)

3-cliques : $(A, B, C); (A, B, D); (A, B, E); (A, C, E); (B, C, E)$

La clique (A, B, D) est adjacente aux cliques (A, B, C) et (A, B, E) mais pas à (A, C, E) et (B, C, E).

3.1.5 Communauté

Une communauté est un groupe de nœuds plus fortement connectés entre eux qu'au reste du graphe. De nombreux travaux traitent de communautés, chacun en donnant une définition légèrement différente (Fortunato, 2010, p. 83), mais l'idée de densité intra-cluster est commune à toutes les définitions. Les communautés sont dites recouvrantes lorsqu'un nœud peut appartenir simultanément à plus d'une communauté.

Les communautés sont parfois notées comme un ensemble de listes de nœuds, chaque liste représentant une communauté, et chaque nœud étant inclus dans une seule liste dans le cas de communautés non recouvrantes, ou dans au moins une liste dans le cas de communautés recouvrantes. La Figure 3.3 présente des exemples de communautés.

3.1.6 Multiplex

Un multiplex est un graphe multi-couche dans lequel chaque couche a son propre ensemble d'arêtes tout en partageant le même ensemble de nœuds V.

3.2 DÉTECTION DE COMMUNAUTÉS

Notre travail s'appuie en grande partie sur la détection de communautés de profils. Afin de familiariser le lecteur avec ce domaine, nous présentons donc de façon non exhaustive dans la section suivante quelques algorithmes incontournables de détection de communautés. Nous nous concentrons tout d'abord sur la détection de communautés non recouvrantes puis, dans un second temps, sur les communautés recouvrantes et leurs spécificités.

3.2.1 *Communautés non recouvrantes*

De nombreux algorithmes ont été proposés dans la littérature pour détecter des communautés non recouvrantes. Nous ne présenterons ici qu'une sélection de certains des algorithmes les plus connus, répartis en trois catégories : les algorithmes fondés sur la maximisation de la modularité, ceux fondés sur la diffusion de l'information, et ceux fondés sur des marches aléatoires. Le Tableau 3.1 résume quelques caractéristiques des algorithmes présentés ci-dessous.

3.2.1.1 *Maximisation de la modularité*

La modularité est une mesure permettant de quantifier la qualité d'un partitionnement de graphe en comparant les arêtes intra-communautés et inter-communautés (Newman, 2006). Elle peut être exprimée comme :

$$Q = \frac{1}{2|E|} \sum_{ij} \left[a_{ij} - \frac{k_i \times k_j}{2|E|} \right] \delta(c_i, c_j)$$

avec k_i la somme des poids des arêtes adjacentes au nœud v_i, c_i la communauté assignée à v_i, et $\delta(c_i, c_j) = \begin{cases} 1 & \text{si } c_i = c_j \\ 0 & \text{sinon} \end{cases}$.

Nous présentons dans cette section trois algorithmes fondés sur la modularité : *Fast-greedy*, un algorithme fondé sur les vecteurs propres de la matrice de modularité et *Multi-level*.

L'algorithme *Fast-greedy* (Wakita et Tsurumi, 2007) détermine la partition donnant la plus grande modularité en fusionnant successivement des paires de communautés, en commençant par les plus petites communautés possibles, c'est-à-dire des communautés ne contenant chacune qu'un seul nœud. Newman (2006) propose un algorithme de détection de communautés fondés sur les *vecteurs propres* de la matrice de modularité d'un graphe, déterminée à l'aide des caractéristiques du graphe et indépendante de toute partition. À partir des valeurs propres et des vecteurs propres de cette matrice, il détermine ensuite les communautés maximisant la modularité. Enfin, l'algorithme *Multi-level* (Blondel et al., 2008) fournit une structure communautaire hiérarchique obtenue en calculant itérativement des maxima locaux de modularité. Chaque maximum local est le résultat de deux étapes simples : (a) chaque

TABLEAU 3.1 – Caractéristiques des algorithmes de détection de communautés non recouvrantes, avec $n = |V|$ et $m = |E|$.

		ÉLÉMENTS GÉRÉS		
CATÉGORIE	ALGORITHME	Poids	Direction	COMPLEXITÉ [a]
Maximisation de la modularité	Fast-greedy	✅	❌	$O(mn \log n)$
	Vecteurs propres	❌	❌	$O(n^2)$
	Multi-level	✅	❌	$O(n)$
Diffusion de l'information	Infomap	✅	✅	$O(n(n+m))$
	Propagation de labels	✅	❌	$O(n+m)$
Marche aléatoire	Walktrap	✅	❌	$O(n^2 \log n)$

a. Sur données éparses typiques.

nœud est déplacé individuellement dans la communauté permettant d'obtenir la partition donnant la meilleure modularité, (b) lorsque que la modularité ne peut plus être améliorée, un nouveau graphe est construit, dont les nœuds sont les communautés détectées à l'étape précédente, et les poids des arêtes sont la somme des poids des arêtes entre nœuds des différentes communautés considérées. Cette méthode est particulièrement reconnue pour sa rapidité d'exécution sur de très grands graphes, grâce à sa complexité linéaire.

3.2.1.2 *Diffusion de l'information*

Certains algorithmes de détection de communautés reposent sur le principe de diffusion de l'information, c'est-à-dire que plus le lien existant entre deux nœuds est important, plus l'information va se propager facilement entre eux. De façon symétrique, il est donc possible d'utiliser la facilité de propagation d'une information pour déterminer quels nœuds sont fortement liés. Nous présentons ici deux algorithmes fonctionnant sur ce principe : tout d'abord *Infomap* puis la propagation de labels. *Infomap* (ROSVALL, AXELSSON et BERGSTROM, 2009) repose sur le flux d'information dans un graphe plutôt que sur ses caractéristiques topographiques. Ce flux peut être imaginé comme un message allant d'un nœud-expéditeur à un nœud-destinataire. Le but d'*Infomap* est de minimiser la représentation du trajet théorique d'un message, et pour cela l'algorithme va regrouper en communautés certains nœuds afin d'optimiser cette représentation. La propagation de labels (RAGHAVAN, ALBERT et KUMARA, 2007) est un autre processus dynamique, fonctionnant par itération : à chaque étape, chaque nœud adopte aléatoirement le label de l'un de ses voisins. L'algorithme s'arrête lorsque la partition est stabilisée, c'est-à-dire lorsque les labels ne changent plus, ou que le nombre maximal d'itérations a été atteint (il peut arriver que certains nœuds aient le même nombre de voisins ayant des labels différents et ne puissent donc pas se stabiliser).

3.2.1.3 *Marche aléatoire*

Une marche aléatoire est une modélisation mathématique d'une succession de pas effectués au hasard dans un système. Les marches aléatoires permettent de simuler certains phénomènes naturels et sont très utilisées en informatique et mathématique pour mesurer des distances entre éléments. L'algorithme *Walktrap* (Pons et Latapy, 2005) s'appuie sur ce concept pour proposer une structure communautaire hiérarchique. Les marches aléatoires dans le graphe permettent de mesurer des distances entre les nœuds, qui peuvent ensuite être regroupés en communautés, elles-même fusionnées par la suite lorsqu'elles sont adjacentes.

3.2.2 *Communautés recouvrantes*

La détection de communautés est une problématique très étudiée dans la littérature, mais la majorité des modèles proposés postulent qu'un nœud n'appartient qu'à une seule communauté. Cependant, il existe de nombreuses situations pour lesquelles il est pertinent qu'un nœud appartienne simultanément à plusieurs communautés. Ce cas de figure est géré par les algorithmes de détection de communautés recouvrantes.

L'une des méthodes les plus connues est la *percolation de cliques* ou *CPM* (*Clique Percolation Method*) (Palla et al., 2005). Une communauté est définie comme l'union de k-cliques adjacentes. La méthode de base considère le graphe comme étant non dirigé et non pondéré mais des variantes ont été proposées, telles que *CPMw* pour gérer les graphes pondérés (Farkas et al., 2007). Le modèle *MMSB* (*Mixed Membership Stochastic Blockmodel*) (Airoldi et al., 2008) est, lui, une extension probabiliste d'un modèle classique de factorisation de matrice à blocs latents et permet d'associer à chaque nœud un vecteur de probabilités d'appartenance aux communautés. *OSLOM* (*Order Statistics Local Optimization Method*) (Lancichinetti et al., 2011) se fonde sur l'optimisation d'une fonction exprimant la signification statistique des communautés détectées et peut prendre en compte l'aspect dirigé et pondéré d'un graphe.

D'autres méthodes sont dérivées du partitionnement classique en k-moyennes (MacQueen, 1967) bien connu en statistiques, qui consiste à diviser en k partitions des observations de façon à ce que chaque observation appartienne à la partition la plus proche d'elle. Dans le *partitionnement flou en c-moyennes* (Dunn, 1973), plutôt que de n'appartenir qu'à une seule partition, chaque observation m possède pour chaque partition n un degré d'appartenance a_m^n, avec $a_m^n \in [0;1]$ et $\sum_{\forall m} a_m^n = 1$. Dans le *partitionnement grossier en k-moyennes* (Lingras et West, 2004), chaque partition est décrite par une approximation inférieure et supérieure, la différence entre les deux étant nommée « zone limite ». Une observation peut soit faire partie d'une seule approximation inférieure ou soit appartenir à au moins deux zones limites.

Il existe également des extensions de la propagation de label classique, comme l'approche *CDLPOV* (*Core Detection Label Propagation with Overlapping*)

(Attal, Malek et Zolghadri, 2016), qui exploite son aspect non-déterministe : le principe est d'exécuter plusieurs fois l'algorithme et de repérer les nœuds appartenant tout le temps à une communauté commune et ceux présents alternativement dans plusieurs communautés, qui appartiennent donc à une zone de recouvrement.

3.3 CONCLUSION

Nous avons présenté dans ce chapitre le vocabulaire et les notions de base nécessaires à la bonne compréhension des graphes et de leurs utilisations. Outre les définitions des termes incontournables de la théorie des graphes et des termes requis pour bien appréhender nos contributions, présentés dans la Section 3.1, nous avons également présenté dans la Section 3.2 un panorama des différents algorithmes de détection de communautés, aussi bien les communautés disjointes que recouvrantes. Il est important de noter que nous nous sommes concentrés ici uniquement sur les algorithmes les plus classiques de détection de communautés, mais il existe de nombreuses autres familles de modèles, tels que ceux basés sur le clustering spectral (Shi et Malik, 2000) ou sur les modèles de mélange gaussiens (McLachlan et Basford, 1988). Pour les lecteurs·trices qui voudraient une vision plus exhaustive de la littérature du domaine, nous recommandons l'état de l'art de Bedi et Sharma (2016). Dû à notre sélection de modèles, nous n'avons pas pris en compte les travaux les plus récents, mais nous devons mentionner que le champ de recherche est très actif et que des approches prometteuses sont apparues ces dernières années, telles que celles basées sur l'apprentissage de représentations et notamment les plongements de graphes (Cavallari et al., 2017 ; Wang et al., 2017).

Le Chapitre 2 ayant introduit les définitions de nos concepts centraux et présenté l'état de l'art des travaux de détection de points de vue sur les médias sociaux et celui-ci ayant posé le cadre théorique, nous allons présenter dans les chapitres suivants nos contributions.

4

CONSTITUTION DU CORPUS #ÉLYSÉE2017FR

La Section 2.4 permet de constater que les jeux de données existants ont d'importantes limites, malgré le fait qu'ils soient essentiels à l'amélioration et la mesure des performances des modèles de détection de points de vue. Nous pouvons notamment citer le manque de jeux de données considérant plus de deux points de vue et la taille limitée de la majorité des jeux de données existants. Nous avons donc construit un jeu de données original, intitulé #Élysée2017fr, afin de proposer un cas d'étude plus proche des situations réelles (Fraisier et al., 2018) [1].

Ses caractéristiques principales sont :

- 22 853 profils Twitter débattant d'un événement politique français majeur, annotés manuellement par des experts ;
- 6 points de vue politiques (pour les 5 principaux partis politiques et une catégorie *indéfini*) ;
- des affiliations multiples aux partis politiques ;
- des informations complémentaires concernant la nature des profils et le sexe des propriétaires, lorsque ces informations étaient inférables ;
- les identifiants des 2 414 584 tweets et 7 763 931 retweets publiés par ces profils, discutant de l'élection présidentielle, en plusieurs langues ;
- les réseaux de retweets et de mentions entre profils annotés.

Ce jeu de données peut être utilisé tel quel par les chercheurs en science politique pour étudier les mécanismes de la campagne présidentielle sur Twitter, mais il peut également être utilisé par les chercheurs en informatique pour évaluer les modèles de détection de points de vue ou les outils d'analyse de réseaux grâce aux annotations manuelles. De plus, sa taille conséquente comparativement à l'existant garantit une excellente robustesse pour les modèles d'apprentissage supervisé, et la présence de communautés politiques recouvrantes permet l'exploration de questions de recherche plus poussées, comme l'identification des électeurs·trices indécis·es.

1. https://dataverse.mpi-sws.org/dataset.xhtml?persistentId=doi:10.5072/FK2/JBNEX2

4.1 PRÉSENTATION DU CONTEXTE : LA CAMPAGNE PRÉSIDENTIELLE FRANÇAISE DE 2017

Ce jeu de données a été collecté durant la campagne présidentielle française de 2017, qui a pris fin le 7 mai avec l'élection du président français actuel, Emmanuel Macron. Au lieu du scénario habituel consistant en une élection présidentielle dominée par les candidat·e·s des partis historiques de gauche, le Parti Socialiste, et de droite, les Républicains [2], cette campagne a été fortement atypique et imprévisible, avec des allégeances mouvantes entre les cinq principaux partis, ce qui en fait un terrain pertinent pour l'analyse et la détection de points de vue.

Avant la campagne présidentielle, deux primaires ont été organisées. Bien que ces primaires aient abouti à la nomination de *François Fillon*, des *Républicains*, comme candidat de la droite et du centre et de *Benoît Hamon*, du *Parti Socialiste*, comme candidat de la gauche, elles ont aussi déclenché de nombreux conflits de chaque côté. La forte présence médiatique du mouvement d'*Emmanuel Macron*, *En Marche*, créé en avril 2016 et présenté comme un mouvement destiné à unir la droite et la gauche, a rendu la situation plus instable encore. Les extrêmes étaient aussi très actifs durant cette campagne. Alors que le *Front National*, le mouvement d'extrême-droite de *Marine Le Pen*, s'était imposé comme une figure incontournable des deux dernières élections présidentielles françaises, la campagne de 2017 a également vu l'émergence d'un important mouvement d'extrême-gauche, la *France Insoumise*, mené par *Jean-Luc Mélenchon*. De plus, la campagne a été fortement ébranlée par les révélations d'emplois fictifs de Penelope Fillon, l'épouse de François Fillon, et de leurs enfants. François Fillon était favori lorsque cette affaire a éclaté, mais sa décision de rester candidat malgré sa mise en examen a conduit à de nouvelles dissensions, certains personnages historiques du parti et des électeurs ayant préféré reporter leur vote sur Marine Le Pen ou sur Emmanuel Macron.

4.2 MÉTHODOLOGIE DE COLLECTE

4.2.1 *Présentation rapide de l'API Streaming de Twitter*

Twitter met à disposition du public plusieurs API permettant d'interagir avec la plateforme et de collecter des données. L'une des plus utilisées est celle permettant de filtrer les publications en temps réel, en fonction de critères définis par l'utilisateur·trice, communément appelée "Streaming API" [3]. Elle permet de (a) collecter les tweets de profils donnés, (b) collecter les tweets contenant des mots-clés spécifiés et (c) collecter les tweets provenant d'une

2. Appelé l'Union pour un mouvement populaire (UMP) jusqu'au 30 mai 2015.

3. Documentation officielle : https://developer.twitter.com/en/docs/tweets/filter-realtime/api-reference/post-statuses-filter.html

zone géographique précise. La limite principale de cette méthode de collecte est le fait que Twitter ne permette pas de collecter plus d'1% de son trafic total à un instant t. Lors d'événements très commentés, il arrive donc fréquemment que Twitter ne restitue qu'une partie des tweets pertinents, sans qu'il soit possible de savoir combien de tweets sont ignorés.

4.2.2 *Sélection des mots-clés*

Le jeu de données #Élysée2017fr a été construit à l'aide de l'outil DMI-TCAT (Borra et Rieder, 2014), en surveillant en temps réel l'utilisation de plusieurs mots-clés faisant référence aux partis politiques et à leurs candidat·e·s, du 25 novembre 2016 au 12 mai 2017. Ces 159 mots-clés ont été sélectionnés par des chercheurs familiers avec le paysage politique français sur Twitter et sont détaillés en annexe dans le Tableau B.2.1. Afin de simplifier l'opération d'annotation, nous nous sommes concentrés sur les cinq principaux partis en lice. Le but étant de capturer au mieux les débats ayant lieu sur Twitter, la répartition des mots-clés n'est pas nécessairement uniforme entre partis : la communauté `LR` a par exemple été très active et a généré de nombreux nouveaux mots-clés durant la campagne comparativement aux communautés `PS` ou `EM`. La liste de mots-clés a donc été mise à jour au fur et à mesure de la campagne afin de s'adapter aux rebondissements et aux évolutions. La plupart d'entre eux sont construits à partir des noms des candidat·e·s, des noms des partis et de leurs slogans (« enbleumarine » et « AuNomDuPeuple » pour le `FN`, « AvenirEnCommun » et « #6èmeRépublique » pour `FI`, ...), parfois accolés à des termes permettant de cibler l'élection : « élection », « présidentielle », « vote », « 2017 », ... Certains sont destinés à cibler des événements précis de la campagne comme les primaires (« PrimaireDeGauche »), des meetings (« #macronlyon ») ou le scandale de l'affaire Fillon (« planB »).

4.3 SÉLECTION DES PROFILS À ANNOTER

Bien que DMI-TCAT ait souvent atteint la limite imposée par Twitter (voir Section 4.2.1) de par l'intensité des débats, la phase de collecte décrite dans la Section 4.2 a permis d'obtenir 42 251 431 tweets et retweets, publiés par 2 941 991 profils. Compte tenu de la quantité de profils collectés, il aurait été impossible de tous les annoter. Nous avons donc dû en échantillonner un sous-ensemble, en considérant les critères suivants :

1. Déterminer l'idéologie politique d'un profil à partir de ses publications est une tâche complexe, même pour une personne, et tout particulièrement sur Twitter à cause de la taille réduite des publications. Afin de produire des annotations fiables, nous n'avons considéré que les profils mentionnant, à n'importe quel moment de la campagne, l'un des cinq principaux partis dans leur pseudonyme, leur nom, ou leur biographie ;

2. Les profils ayant moins de 10 publications (en considérant tweets et retweets) durant les 7 mois de collecte ont aussi été écartés. Cela nous a permis d'exclure les profils les plus inactifs, c'est-à-dire ayant tweeté ou retweeté en moyenne moins de 1,4 fois par mois.

L'échantillon résultant de cette sélection comptait 23 169 profils. Bien que cela ne représente qu'1% des profils présents dans le jeu de données originel, il s'agit d'un ensemble bien plus imposant que les jeux de données existants jusqu'à présent (cf. Section 2.4).

4.4 PROTOCOLE D'ANNOTATION DE L'IDÉOLOGIE POLITIQUE DES PROFILS

4.4.1 *Pourquoi choisir une annotation par experts ?*

De nombreux jeux de données sont annotés à l'aide du « crowdsourcing », c'est-à-dire en sous-traitant l'annotation à une grande quantité de personnes. L'idée est de diviser cette dernière en une multitude de petites tâches rapides à compléter et payées à l'unité. Ainsi, un processus d'annotation qui serait très long à réaliser pour une poignée d'annotateurs·trices peut être réalisé en quelques jours, voire en quelques heures, grâce à des plateformes [4] co-ordonnant des centaines ou des milliers de personnes.

Mais dans notre cas, le crowdsourcing présentait de nombreuses limites. En effet, nous aurions eu besoin de personnes ayant un excellent niveau de français, étant donné qu'il s'agissait de la langue utilisée par la majorité des profils collectés, ainsi qu'une connaissance non négligeable du paysage politique français, actuel et passé, et des événements importants de la campagne. Des qualifications aussi strictes auraient considérablement réduit la quantité de personnes susceptibles de compléter nos tâches d'annotations, diminuant du même coup l'attrait principal du crowdsourcing. Étant donné le niveau d'expertise demandé, cela se serait également traduit par un prix unitaire par tâche bien plus élevé que la moyenne, ce qui, au vu de la taille importante de notre dataset, aurait représenté un coût global prohibitif.

Nous avons donc décidé de ne pas faire appel au crowdsourcing, et au contraire de ne faire intervenir que des expert·e·s. Ceci nous permettait également de garantir une meilleure qualité pour les annotations fournies dans ce jeu de données.

4. Les deux plateformes les plus connues sont à l'heure actuelle Amazon Mechanical Turk et Figure Eight (précédemment CrowdFlower).

4.4.2 *Annotation d'un profil*

4.4.2.1 *Éléments présentés à l'annotateur·trice*

Les données collectées pour chaque profil étaient extrêmement denses et difficilement exploitables présentées telles quelles. Afin de faciliter la tâche des 16 annotateurs·trices, chaque profil a donc été résumé par :

— l'ensemble de ses pseudonymes utilisés pendant la campagne [5] ;

— l'ensemble de ses noms utilisés pendant la campagne [5] ;

— l'ensemble de ses biographies utilisées pendant la campagne [5] ;

— ses 10 publications les plus partagées (en incluant ses propres tweets mais aussi les retweets de sa part).

La Figure 4.1 présente un exemple de profil résumé à l'aide des éléments décrits ci-dessus.

4.4.2.2 *Annotations attendues*

L'annotation principale pour chaque profil était son affiliation politique. Comme présenté dans la Section 4.1, la campagne présidentielle française de 2017 a été dominée par cinq partis politiques, sur lesquels nous nous sommes concentrés. Ils sont présentés ici de l'extrême-gauche à l'extrême-droite :

`FI` La France Insoumise, menée par Jean-Luc Mélenchon,

`PS` Le Parti Socialiste, mené par Benoît Hamon,

`EM` En Marche !, mené par Emmanuel Macron,

`LR` Les Républicains, menés par François Fillon,

`FN` Le Front National, mené par Marine Le Pen.

Les annotateurs·trices devaient déterminer l'idéologie politique en fonction du soutien exprimé envers les partis ou les candidat·e·s. Une affiliation officielle (être membre d'un parti par exemple) n'était pas nécessaire pour considérer qu'un profil était proche d'un parti ou d'un·e candidat·e. De même, un vote par défaut ne constituait pas une preuve d'affiliation à nos yeux (par exemple, un tweet indiquant « Je vote contre Le Pen avec un bulletin Macron » ne comptait pas comme un soutien à `EM`). Étant donnée la nature complexe de la campagne, certains profils Twitter pouvaient être affiliés à plusieurs partis. Le Tableau 4.1 présente quelques exemples de ce phénomène. Si les éléments présentés pour un profil (voir Section 4.4.2.1) n'étaient pas suffisants pour déterminer une allégeance politique, le profil était considéré comme étant politiquement *indéterminé*. Cela incluait donc les profils discutant de la campagne de façon neutre mais également les gens qui semblaient exhiber une orientation politique mais

5. DMI-TCAT collectant pour chaque tweet les informations sur son auteur·e, les changements de pseudonymes, de noms ou de biographies était enregistrés.

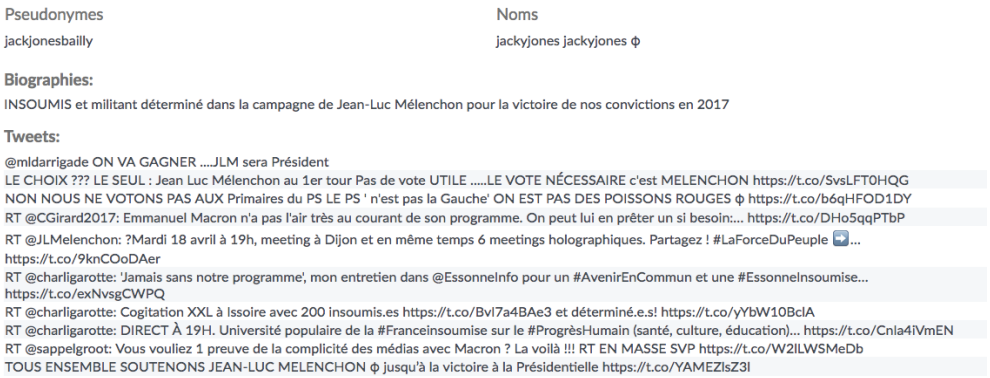


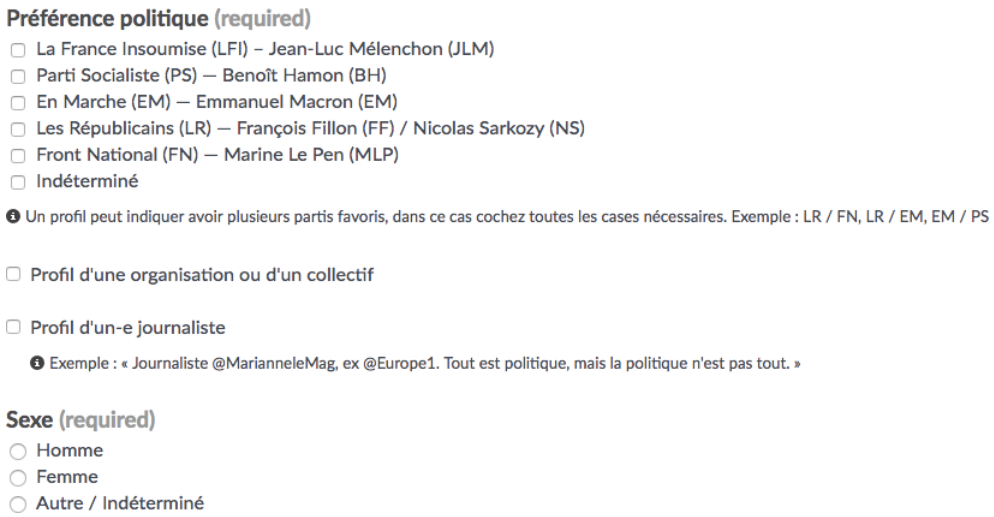
FIGURE 4.1 – Exemple de présentation d'un profil Twitter lors de la tâche d'annotation.

TABLEAU 4.1 – Exemples d'extraits de profils Twitter présentant des affiliations politiques multiples durant la campagne présidentielle française de 2017.

PARTIS	EXTRAITS	EXPLICATIONS
PS / EM	Fidélité @fhollande soutien #EnMarche	« @fhollande » fait référence à François Hollande, le précédent président socialiste français.
LR / FN	#FillonisteAvecMarine [...] #JamaisMacron	« Marine » fait référence à Marine Le Pen, leader du FN.
EM / LR	#EnMarche Ex-Jeune avec Juppé	Alain Juppé était l'un des leaders de LR éliminé durant la primaire de droite.

pour lesquels nous ne disposions pas de suffisamment d'éléments pour les catégoriser. Dû à la nature de la collecte, avec un simple filtre par mots-clés, certains profils pouvaient éventuellement être hors sujet et bruiter notre jeu de données. Les cas les plus fréquents de hors sujet étaient dus à la marque de vêtements de sport *Macron* et au hashtag #MLP, qui dans notre cas faisait référence à *Marine Le Pen*, mais est également un anagramme utilisé pour faire référence à la marque de jouets *My Little Pony* ainsi qu'à tous ses produits dérivés. Les 316 profils considérés comme étant hors sujet ont été simplement exclus du jeu de données.

En plus de l'affiliation politique, de nombreux éléments pourraient entrer en compte dans la façon dont le point de vue politique transparaît dans le profil : il semble assez naturel que l'expression du point de vue soit différente entre un profil géré par un groupe de militant·e·s et le profil personnel d'un individu, le profil professionnel d'une journaliste, ou le profil officiel d'un journal régional.



(a) Pour un profil individuel.

(b) Pour un profil non individuel.

FIGURE 4.2 – Formulaire d'annotation d'un profil Twitter.

Les annotations suivantes étaient donc également demandées afin d'améliorer la description des profils :

— la nature du profil : géré par une personne individuelle ou non ;
— pour les profils gérés par une personne : le sexe de celle-ci lorsque l'information était inférable, et si cette dernière s'identifiait comme étant un·e professionnel·le des médias (journaliste, éditorialiste, ...) ;
— pour les autres profils, si le groupe ou l'entité morale gérant le profil était de nature politique (par exemple un groupe de militant·e·s ou un profil officiel de parti politique), médiatique (par exemple un profil officiel de journal), ou autre.

CrowdFlower a été utilisé pour créer une tâche d'annotation accessible exclusivement aux annotateurs·trices que nous avions sollicité. L'utilisation de la plateforme présentait deux avantages principaux : CrowdFlower permettait de créer un formulaire présentant les annotations attendues de façon plus agréable et efficace (voir Figure 4.2) et gérait automatiquement la répartition des annotations entre les annotateurs·trices connecté·e·s.

4.4.3 *Mesure de l'accord inter-annotateurs*

L'un des indicateurs de qualité incontournables pour un jeu de données annoté est l'accord inter-annotateurs : afin de garantir que les annotations soient les plus objectives et pertinentes possible, tout ou partie du jeu de données est annoté par plusieurs personnes, et la similarité entre leurs annotations respectives est calculée. Plus cette similarité est importante, plus les annotations sont considérées fiables, alors que si elle est trop faible, les annotations seront difficilement exploitables. Cela peut notamment arriver lorsque les consignes de la tâche d'annotation sont trop ambiguës et peuvent donc être interprétées de plusieurs façons. Traditionnellement, un accord inter-annotateur est mesuré de la façon suivante :

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad \in [-1; 1]$$

avec p_o l'accord observé, c'est-à-dire l'accord moyen sur tous les éléments, et p_e l'accord attendu, c'est-à-dire l'accord mesuré lorsque les éléments sont assignés à des catégories de façon aléatoire.

Bien que de nombreuses mesures de l'accord inter-annotateurs existent dans la littérature (Artstein et Poesio, 2008), les plus connues ne peuvent prendre en compte qu'une catégorisation unique par élément. Le kappa de Cohen (1960) mesure l'accord entre deux évaluateurs·trices A et B classant chacun N éléments dans K catégories mutuellement exclusives :

$$p_o^{\text{COHEN}} = \frac{1}{N} \sum_{k \in K} n_k^{A\&B} \qquad p_e^{\text{COHEN}} = \frac{1}{N^2} \sum_{k \in K} \left(n_k^A \times n_k^B \right)$$

avec n_k^i le nombre d'éléments classés dans la catégorie k par l'annotateur i.

Le kappa de FLEISS (1971) est une généralisation du kappa de COHEN permettant de prendre en compte M annotateurs·trices :

$$p_o^{\text{FLEISS}} = \frac{1}{N \times M(M-1)} \left(\sum_{i=1}^{N} \sum_{k=1}^{K} \left(n_k^i \right)^2 - N \times M \right)$$

$$p_e^{\text{FLEISS}} = \sum_{k=1}^{K} \left(\frac{\sum_{i=1}^{N} n_k^i}{N \times M} \right)^2$$

avec n_k^i le nombre de fois où l'élément i est attribué à la catégorie k.

Nous avons opté pour la variante de BHOWMICK, MITRA et BASU (2008), qui est spécifiquement conçue pour mesurer l'accord quand les éléments peuvent appartenir à plusieurs classes[6]. Pour permettre aux éléments d'appartenir à plusieurs catégories, l'accord est calculé entre les paires possibles de catégories :

$$p_o^{\text{BHOWMICK}} = \frac{1}{N \times K(K-1) \times M(M-1)} \sum_{i=1}^{N} \sum_{pk \in S} n_{pk}^i$$

$$p_e^{\text{BHOWMICK}} = \frac{1}{|S|} \sum_{pk \in S} \sum_{g \in G} \frac{1}{|W|} \sum_{(x,y) \in W} \frac{pk_g^x \times pk_g^y}{N^2}$$

avec S l'ensemble des paires de catégories, W l'ensemble des paires d'annotateurs·trices, n_{pk}^i le nombre de fois où l'élément i est attribué à la paire de catégories pk et pk_g^x le nombre d'éléments assignés à pk par l'annotateur x avec la combinaison d'affectation g. En effet, étant donné que l'accord porte sur des paires de catégories, nous avons quatre types d'affectations possible pour une paire $pk = (k_1, k_2)$, représenté chacun par une combinaison d'affectation g :

— l'élément n'est assigné ni à k_1 ni à k_2 : $g = (0, 0)$;

— l'élément est assigné uniquement à k_1 : $g = (1, 0)$;

— l'élément est assigné uniquement à k_2 : $g = (0, 1)$;

— l'élément est assigné à la fois à k_1 et k_2 : $g = (1, 1)$.

Nous notons G l'ensemble des combinaisons d'affectations. Cela permet de prendre en compte les accords inter-annotateurs positifs – les annotateurs x et y sont d'accord que l'élément i appartient aux catégories k_1 et k_2 – aussi bien que les accords négatifs – les annotateurs x et y sont d'accord que l'élément i n'appartient pas aux catégories k_1 et k_2.

Pour mesurer la qualité de nos annotations d'idéologie politique, nous avons aléatoirement sélectionné 1 000 profils destinés à être annotés par trois annotateurs·trices chacun. Pour ces 1 000 profils, les annotations ont abouti à un accord observé $P_o = 0{,}89$, un accord attendu $P_e = 0{,}57$ et un accord final mesuré $A_m = 0{,}75$, indiquant un consensus très important entre les annotateurs·trices

6. Notre implémentation Python pour cette mesure est disponible à l'adresse suivante : https://github.com/SyrupType/bhowmick_agreement_measure

selon la grille d'interprétation de LANDIS et KOCH (1977, p. 165), un accord étant considéré comme « presque parfait » à partir de 0,81. Compte tenu de cet accord et de la taille de notre ensemble de données, nous avons choisi d'annoter le reste du corpus ($N = 22\,169$) avec un seul annotateur par profil. Les annotations finales des 1 000 profils annotés trois fois ont été obtenues par vote majoritaire – nous n'avons par chance pas eu de profils avec trois annotations différentes à départager, cas pour lequel nous aurions eu besoin de faire appel à une quatrième personne.

4.5 PRÉSENTATION DU DATASET

4.5.1 *Statistiques globales*

Le Tableau 4.2 présente le nombre de profils par parti, ainsi que le volume de tweets et retweets que ces derniers ont publié. La Figure 4.3 compare, quant à elle, la proportion de profils, tweets et retweets représentée par chaque parti. En matière de profils et de tweets, les deux partis les plus représentés sont `FI` et `LR`, alors que le `PS` est largement distancé par les autres partis. Pour les retweets en revanche, `LR` représente environ un tiers des publications, et le `FN` est presque aussi présent que `FI`, suggérant que les profils `LR` et `FN` étaient particulièrement actifs. En effet, ils ont produit respectivement 28% et 20% des retweets alors qu'ils ne représentent que 19% et 15% du nombre total de profils.

Cette observation est confirmée par le Tableau 4.3, avec un nombre médian de retweets par profil de 66 pour `LR` et de 98 pour le `FN`. De façon intéressante, les 77 retweets médian du `PS` suggèrent des profils qui seraient également très actifs, mais cela ne transparaît pas au niveau du volume total de publications. Ce tableau confirme également que les profils publiant un très grand nombre de tweets sont rares, alors que des profils produisant un grand nombre de retweets sont bien plus fréquents : le nombre médian de tweets par profils varie entre 15 et 29, alors que le nombre médian de retweets se trouve entre 26 et 98.

La Figure 4.4 compare la distribution des différentes natures de profils en fonction du parti. Au niveau des profils individuels, il n'existe pas de différence notable entre partis, avec une sur-représentation masculine notable, à l'exception du `PS` qui compte presque autant de profils gérés par des femmes que par des hommes. En revanche, `EM`, `LR`, et le `PS` ont bien plus de profils non individuels, gérés par des groupes de militant·e·s ou des associations. C'est un élément intéressant car, alors que pour `LR` et le `PS` cela est probablement dû à l'ancienneté des partis, ayant donné le temps aux groupes d'activistes d'investir Twitter, `EM` a été fondé quelques mois avant la campagne. Cela pourrait donc représenter un effort actif de la part du parti pour être organisé et présent sur la plateforme.

TABLEAU 4.2 – Nombre de profils par parti politique.

PARTI	INDIVIDUELS (# professionels des médias)						NON INDIVIDUELS		TOTAL	# TWEETS	# RETWEETS
	Homme		Femme		Autre / Ind.		Politique	Autre			
FI	2696	(12)	1093	(1)	1023	(2)	298	3	**5 113**	528 401	1 664 144
PS	697	(1)	610	(2)	322		200	3	**1 832**	162 368	533 468
EM	2056	(2)	766	(8)	473	(3)	654	13	**3 962**	379 223	1 142 000
LR	2224	(3)	922	(2)	606	(1)	604	10	**4 366**	544 259	2 136 897
FN	1878	(4)	567	(2)	812	(1)	114	5	**3 376**	330 614	1 525 264
FI / PS	91	(1)	67		68		2		**228**	20 389	29 957
FI / EM	19		9		5				**33**	2 987	6 180
FI / LR	3				1				**4**	426	323
FI / FN	14		2		6				**22**	3 804	4 836
PS / EM	84		43		22		2		**151**	21 557	70 449
PS / LR					2				**2**	353	1 309
PS / FN	2				1				**3**	1 760	9 410
EM / LR	86		34		18		9	1	**148**	24 504	37 303
EM / FN	1								**1**	1 070	1 459
LR / FN	113	(1)	37		56		4	1	**211**	24 186	113 995
Ind.	1428	(198)	700	(92)	1049	(28)	35	189	**3 401**	368 683	486 937
TOTAL	**11 392**	**(222)**	**4 850**	**(107)**	**4 464**	**(35)**	**1 922**	**225**	**22 853**	**2 414 584**	**7 763 931**
Hors sujet									316		

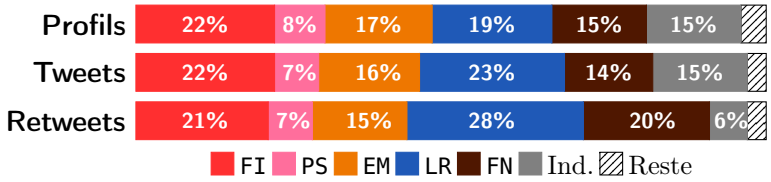


FIGURE 4.3 – Part de profils, tweets et retweets par parti.

TABLEAU 4.3 – Nombre médian et moyen de publications par profil et par parti.

		FI	PS	EM	LR	FN	Ind.
TWEETS	Médiane	23	25	25	28	29	15
	Moyenne	107	93	99	133	106	118
RETWEETS	Médiane	46	77	54	66	98	26
	Moyenne	330	293	292	496	457	142

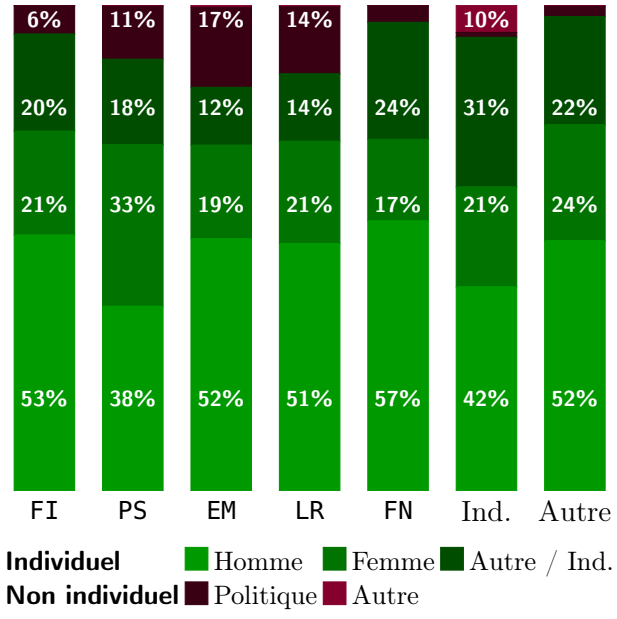


FIGURE 4.4 – Distribution des natures de profils par parti.

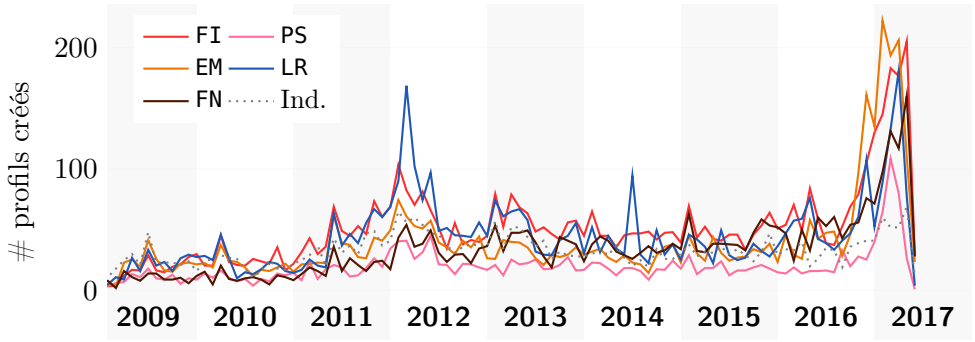


FIGURE 4.5 – Nombre de profils créés par mois. Les 209 profils créés avant 2009 n'apparaissent pas sur cette visualisation.

4.5.2 *Aspects temporels*

4.5.2.1 *Ancienneté des profils*

Afin de déterminer si les profils de notre jeu de données avaient été créés spécifiquement pour la campagne présidentielle ou s'ils étaient déjà présents sur Twitter, nous avons observé leur mois de création. Les résultats sont présentés dans la Figure 4.5 – par souci de lisibilité nous n'avons représenté que les cinq partis principaux et le point de vue *indéterminé*. Une large proportion de profils semblent avoir été créés juste avant ou pendant la campagne : 19 % sont apparus après septembre 2016. Cela pourrait indiquer que la majorité de ces profils sont dédiés à l'expression politique. Outre ces créations récentes, le seul phénomène notable est un pic de création au début de l'année 2012, probablement lié à l'élection présidentielle précédente, dont les tours ont pris place le 22 avril et le 6 mai 2012. Les tests de Kolmogorov-Smirnoff et de Mann-Whitney – permettant de détecter si des échantillons suivent une même loi statistique – n'indiquent pas de différence significative entre les différentes distributions de création de profils par parti (avec un risque α à 5 %).

4.5.2.2 *Chronologie des publications*

La Figure 4.6 montre l'évolution du nombre de tweets et de retweets publiés par les profils de notre jeu de données ainsi que certaines dates importantes de la campagne. Afin d'avoir une meilleure lisibilité, cette visualisation est concentrée sur les cinq partis principaux, les publications des profils affiliés à plusieurs partis étant également réparties entre ces derniers. Les pics d'activité les plus importants sont corrélés avec les éléments clés de la campagne, particulièrement les débats et les deux tours de l'élection. Le schéma global semble similaire pour tous les partis : après le second tour de la primaire de la droite et du centre, le nombre de publications diminue jusqu'en janvier 2017, où l'activité augmente à nouveau pour atteindre son maximum juste avant le premier tour de l'élection présidentielle. L'exception notable est le `FN`, dont la

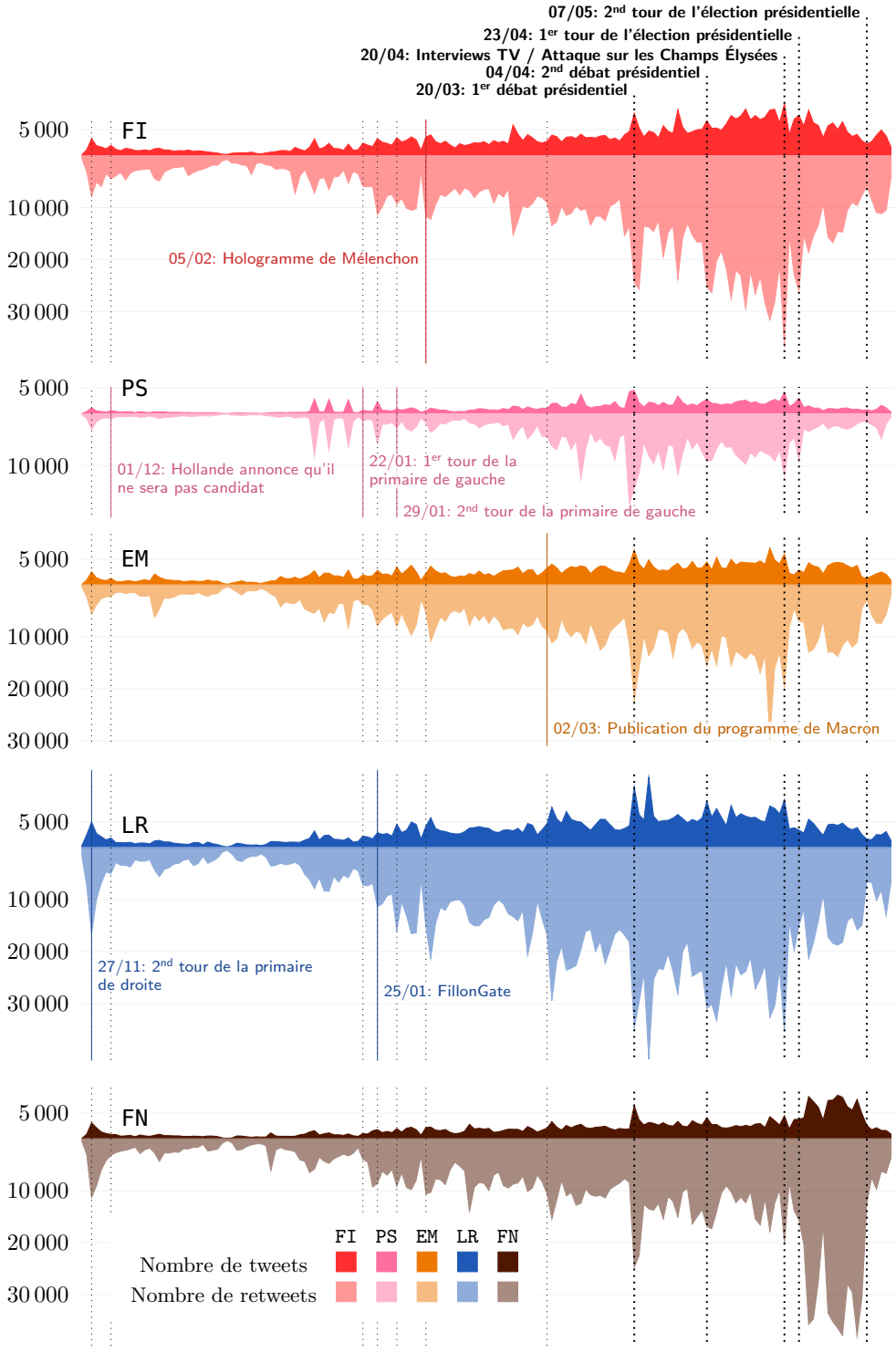


FIGURE 4.6 – Évolution du nombre de tweets et de retweets publiés par les profils annotés entre le 25 novembre 2016 et le 12 mai 2017.

TABLEAU 4.4 – Pays présents dans #Élysée2017fr par parti.
Seuls les pays indiqués par 20 profils ou plus sont détaillés ici, la liste complète est disponible dans le Tableau B.2.2.

PAYS	FI	PS	EM	LR	FN	Reste
France	2 689	1 090	2 704	2 848	1 723	1 899
États-Unis	26	8	35	21	69	35
Royaume-Uni	30	6	40	18	22	25
Belgique	23	1	15	19	14	27
Canada	16	4	15	10	13	13
Espagne	12	1	13	6	13	25
Suisse	11	4	10	10	6	10
Italie	2	1	10	6	13	13
Maroc	6	4	4	1	2	23
Allemagne	8	1	6	3	3	8
Grèce	2	1	2	1	1	15
Autre	61	17	57	49	53	117
TOTAL	2 886	1 138	2 911	2 992	1 932	2 210
À L'ÉTRANGER	**7%**	**4%**	**7%**	**5%**	**11%**	**14%**

majorité de l'activité est concentrée entre la qualification de Marine Le Pen, au premier tour de l'élection présidentielle, et le second tour. En effet, cette semaine représente 26 % du nombre total de tweets du FN, comparé à 15 % pour FI, 9 % pour le PS, 12 % pour EM et 10 % pour LR. La différence est encore plus notable pour les retweets, avec 27 % des retweets du FN ayant eu lieu durant cette semaine là, contre 12 % pour FI, 7 % pour le PS, 12 % pour EM et 10 % pour LR.

4.5.3 *Aspects géographiques*

#Élysée2017fr n'exclut pas les tweets n'étant pas rédigés en français ou les profils n'étant pas localisés en France, afin de prendre en compte les expatriés, les Français tweetant dans une autre langue et les étrangers discutant de l'élection présidentielle. Le Tableau 4.4 présente les pays les plus représentés dans le jeu de donnés, en fonction de la localisation indiquée par les utilisateurs·trices dans les profils Twitter. Nous avons exclu les 192 profils dont la localisation a changé durant la campagne, puis inféré manuellement les pays lorsque c'était possible. Cela nous a permis de déterminer un pays pour 86 % des 16 396 profils dont la localisation n'était pas vide. Le Tableau 4.5 présente les langues les plus utilisées par les profils, d'après la détection automatique de langue de Twitter [7].

7. http://support.gnip.com/apis/powertrack2.0/rules.html#Operators

TABLEAU 4.5 – Langues principales et secondaires utilisées par les profils d'#Élysée2017fr.
Seuls les cinq langues les plus utilisées sont présentées ici, les listes complètes sont disponibles dans le Tableau B.2.3 et le Tableau B.2.3.

	LANGUE SECONDAIRE					
LANGUE PRINCIPALE	Français	Anglais	Espagnol	Italien	Grec	TOTAL
Français	9 376	3 612	506	204	7	**13 705**
Anglais	161	264	7	4	1	**437**
Espagnol	12	5	46	1		**64**
Italien	4	5		18		**27**
Grec	3	1			9	**13**
TOTAL	**9 556**	**3 887**	**559**	**227**	**17**	**14 266**

Les langues principales et secondaires sont déterminées pour chaque profil en fonction du nombre de tweets publiés : il s'agit des deux langues les plus utilisées par le profil lors de ses publications. Pour cette analyse, nous n'avons pas considéré les retweets et les tweets dont la langue était indéterminée. Nous avons également écarté les tweets catégorisés par Twitter comme étant écrits en indonésien, haïtien ou tagalog, car une étude manuelle de ces derniers a rapidement révélé qu'ils avaient été mal catégorisés, probablement dû à leur petite taille et à la présence d'un grand nombre de mentions, d'hashtags et d'urls.

Ces éléments nous permettent de dire que, bien que l'élection présidentielle soit un événement national, elle a également été largement commentée à l'étranger, notamment dans les pays européens voisins. De façon intéressante, 11 % des profils affiliés au `FN` indiquent un pays étranger, soit presque le double comparé aux autres partis. Cette observation est confirmée en observant plus en détails les partis associés aux profils dont la langue principale n'est pas le français, comme présenté dans le Tableau 4.6. Le nombre de profils `FN` dont la langue principale n'est pas le français surpasse largement les autres partis. De plus, alors que pour les autres partis la seule langue étrangère largement représentée est l'anglais, les profils du `FN` sont bien plus variés, avec un nombre non négligeable de profils parlant espagnol ou italien. Cela peut en partie être expliqué par la campagne de désinformation qui a visé Emmanuel Macron à la fin de la campagne, communément appelée les *MacronLeaks*. En effet, les travaux de FERRARA (2017) ont montré qu'une partie de l'activité liée aux MacronLeaks était causée par des bots politiques anglophones qui avait été précédemment repérés lors de la campagne présidentielle 2016 aux États-Unis.

TABLEAU 4.6 – Partis d'affiliation des profils dont la langue principale n'est pas le français.
Seules les langues utilisées par au moins cinq profils sont détaillées ici, la liste complète est disponible dans le Tableau B.2.5.

LANGUE PRINCIPALE	FI	PS	EM	LR	FN	TOTAL
Anglais	50	9	94	10	161	**324**
Espagnol	8	1	7	5	17	**38**
Italien	1		3	1	19	**24**
Allemand	1		2		5	**8**
Portugais	1		1	1	5	**8**
Grec	3		1		1	**5**
Néerlandais			1		4	**5**
Polonais				1	4	**5**
Autre	1	2	2	1	7	**28**
TOTAL	**66**	**13**	**112**	**19**	**223**	**433**

TABLEAU 4.7 – Caractéristiques globales des réseaux de retweets et de mentions.

	RETWEETS	MENTIONS
Nombre de nœuds	22 048	22 569
Nombre d'arêtes	1 321 948	1 896 262
Densité	0,003	0,004
Diamètre	333	122

TABLEAU 4.8 – Détails sur les nœuds et les arêtes des réseaux de retweets et de mentions.

		DEGRÉ DES NŒUDS		POIDS DES ARÊTES
		Sortant	Entrant	
RETWEETS	Min	1	1	1
	Médiane	14	16	1
	Moyenne	86	61	3
	Max	7 183	1 783	7 983
MENTIONS	Min	1	1	1
	Médiane	26	16	1
	Moyenne	84	107	4
	Max	2 168	14 469	10 473

TABLEAU 4.9 – Assortativité des réseaux de retweets et de mentions. Les autres réseaux sont présentés ici pour comparaison et sont tirés des travaux de NEWMAN (2002, 2003).

RÉSEAU	CRITÈRE	ASSORTATIVITÉ
Co-auteur·e·s en mathématiques	Degré	0,12
Co-auteur·e·s en biologie	Degré	0,13
Mentions #Élysée2017fr	Parti politique	**0,14**
Retweets #Élysée2017fr	Parti politique	**0,17**
Collaborations entre comédien·ne·s	Degré	0,21
Partenaires sexuel·le·s	Origine ethnique	0,25 [a]
Collaborations entre chefs d'entreprise	Degré	0,28
Co-auteur·e·s en physique	Degré	0,36

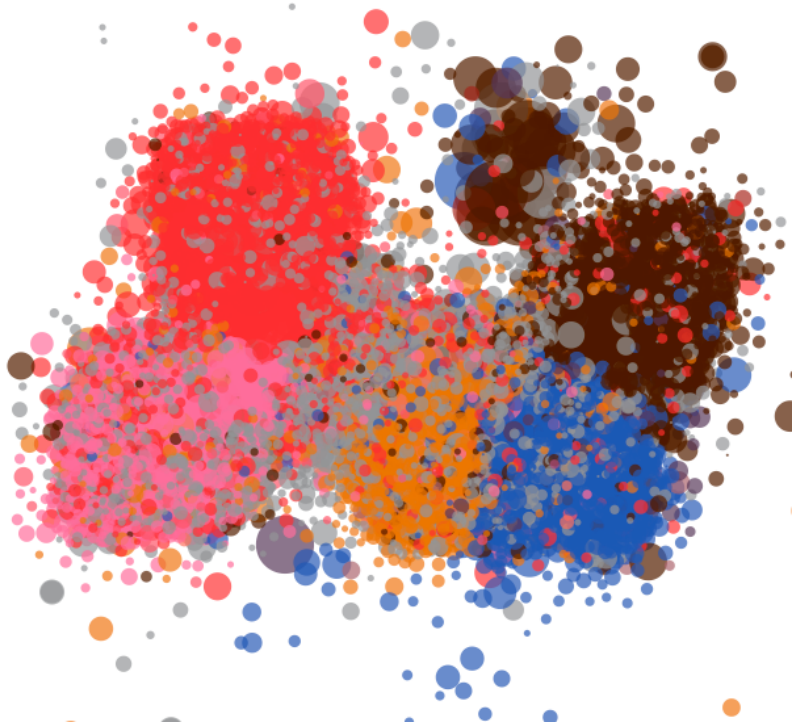
a. Nous présentons ici la moyenne des coefficients individuels d'assortativité présenté dans (NEWMAN, 2003, Tableau 1) par souci de lisibilité.

4.5.4 *Interactions sociales*

4.5.4.1 *Réseaux de retweets et de mentions*

En plus des annotations, #Élysée2017fr contient également les réseaux de retweets et de mentions liant les profils, éléments qui sont rarement fournis dans ce type de jeu de données. Ces réseaux sont créés à partir de 3 203 187 retweets et 6 371 852 mentions. Il s'agit de réseaux dirigés suivant le sens de l'information : des profils retweetés (ie. ayant créés les tweets initiaux) aux profils retweetant et des profils mentionnant aux profils mentionnés. Nous restons de cette façon fidèle à la temporalité de création de l'information, les retweets n'étant possibles et les mentions visibles qu'après création du tweet initial. Le poids d'une arête représente le nombre d'interactions liant la source à la cible. Le Tableau 4.7 résume les caractéristiques principales de ces réseaux et le Tableau 4.8 décrit de façon plus détaillée les nœuds et les arêtes. La Figure 4.7 permet de les visualiser avec une disposition calculée par l'algorithme OpenOrd (MARTIN et al., 2011). Le profil médian du jeu de données tend à retweeter un peu plus qu'il n'est retweeté, et mentionne d'autres profils de façon bien plus importante qu'il n'est mentionné. Le réseau de mentions est plus dense que le réseau de retweets et est moins divisé par parti. Cette observation est confirmée par l'assortativité mesurée sur les réseaux (voir Tableau 4.9). L'assortativité mesure la préférence des nœuds à s'attacher à des nœuds similaires, en comparant la connectivité réelle des nœuds avec une connectivité attendue aléatoire (NEWMAN, 2003). Les réseaux de retweets et de mentions #Élysée2017fr sont proches en matière d'assortativité de réseaux de co-auteur·e·s – chercheur·se·s apparaissant conjointement sur une publication scientifique – et de comédien·ne·s, montrant un attachement préférentiel entre nœuds affiliés à un même parti politique notablement supérieur à un attachement aléatoire.

(a) Réseau des retweets.



(b) Réseau des mentions.

FIGURE 4.7 – Réseau des retweets et des mentions (visualisés avec Gephi). La taille des nœuds est proportionnelle à leur degré sortant et les couleurs sont identiques à celles de la Figure 4.3, avec des couleurs intermédiaires pour les profils ayant plusieurs affiliations. Les arcs ne sont pas représentés pour des raisons de lisibilité.

TABLEAU 4.10 – Nombre moyen de retweets / mentions par profil en fonction des affiliations politiques des profils impliqués.

(a) Retweets

PROFIL RETWEETÉ	PROFIL RETWEETANT FI	PS	EM	LR	FN
FI	143	2	1	1	1
PS	2	82	4		
EM	3	9	122	3	1
LR	3	1	3	204	20
FN	1	1	1	18	194

(b) Mentions

PROFIL MENTIONNANT	PROFIL MENTIONNÉ FI	PS	EM	LR	FN
FI	249	24	16	11	6
PS	14	240	21	7	3
EM	8	13	232	22	7
LR	4	4	23	382	29
FN	6	4	16	39	301

4.5.4.2 *Interactions entre partis*

Mesurer les interactions entre profils en prenant en compte leurs affiliations politiques nous donne des informations intéressantes quant aux dynamiques existantes entre les partis. Le Tableau 4.10 présente les nombres moyens de retweets et de mentions agrégés par parti. Les résultats sont concordants avec la Figure 4.7a et la Figure 4.7b. Sans surprise, l'immense majorité des échanges a lieu au sein d'une même communauté politique. Les retweets suggèrent une proximité certaine entre LR et le FN, ainsi qu'entre le PS et EM, bien que celle-ci soit moins prononcée. Les mentions sont un peu moins partisanes. Presque chaque parti mentionne occasionnellement des profils EM et LR, alors que les profils de FI sont principalement mentionnés par ceux du PS, ceux du FN par des profils LR, et ceux du PS par des profils FI ou EM.

4.6 CONSIDÉRATIONS ÉTHIQUES

Au vu du sujet traité, il nous paraît pertinent d'aborder les questionnements éthiques que peut soulever un jeu de données de cette nature. Les points de vue étant des états subjectifs, il s'agit bien évidemment de sujets délicats à traiter, d'autant plus lorsqu'ils portent sur des problématiques sensibles telles que les valeurs politiques. Il s'agit néanmoins de matériaux de grande valeur

pour les chercheurs travaillant sur ces thématiques, particulièrement à l'heure actuelle où les médias sociaux ont une influence de plus en plus marquée dans les processus démocratiques – comme nous l'avons vu dans la Section 4.5.3 avec la présence de bots politiques. Nous avons donc œuvré à être les plus responsables possible dans le traitement de ces données.

Les tweets présents dans ce jeu de données ont été collectés via l'API Twitter afin de respecter les paramétrages de confidentialité des profils en ne collectant que les données explicitement publiques. De plus, ils n'ont été partagés que via leurs identifiants, afin de respecter le droit à l'oubli de leur auteur·e si ce dernier ou cette dernière a décidé de les supprimer depuis la collecte. Nous avons également voulu partager ce jeu de données via une communauté reconnue pour être sensible aux impacts sociaux des recherches en informatique. En effet, la communauté ICWSM [8] est hautement interdisciplinaire – mêlant informaticiens·nes, linguistes, sociologues, psychologues, etc. – permettant une grande réflexivité sur les sujets traités et donc une utilisation pertinente de ces données.

De plus, ce jeu de données a vocation à réaliser des analyses macroscopiques. En effet, comme l'énonce Liu (2012) :

> « Contrairement aux informations factuelles, les opinions et les sentiments [...] sont subjectifs. Il est donc important d'examiner une collection d'opinions provenant de nombreuses personnes plutôt que l'opinion unique d'une seule personne car une telle opinion ne représente que la vue subjective de cette personne. »

Bien que nos intentions ne soient en rien une garantie, nous avons choisi de faire confiance à nos consœurs et confrères pour utiliser à leur tour de façon responsable ces données et, à travers leurs travaux, permettre une meilleure compréhension et potentiellement une meilleure protection contre les utilisations frauduleuses de données personnelles.

4.7 CONCLUSION

Nous avons présenté dans ce chapitre le jeu de données original #Élysée2017fr, un jeu de données de points de vue politiques permettant d'analyser la campagne présidentielle française de 2017. Les analyses préliminaires ont permis de comparer la participation des partis au cours de la campagne, en termes de volume de tweets, de retweets, ou de proportion de comptes d'activistes. De plus, elles ont permis de constater l'intérêt international pour cet événement, ainsi que la présence de nombreuses communautés étrangères d'extrême-droite. Ce jeu de données nous a également permis d'évaluer nos modèles de détection de points de vue, présentés dans le chapitre suivant. L'Annexe B présente les détails techniques de ce jeu de données.

8. International Conference on Web and Social Media : http://www.icwsm.org

Deuxième partie

MODÈLES SEMI-SUPERVISÉS DE DÉTECTION DE POINTS DE VUE SUR LES MÉDIAS SOCIAUX

5

FONDATIONS DES MODÈLES

Nous pouvons à présent nous pencher sur la conception de modèles de détection de points de vue étant donné que nous disposons d'un jeu de données complexe permettant une évaluation approfondie (présenté dans le Chapitre 4). Nous présentons dans ce chapitre les fondations théoriques sur lesquelles s'appuient nos modèles, présentés dans le Chapitre 6 et le Chapitre 7. Nous énonçons tout d'abord les hypothèses sous-tendant notre travail avant de décrire les expérimentations réalisées pour les confirmer.

5.1 PRÉSENTATION DES HYPOTHÈSES

Notre travail repose sur deux hypothèses principales détaillées ci-après : en premier lieu la pertinence de l'utilisation de l'homophilie présente sur les médias sociaux, puis l'importance du fait de considérer plusieurs faisceaux d'indices lorsque ceux-ci sont disponibles.

5.1.1 *(H1) Utilisation de l'homophilie sur les médias sociaux*

Comme présenté dans la Section 2.2.2, plusieurs travaux ont montré que les médias sociaux numériques étaient fortement polarisés dans certains contextes, particulièrement sur les sujets politiques (Barberá et al., 2015 ; Iyengar et Westwood, 2015 ; McPherson, Smith-Lovin et Cook, 2001). Ils ont révélé la présence d'une forte *homophilie* sur plusieurs plateformes : les utilisateurs et utilisatrices préfèrent intéragir avec des personnes partageant leurs idées, allant parfois jusqu'à créer des *« chambres d'écho »* dans lesquelles leurs convictions ne sont jamais mises à mal (Sunstein, 2009). Les blogs politiques ont tendance à faire référence à des blogs de leur propre camp (Adamic et Glance, 2005), et de façon similaire, les réseaux de retweets et de mentions sur Twitter lors d'événements politiques sont habituellement hautement polarisés (Conover et al., 2011a ; Fraisier et al., 2017). À la vue de ces observations, nous pensons qu'il est possible de tirer parti de cette homophilie pour trouver des communautés d'utilisateurs·trices partageant un même point de vue. Notons tout de même que, comme le rappellent Boullier et Lohard (2012), les profils peuvent faire partie de plusieurs univers sociaux, et que les communautés ne seront donc pas omni-domaines mais dépendront du sujet étudié.

5.1.2 *(H2) Importance des faisceaux d'indices multiples*

Les profils de médias sociaux sont liés par une variété d'informations : des éléments de langage, des interactions sociales variées, des situations géographiques, etc. Si l'on essaie d'inférer leurs points de vue sur un sujet, nous comprenons intuitivement qu'utiliser plusieurs faisceaux d'indices est bénéfique : alors que l'un des profils pourrait massivement partager les publications d'une candidate, un autre pourrait être plus discret en matière de partages mais utiliser la même rhétorique dans ses publications que l'un des partis, ou être passif mais indiquer vivre dans une zone connue pour être en faveur d'un point de vue spécifique. Pouvoir exploiter plusieurs types de *proximités* inter-profils présente donc plusieurs avantages :

— la possibilité de prendre en compte des informations variées afin d'obtenir une prédiction plus fine ;

— la possibilité de détecter des points de vue quelle que soit la plateforme considérée.

5.2 FORMALISATION

Pour tirer parti au mieux des multiples proximités étudiées, nous les représentons sous forme de graphes de proximité et nous exploitons leur structure pour prédire les points de vue. Les notations utilisées sont résumées dans le Tableau 5.1.

Soit $V = \{v_1, v_2, \ldots\}$ un ensemble de profils, $\boldsymbol{\sigma}$ l'ensemble des points de vue distincts exprimés et $\boldsymbol{\Sigma}$ la matrice $|V| \times |\boldsymbol{\sigma}|$ notant les points de vue exprimés. $\boldsymbol{\Sigma}_{i,\bullet}$ représente les points de vue exprimés par le profil v_i, avec

$$\boldsymbol{\Sigma}_{i,j} = \begin{cases} 1 & \text{si } v_i \text{ exprime le point de vue } \boldsymbol{\sigma}_j \\ 0 & \text{sinon.} \end{cases}$$

Considérons $S \subset V$ l'ensemble de profils avec un point de vue connu, avec $|S| \ll |V|$. $A = \{A_1, \ldots, A_k\}$ est un ensemble d'attributs, avec A_i^j la valeur de l'attribut A_i pour le profil v_j. Nous définissons un multiplex $\mathcal{G}$. Chaque couche est définie comme un graphe $G_i = (V, \mathrm{Sim}_i)$ telle que Sim_i représente la proximité entre les profils contenus dans V en fonction de l'attribut A_i. Nous considérons les proximités comme étant symétriques, c'est-à-dire que $\mathrm{Sim}_i(v_i, v_j) = \mathrm{Sim}_i(v_j, v_i)$, par conséquent G_i est non dirigé.

FORMULATION DE LA TÂCHE. Étant donné $\mathcal{G} = \{G_i,\ i \in [\![1, k]\!]\}$ et S, est-il possible de déterminer avec exactitude $\boldsymbol{\Sigma}_{v_i,\bullet}$ avec $v_i \in V \setminus S$?

TABLEAU 5.1 – Notations utilisées.

	MODÈLE THÉORIQUE (Section 5.2)
V	Ensemble des profils
$\boldsymbol{\sigma}$	Ensemble des points de vue exprimés
$\boldsymbol{\Sigma}$	Matrice $\lvert V\rvert \times \lvert\boldsymbol{\sigma}\rvert$ notant les points de vue exprimés
S	Ensemble des profils dont les points de vue sont connus
A	Ensemble d'attributs permettant de calculer des similarités inter-profils
$\mathcal{G}$	Multiplex avec chaque couche $G_i = (V, \mathrm{Sim}_i)$ représentant une proximité calculée à partir de A_i
	MODÈLE SÉQUENTIEL COMMUNAUTAIRE (Chapitre 6)
V_T	Ensemble des profils servant à l'évaluation
X	Liste des proximités utilisées
$\boldsymbol{\omega}$	Fonction d'ordonnancement des proximités
X^{ord}	Séquence *ordonnée* des proximités utilisées
$\boldsymbol{\varphi}$	Fonction d'importance des profils-graines
s	Nombre de profils-graines à sélectionner
s_{com}	Nombre minimal de communautés-graines à considérer
s_{min}	Nombre minimal de profils-graines à sélectionner dans chaque communauté-graine
	MODÈLE DYNAMIQUE (Chapitre 7)
Π	Ensemble des périodes considérées
$\boldsymbol{\Sigma}^p$	Tenseur d'ordre 3 et de dimensions $\lvert V\rvert \times \lvert\boldsymbol{\sigma}\rvert \times \lvert\Pi\rvert$ notant les points de vue exprimés

TABLEAU 5.2 – Jeux de données utilisés lors des expérimentations. V est l'ensemble des profils présents dans le jeu de données et V_T les profils annotés utilisés pour l'évaluation.

JEU DE DONNÉES		POINTS DE VUE σ	$\lvert V_T \rvert$	$\lvert V \rvert$
[#E]	#Élysée2017fr	France Insoumise	5 113	22 843
		Parti Socialiste	1 832	
		En Marche !	3 962	
		Les Républicains	4 366	
		Front National	3 376	
		Total	*18 649*	
[SR]	Référendum sur l'indépendance écossaise	Pour	564	604 399
		Contre	537	
		Total	*1 101*	
[ME]	Élection de mi-mandat aux États-Unis	Démocrate	761	1 718 131
		Républicain	810	
		Total	*1 571*	
[PE]	Élection présidentielle aux États-Unis	Démocrate	481	2 042 025
		Républicain	427	
		Total	*918*	
[GC]	Contrôle strict des armes à feu	Pour	312	1 420
		Contre	489	
		Total	*801*	

5.3 JEUX DE DONNÉES UTILISÉS

En plus d'#Élysée2017fr, présenté dans le Chapitre 4, nous utilisons trois autres jeux de données évoqués dans la Section 2.4 pour étudier cette tâche : le référendum sur l'indépendance écossaise (BRIGADIR, GREENE et CUNNINGHAM, 2015), l'élection de mi-mandat aux États-Unis de 2014 (BRIGADIR, GREENE et CUNNINGHAM, 2015) et l'élection présidentielle aux États-Unis (GARIMELLA et al., 2017). Par souci de lisibilité, ces jeux de données seront par la suite abrégés `#E` pour #Élysée2017fr, `SR` pour le référendum écossais, `ME` pour l'élection de mi-mandat et `PE` pour l'élection présidentielle aux États-Unis par la suite. Les profils de `PE` ayant un point de vue connu ont été déterminés grâce aux profils officiels des candidat·e·s, à savoir Donald Trump, Mike Pence, Hillary Clinton et Bernie Sanders, et aux profils de `ME`, en considérant que les politicien·ne·s ayant changé de parti entre 2014 et 2016 étaient négligeables.

Un jeu de données provenant d'une autre plateforme a également été utilisé, portant sur le contrôle des armes à feu aux États-Unis (`GC`) (ABBOTT et al., 2016). Il est constitué de discussions provenant de CreateDebate.com, un site

de débats. Chaque discussion a deux points de vue possibles, sélectionnés par son auteur·e. Ces points de vue « locaux » ont été liés par ABBOTT et al. à deux points de vue globaux : *pour* ou *contre* un contrôle strict des armes à feu. Par exemple, dans une discussion nommée « Le droit à porter une arme, nécessaire ? », le point de vue local « Oui, pour nous protéger » a été mis en correspondance avec « *contre un contrôle strict* » et le point de vue « Non, cela ne fait qu'augmenter la criminalité » avec « *pour un contrôle strict* ». Lorsqu'un profil ajoute une publication dans une discussion, il doit indiquer son point de vue, et si sa publication soutient, clarifie ou conteste une autre publication. Chaque profil contient également des informations biographiques, une liste de relations, le nombre de débats auxquels le profil a pris part, son nombre de publications et son efficacité. Les relations sont des profils indiqués comme étant des alliés, des ennemis ou comme étant hostiles envers le profil, et l'efficacité mesure le pourcentage de votes positifs laissés par les autres membres de la plateforme sur les publications du profil. Dans nos expériences, nous n'avons pas considéré les 88 profils dont le point de vue global changeait au cours des débats.

Le Tableau 5.2 présente en détail le nombre de profils annotés par point de vue (V_T), sur lesquels nous basons nos évaluations, ainsi que le nombre total de profils présents dans les jeux de données (V). Les variations dans les nombres de profils par rapport au Tableau 2.5 sont dues aux profils supprimés depuis la collecte initiale et au fait que pour #Élysée2017fr nous n'avons pas considéré les profils ayant des affiliations multiples et les profils n'ayant pas d'affiliation politique, afin de nous rapprocher des autres jeux de données à notre disposition.

5.4 PROXIMITÉS

De nombreuses proximités entre profils peuvent être utilisées, exploitant des types d'informations variées : textuelles, sociales, etc. Chaque proximité définit une mesure de similarité Sim_i et par extension le graphe G_i représentant les profils V et les arêtes E_i les reliant. Cette définition flexible permet de mettre en place un modèle générique pouvant être adapté à n'importe quelle plateforme sociale. De plus, chaque proximité étant utilisée de façon isolée, il n'est pas nécessaire de normaliser les poids des arêtes. Les proximités utilisées dans ce travail sont présentées ci-dessous et résumées dans le Tableau 5.3. Certaines valeurs observées pour avg_w et $\max_w$ semblent surprenantes au premier abord. Après une étude plus poussée, il semblerait qu'elles soient dues à un petit nombre de profils particulièrement actifs et utilisant les mêmes éléments de langage, probablement grâce à une organisation décidée en amont [1].

1. Pour le référendum écossais, on retrouve notamment plusieurs des profils connus en faveur de l'indépendance comme @YesScotland, @WeAreNational ou @_WeAreScotland_.

5.4.1 *Proximités fondées sur le contenu textuel*

Ce type de proximités lie les profils en utilisant des éléments textuels communs dans leurs publications.

Utilisation de mots-clés (kw)

$$\mathrm{Sim}_{kw}(v_i, v_j) = \left| A^i_{kw} \cap A^j_{kw} \right|$$

avec A^i_{kw} les hashtags utilisés dans les publications de v_i pour les jeux de données Twitter (`#E`, `SR`, `ME`, `PE`), ou les substantifs pour le jeu de données CreateDebate (`GC`) étant donné que cette plateforme ne propose pas de système de mots-clés à ses utilisateurs·trices.

Référence à une autre information (ref)

$$\mathrm{Sim}_{ref}(v_i, v_j) = \left| A^i_{ref} \cap A^j_{ref} \right|$$

avec A^i_{ref} les sites web mentionnés par v_i. Les URL provenant d'un service de raccourcissement d'URL, tels que `bit.ly` ou `goo.gl`, ont été étendues lorsque possible pour une correspondance optimale. Pour CreateDebate, seuls les noms de domaines ont été considérés à cause de la faible taille du jeu de données : seules trois URL complètes étaient partagées par plusieurs profils, les autres étant utilisées une seule fois.

Nous sommes conscients que nos mesures de similarités textuelles sont limitées et que des mesures bien plus complètes existent, ne serait-ce que les classiques coefficient de Jaccard (Jaccard, 1901) ou similarité cosinus (Singhal, 2001). Les proximités présentées ici ont cependant l'avantage d'être très peu coûteuses à calculer et d'être indépendantes de la langue, renforçant la généricité du modèle. En effet, la plupart des similarités textuelles avancées reposent sur des outils de traitement de la langue nécessitant un paramétrage spécifique pour la collection considérée.

5.4.2 *Proximités fondées sur les interactions sociales*

Ces proximités sont établies sur le contexte social et reposent sur le nombre d'interactions sociales entre les profils. Nous définissons $A^i_{soc} = \left\{s^i_k\right\}$ comme étant l'ensemble des interactions $s^i_k = (v_k, t_k)$, avec *soc* l'interaction sociale considérée, v_k un profil ayant interagi avec v_i et t_k la date de l'interaction. Afin de prendre en compte le fait que les logiques d'utilisation varient grandement selon la nature symétrique ou asymétrique des relations (Colleoni, Rozza et Arvidsson, 2014), nous avons considéré trois versions pour ces proximités. La première (*all*) considère toutes les interactions, alors que la seconde (*rec*)

se concentre sur les interactions réciproques et la troisième ($\overline{rec}$) sur les interactions non-réciproques. Leurs définitions formelles sont données par les équations 1, 2 et 3 respectivement. Par définition, $A_{soc_{all}} = A_{soc_{rec}} \cup A_{soc_{\overline{rec}}}$.

$$\mathrm{Sim}_{soc_{all}}(v_i, v_j) = \left|\left\{s_k^i \mid v_k = v_j\right\}\right| + \left|\left\{s_l^j \mid v_l = v_i\right\}\right| \tag{1}$$

$$\mathrm{Sim}_{soc_{rec}}(v_i, v_j) = \min\left(\left|\left\{s_k^i \mid v_k = v_j\right\}\right|, \left|\left\{s_l^j \mid v_l = v_i\right\}\right|\right) \tag{2}$$

$$\mathrm{Sim}_{soc_{\overline{rec}}}(v_i, v_j) = \left|\left|\left\{s_k^i \mid v_k = v_j\right\}\right| - \left|\left\{s_l^j \mid v_l = v_i\right\}\right|\right| \tag{3}$$

Nous avons considéré les proximités présentée ci-après.

Citation (cite)

Cette proximité représente les partages de la publication (tout ou partie) d'un autre profil, via les *retweets* pour Twitter et les *citations* pour CreateDebate.

Interpellation (call)

Cette proximité capture le fait qu'un profil en interpelle un autre, à savoir les *mentions* pour Twitter, et les publications *soutenant ou clarifiant* explicitement une autre publication pour CreateDebate.

Association (asso)

Cette proximité représente les profils explicitement suivis, c'est-à-dire les *amis* sur Twitter et les *alliés* sur CreateDebate. Seules les associations des profils dans V_T ont été collectés afin de conserver une taille de graphe exploitable sur une machine standard.

Les profils de CreateDebate contenaient des informations supplémentaires qui ont été également utilisées pour mesurer la similarité sociale entre profils.

Critères socio-démographiques (socio)

$$\mathrm{Sim}_{socio}(v_i, v_j) = \left|A_{sexe}^i \cap A_{sexe}^j\right| + \left|A_{age}^i \cap A_{age}^j\right| + \left|A_{ecole}^i \cap A_{ecole}^j\right|$$

avec *age* l'âge indiqué dans le profil arrondi à la décennie et *ecole* le niveau scolaire.

Croyances religieuses et politiques (beliefs)

$$\mathrm{Sim}_{beliefs}(v_i, v_j) = \left|A_{religion}^i \cap A_{religion}^j\right| + \left|A_{parti}^i \cap A_{parti}^j\right|$$

TABLEAU 5.3 – Nombre d'arêtes $|E_i|$ dans G_i.
avg_w indique le poids moyen et $\max_w$ le poids maximal.

		#E	SR	ME	PE	GC
ref	$\|E_i\|$	17 367 840	314 303	15 640	50 429 635	74
	avg_w	10	4	6	8	3
	$\max_w$	99 704	36 444	7 592	233 620 968	32
kw	$\|E_i\|$	78 440 288	422 737	1 304 290	144 891	236 976
	avg_w	80	5 142	24	12 402	49
	$\max_w$	29 006	5 319 272	222 379	559 657 926	29 854
$cite_{all}$	$\|E_i\|$	1 262 484	1 426 334	712 734	1 848 244	128
	avg_w	3	1	2	2	6
	$\max_w$	7 983	324	672	1 512	68
$cite_{rec}$	$\|E_i\|$	123 853	11 563	7 167	6 588	23
	avg_w	7	1	2	2	8
	$\max_w$	7 983	65	73	483	53
$call_{all}$	$\|E_i\|$	1 787 127	335 171	66 066	457 683	226
	avg_w	5	3	2	3	2
	$\max_w$	10 473	1 296	377	35 725	14
$call_{rec}$	$\|E_i\|$	227 400	21 682	2 640	5 080	48
	avg_w	8	2	1	4	2
	$\max_w$	8 658	197	55	2 657	11
$asso_{all}$	$\|E_i\|$	3 135 200	1 743 105	3 121 674	3 661 573	
	avg_w	1	1	1	1	*NA*
	$\max_w$	1	1	1	1	
$asso_{rec}$	$\|E_i\|$	1 532 825	1 582 919	2 648 403	3 324 920	4 034
	avg_w	1	1	1	1	1
	$\max_w$	1	1	1	1	1
$asso_{\overline{rec}}$	$\|E_i\|$	1 602 375	160 186	473 271	160 186	
	avg_w	1	1	1	1	*NA*
	$\max_w$	1	1	1	1	
socio	$\|E_i\|$					312 795
	avg_w	*NA*	*NA*	*NA*	*NA*	1
	$\max_w$					3
beliefs	$\|E_i\|$					240 323
	avg_w	*NA*	*NA*	*NA*	*NA*	1
	$\max_w$					2
city	$\|E_i\|$	2 929 622	11 833 595	2 023	2 838	
	avg_w	1	1	1	1	*NA*
	$\max_w$	1	1	1	1	
region	$\|E_i\|$	7 780 336	12 196 762	20 447	11 101	
	avg_w	1	1	1	1	*NA*
	$\max_w$	1	1	1	1	
country	$\|E_i\|$					445 407
	avg_w	*NA*	*NA*	*NA*	*NA*	1
	$\max_w$					1

5.4.3 *Proximités fondées sur le contexte géographique*

Ce type de proximité repose sur les localisations géographiques indiquées dans les profils. La construction est similaire pour les trois granularités étudiées : *city*, *region* et *country*.

$$\text{Sim}_{loc}(v_i, v_j) = \left| A^i_{loc} \cap A^j_{loc} \right|, \text{ avec } loc \in \{city, region, country\}$$

Étant donné que les situations sont indiquées librement par les utilisateurs·trices, une étape de nettoyage et d'extraction manuelle des lieux était nécessaire afin de prendre en compte les nombreuses variations de format. Afin de limiter le coût de cette étape, seules les situations géographiques des profils dans V_T ont été prises en compte.

5.5 VALIDATION DES HYPOTHÈSES

Nous présentons dans cette section les expérimentations réalisées afin de valider les hypothèses présentées dans la Section 5.1. Pour confirmer *(H1)* et prouver que nous pouvons exploiter l'homophilie, nous répondons aux questions suivantes :

QP1. Nos proximités ont-elles une structure communautaire bien définie ?

QP2. Les principales communautés détectées sont-elles homogènes pour ce qui est des points de vue, quelles que soient les proximités considérées ?

Enfin, pour valider *(H2)* et prouver la complémentarité des proximités, et donc l'importance de faisceaux d'indices multiples, nous examinons la question suivante :

QP3. Les communautés détectées varient-elles en fonction des proximités considérées ?

5.5.1 *Structure communautaire*

Pour répondre à **QP1**, nous avons mesuré la *transitivité* de nos graphes de proximité. La transitivité mesure la probabilité que deux voisins d'un nœud soient connectés, et plus cette probabilité est importante, plus le graphe possède une structure communautaire apparente (Orman, Labatut et Cherifi, 2013). Elle peut être exprimée comme la moyenne du coefficient de clustering pondéré (Barrat et al., 2004) défini comme :

$$\frac{1}{s_i(k_i+1)} \sum_{j,h} \frac{w_{ij}+w_{ih}}{2} e_{ij} \times e_{ih} \times e_{jh}$$

avec s_i la force du profil v_i, k_i son degré, w_{ij} le poids de l'arête connectant v_i à v_j et $e_{ij} = \begin{cases} 1 & \text{si } w_{ij} > 0 \\ 0 & \text{sinon.} \end{cases}$

TABLEAU 5.4 – Transitivité des graphes de proximité.

	#E	SR	ME	PE	GC
ref	0,69	0,61	0,44	0,74	0,59
kw	0,87	0,95	0,79	0,77	0,91
$cite_{all}$	0,38	0,02	0,06	0,05	0,03
$cite_{rec}$	0,24	0,09	0,12	0,03	0,07
$call_{all}$	0,48	0,15	0,11	0,16	0,05
$call_{rec}$	0,24	0,11	0,04	0,04	0,01
$asso_{all}$	0,41	0,16	0,12	0,13	
$asso_{rec}$	0,31	0,16	0,10	0,12	0,09
$asso_{\overline{rec}}$	0,26	0,03	0,03	0,03	
socio					0,80
beliefs					0,82
city	0,95	1,00	0,73	0,83	
region	0,99	1,00	1,00	0,99	
country					0,98

Le Tableau 5.4 présente la transitivité des proximités étudiées. La plupart d'entre elles ont une transitivité non négligeable, suggérant une structure communautaire sous-jacente plus ou moins prononcée selon le jeu de données. Malgré leur utilisation fréquente et reconnue dans les modèles de détection de points de vue, toutes les versions de *cite* présentent néanmoins une transitivité faible ou modérée, allant de 0,02 à 0,38. La transitivité des proximités sociales d'#Élysée2017fr (#E) est en moyenne supérieure de 0,25 à celle des autres jeux de données Twitter ; nous supposons que cela est principalement dû au fait que nous l'ayons collecté nous-mêmes. En effet les autres jeux de données ont été reconstitués plusieurs années après la collecte initiale, ce qui occasionne généralement une perte non-négligeable – mais malheureusement difficilement quantifiable – de données.

5.5.2 *Homogénéité des communautés*

Afin de répondre à **QP2**, nous avons mesuré l'homogénéité des communautés détectées pour chaque proximité.

5.5.2.1 *Choix de l'algorithme*

Au vu des divers algorithmes de détection de communautés non recouvrantes présentés dans la Section 3.2.1 (synthétisés dans le Tableau 3.1), une question se pose : quel algorithme utiliser pour notre tâche ? Afin de faire un choix éclairé, nous avons décidé de comparer les différents algorithmes à notre dispo-

sition. Le but étant d'étudier les différences de comportements des algorithmes sur un même graphe, nous n'avons pas effectué une comparaison exhaustive – qui au vu des cinq jeux de données, 14 proximités et six algorithmes impliqués, nous aurait donné 420 cas différents à étudier – mais nous avons plutôt sélectionné un échantillon de nos données, à savoir les proximités $cite_{all}$ et $call_{all}$ des jeux de données `SR` et `ME`. Nous avons sélectionné ces proximités car elles sont reconnues dans la littérature pour fournir des communautés fortement homogènes (Conover et al., 2011a ; Conover et al., 2011b ; Merry, 2016 ; Weber, Garimella et Batayneh, 2013), ce qui nous permettait d'obtenir un premier critère d'analyse des algorithmes : si l'un d'entre eux donnait de mauvais résultats sur ces proximités, nous pouvions le disqualifier. Pour Infomap, qui nécessite un graphe dirigé, nous avons considéré le sens de l'information sur Twitter, à savoir des profils retweetés vers les profils retweetant et des profils mentionnant vers les profils mentionnés[2]. `SR` et `ME` ont été choisis car il s'agissait des jeux de données les plus semblables, ayant la même technique et la même année de collecte initiale, et pouvaient donc nous donner des valeurs comparables. Nous avons mesuré la pureté intra-communautaire moyenne des communautés détectées (Girvan et Newman, 2002) pour évaluer leur homogénéité :

$$\text{Pureté}(x) = \frac{1}{|C_x|} \sum_{c_i \in C_x} \frac{\max_j \left| c_i \cap \mathbf{\Sigma}_{\bullet,j} \right|}{|c_i|}$$

avec C_x la liste des communautés contenant au moins cinq profils dans V_T pour la proximité x et $\left| c_i \cap \mathbf{\Sigma}_{\bullet,j} \right|$ le nombre de profils exprimant le point de vue $\boldsymbol{\sigma}_j$ présent dans la communauté c_i.

Les résultats sont présentés dans le Tableau 5.5. Dans l'ensemble, à l'exception de la détection via vecteurs propres qui obtient systématiquement des scores significativement inférieurs aux autres, les différents algorithmes semblent tous efficaces, avec une pureté moyenne de 84 %. La famille d'algorithmes fondés sur la diffusion de l'information semble présenter un léger avantage, nous allons donc nous orienter vers ceux-ci. Malgré ses bons résultats, Infomap ne peut rivaliser avec la vitesse d'exécution de la propagation de labels (moins de 2 secondes par graphe), ce qui était à attendre au vu de sa complexité. De plus, cet algorithme nécessite des arêtes dirigées, ce qui nous priverait de nombreuses proximités pour lesquelles une direction n'aurait pas de raison d'être. Nous avons donc utilisé pour la suite de nos expérimentations la propagation de labels pour détecter les communautés.

2. Nos graphes de proximités sont définis comme étant non dirigés mais nous incluons Infomap dans la comparaison afin de confirmer que le fait de considérer des similarités symétriques ne dégrade pas les résultats.

TABLEAU 5.5 – Pureté moyenne des communautés contenant au moins cinq profils dans V_T.
Les valeurs maximales sont en gras pour plus de lisibilité, et le nombre de communautés inclues dans le calcul est indiqué entre parenthèses.

		$cite_{all}$		$call_{all}$	
CATÉGORIE	ALGORITHME	SR	ME	SR	ME
Maximisation de la modularité	Fast-greedy	0,977 (10)	0,926 (37)	**0,859** (25)	0,630 (40)
	Vecteurs propres	0,873 (14)	0,620 (3)	0,513 (1)	0,575 (19)
	Multi-level	0,975 (14)	0,866 (37)	0,843 (29)	0,659 (45)
Diffusion de l'information	Infomap	0,934 (33)	**0,975** (48)	0,767 (21)	**0,773** (46)
	Propagation de labels	**0,981** (2)	0,910 (26)	0,534 (1)	0,762 (58)
Marche aléatoire	Walktrap	0,977 (10)	0,926 (37)	**0,859** (25)	0,630 (40)

TABLEAU 5.6 – Pureté moyenne des 10 plus grosses communautés contenant au moins cinq profils dans V_T.
Les valeurs supérieures à 0,80 sont en gras pour plus de lisibilité.

	#E	SR	ME	PE	GC
ref	0,68	**0,82**	**0,84**	**0,87**	0,76
kw	0,25	0,52	0,62	0,53	0,63
$cite_{all}$	**0,93**	**0,94**	**0,99**	**0,97**	0,67
$cite_{rec}$	**0,95**	**1,00**	**0,98**	**0,99**	0,60
$call_{all}$	**0,93**	0,52	0,78	**0,86**	0,66
$call_{rec}$	**0,85**	0,79	**0,96**	**0,96**	0,66
$asso_{all}$	**0,87**	**0,98**	**0,94**	**0,98**	
$asso_{rec}$	0,79	**0,99**	0,52	**0,97**	0,66
$asso_{\overline{rec}}$	**0,84**	0,75	0,52	0,53	
socio					0,64
beliefs					0,64
city	0,33	0,61	0,69	0,70	
region	0,30	0,61	0,58	0,61	
country					0,62

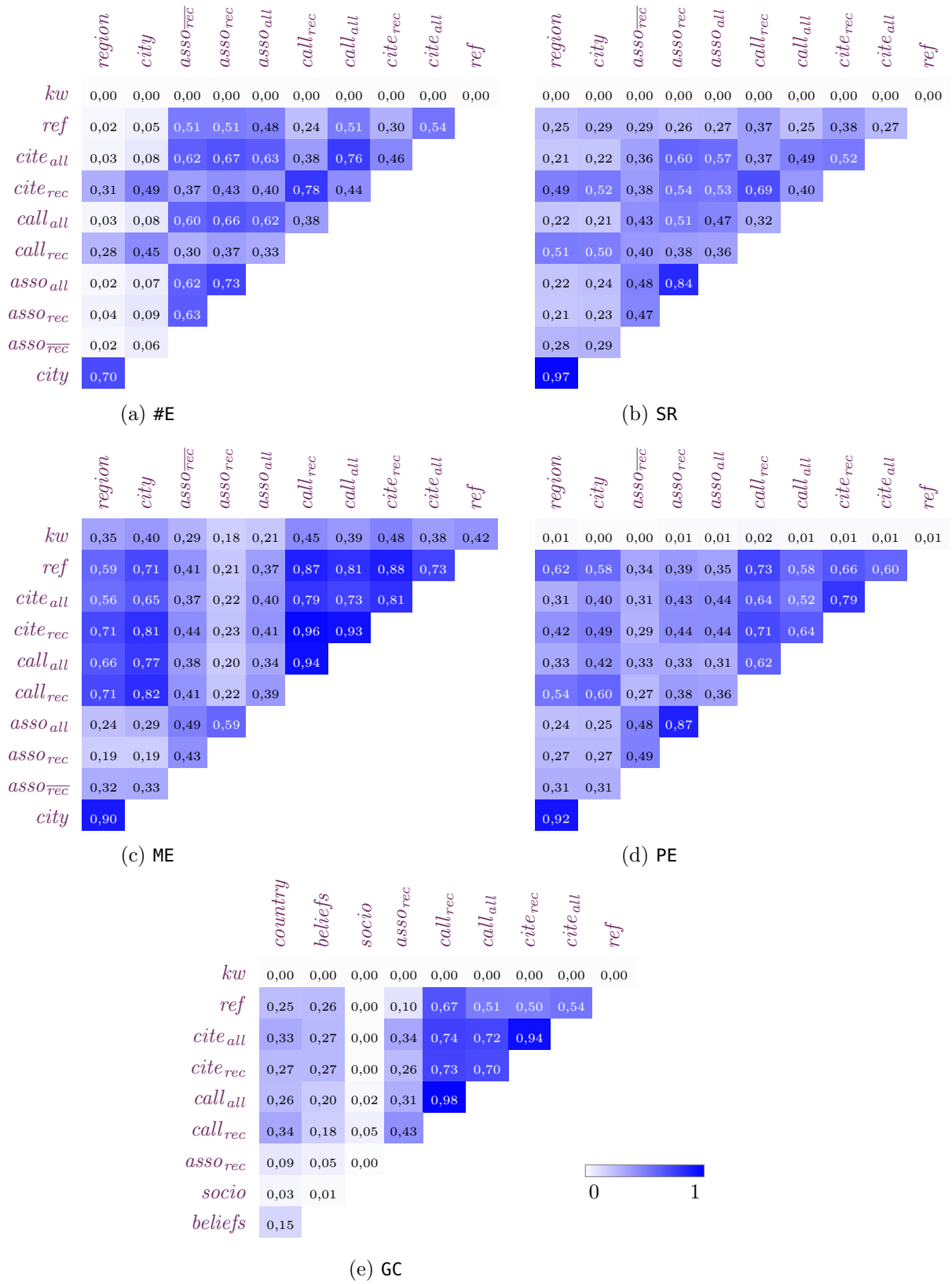


FIGURE 5.1 – Information mutuelle normalisée entre les communautés des proximités.

5.5.2.2 *Analyse des proximités*

Notre algorithme de détection de communautés étant sélectionné, nous pouvons comparer en détail nos différentes proximités. Nous avons conservé pour cela la pureté, mais avec C_x la liste des 10 plus grosses communautés contenant au moins cinq profils dans V_T pour la proximité x (voir Tableau 5.6)[3]. Nous nous concentrons sur les plus grandes communautés car ce sont celles qui nous permettraient de toucher le plus grand nombre de profils dans le cas d'une propagation communautaire des points de vue. Bien qu'il s'agisse d'une vision partielle – nous avons une médiane de 50 % de profils annotés par communauté – nous pouvons constater que certaines proximités présentent des puretés moyennes extrêmement hautes. Sur les jeux de données Twitter, $cite_{all}$, $cite_{rec}$ et $asso_{all}$ obtiennent dans tous les cas une pureté moyenne proche ou supérieure à 0,90. Les proximités ref et $call_{rec}$ ne sont pas aussi performantes mais ont tout de même de bons résultats. En revanche, les scores de $call_{all}$, $asso_{rec}$ et $asso_{\overline{rec}}$ varient beaucoup selon de jeu de données considéré.

5.5.3 *Information partagée entre proximités*

Pour mesurer la quantité d'information partagée entre proximités et donc répondre à **QP3**, nous avons mesuré l'information mutuelle normalisée (NMI) (Danon et al., 2005) pour chaque paire de proximités (voir Figure 5.1), définie comme :

$$\mathrm{NMI}(x_1, x_2) = \frac{2 \times I(x_1, x_2)}{H(x_1) + H(x_2)}$$

avec

$$I(x_1, x_2) = \sum_{c_i^1 \in C_{x_1}} \sum_{c_j^2 \in C_{x_2}} \left(\frac{|c_i^1 \cap c_j^2|}{|V|} \times \log \frac{|V| \times |c_i^1 \cap c_j^2|}{|c_i^1| \times |c_j^2|} \right)$$

l'information mutuelle entre les proximités x_1 et x_2 et

$$H(x) = -\sum_{c_i \in C_x} \left(\frac{|c_i|}{|C_x|} \times \log \frac{|c_i|}{|C_x|} \right)$$

l'entropie.

La première observation que nous pouvons faire est que les relations entre proximités semblent dépendre fortement du jeu de données. Dans la plupart des cas, chaque proximité apporte une part d'information propre sur les profils. Les

3. C_x peut occasionnellement contenir moins de 10 communautés, soit parce que la liste complète des communautés du graphe de proximité contient moins de 10 communautés, soit parce que, parmi toutes les communautés détectées, il y en a moins de 10 contenant plus de cinq profils présents dans V_T.

proximités ayant les partitions communautaires les plus semblables sont sans surprises *city* et *region*, ainsi que les versions complètes et réciproques des proximités sociales. Il y a beaucoup d'informations redondantes entre proximités sur ME, $call_{all}$, $cite_{all}$ et *ref* étant fortement apparentées les unes aux autres. C'est une observation intéressante étant donné qu'il s'agit du seul jeu de données Twitter exhibant ce phénomène. Les valeurs extrêmement faibles observées entre *kw* et les autres proximités sur SR, #E, PE et GC sont dues au fait que sur ces jeux de données, tous les profils dans V_T sont assignés à la même communauté lorsque la proximité *kw* est considérée.

5.6 IMPLICATIONS POUR LA DÉTECTION DE POINTS DE VUE

Les résultats de ces expérimentations montrent que les communautés détectées à l'aide d'éléments présents sur les médias sociaux sont extrêmement homogènes en termes de points de vue. Ces communautés peuvent donc être un moyen efficace pour propager ceux-ci à partir de quelques profils ayant un point de vue connu. De plus, chaque proximité semble apporter une information spécifique, pouvant potentiellement amener à une caractérisation plus fine des profils. Sans surprise, les versions réciproques des proximités sociales sont sémantiquement proches des versions complètes (comme le montre la quantité d'information mutuelle entre ces paires de proximités) mais leur meilleure homogénéité peut présenter un intérêt non négligeable pour notre tâche.

Même lorsque les jeux de données proviennent de la même plateforme, chacun dispose de ses singularités et, par conséquent, certaines proximités pourraient aussi bien être utiles que nuisibles pour notre tâche en fonction du jeu de données. Sur Twitter, $cite_{all}$ et $asso_{all}$ semblent particulièrement encourageantes : elles produisent des communautés fortement homogènes et, à l'exception de leurs versions réciproques, amènent une information unique. Sur CreateDebate, *ref* et $asso_{rec}$ semblent intéressantes pour les mêmes raisons.

6

MODÈLE COMMUNAUTAIRE SÉQUENTIEL DE DÉTECTION DE POINTS DE VUE

Compte-tenu des hypothèses validées dans le Chapitre 5, nous avons développé un modèle séquentiel de détection de points de vue s'appuyant sur les communautés. Nous le présentons en détail dans ce chapitre, en commençant par ses principes de base, son fonctionnement général puis le fonctionnement de ses modules principaux, permettant l'ordonnancement des proximités et la sélection des profils-graines. Nous l'évaluons ensuite sur cinq jeux de données provenant de deux plateformes différentes pour démontrer son efficacité malgré la quantité extrêmement limitée de profils-graines employés. Les notations utilisées tout au long de ce chapitre sont résumées dans le Tableau 5.1.

6.1 PRINCIPES DE BASE

Au vu des résultats préliminaires présentés précédemment, nous proposons `SCSD`, un modèle communautaire séquentiel de détection de points de vue (Sequential Community-base Stance Detection model). Ce modèle repose sur deux grands principes :

PROPAGATION COMMUNAUTAIRE DES POINTS DE VUE. Les expérimentations précédentes ont confirmé que les communautés détectées sur les médias sociaux pouvaient être extrêmement homogènes en matière de points de vue. Il paraît donc logique de penser que le point de vue des profils d'une communauté peut être aisément déterminé à partir du moment où l'on connaît le point de vue d'une poignée d'entre eux, que nous nommons profils-graines. Ceci est particulièrement utile si nous nous plaçons dans le cadre de chercheurs·ses ayant la volonté d'exploiter de larges jeux de données sans avoir les ressources pour annoter une portion significative de ceux-ci.

IMPORTANCE VARIABLE DES PROXIMITÉS DANS LA DIFFUSION DE POINTS DE VUE. Il est important de noter cependant que les proximités ne sont pas toutes équivalentes pour ce qui est de l'homogénéité de leurs communautés. Certaines sont particulièrement précises mais ne touchent que peu de profils alors que d'autres sont bien plus répandues mais ont une précision moindre. Nous allons tenter d'exploiter au mieux ces différences afin que les proximités se renforcent mutuellement durant la phase de détection des points de vue.

6.2 FONCTIONNEMENT DU MODÈLE

L'idée de base de notre modèle est de propager les points de vue à partir d'une petite quantité de profils dont nous connaissons le point de vue. Nous nous appuyons en priorité sur les proximités les plus fiables, c'est-à-dire les proximités représentant les liens inter-profils les plus forts et qui sont donc les plus à même de relier les profils partageant le même point de vue, compte-tenu du niveau élevé l'homophilie sur les médias sociaux. Nous ferons référence tout au long de cette section à l'exemple d'un jeu de données Twitter sur lequel nous souhaitons détecter deux points de vue afin d'illustrer le fonctionnement du modèle `SCSD`, présenté dans la Figure 6.1 pour les phases d'initialisation et la Figure 6.2 pour la phase de propagation des points de vue.

Le mécanisme central de notre modèle repose sur l'assignation itérative d'un point de vue par communauté. Tout d'abord, plusieurs partitions de profils sont déterminées en fonction de proximités variées (Figure 6.1a). Ces dernières sont ensuite ordonnées par une fonction sélectionnée par l'utilisateur·trice du modèle (Figure 6.1b). Un ensemble de profils est choisi parmi les profils de la première proximité considérée, en prenant en compte les communautés détectées sur celle-ci (Figure 6.1c). Enfin, une étape itérative de détection de point de vue permet de déterminer les points de vue des profils de V à partir des profils sélectionnés à l'étape précédente, nommés les profils-graines, en propageant le point de vue majoritaire de chaque communauté aux profils de cette communautés ayant un point de vue indéterminé (Figure 6.2). L'Algorithme 1 et l'Algorithme 2 synthétisent ce processus.

ALGORITHME 1 – Fonctionnement global de `SCSD`.
$X = (x_1, \ldots, x_n)$ est la séquence de proximités considérées. L'Algorithme 3 détaille la sélection des profils-graines et l'Algorithme 2 la fonction déterminant le point de vue majoritaire d'une communauté.

```
# Initialisation
POUR CHAQUE x_i DANS X :
    C_{x_i} ← détecterCommunautés(x_i)
X^ord ← ordonnerProximités(X, ω(X))
S ← sélectionnerProfilsGraine(x_1^ord, s, s_com, s_min, φ)
POUR CHAQUE v_i DANS S :
    P(v_i) ← récupérerPointDeVue(v_i)
# Processus de détection de points de vue
POUR CHAQUE x_i^ord DANS X^ord :
    POUR CHAQUE c DANS C_{x_i^ord} :
        POUR CHAQUE v_i DANS c :
            SI Σ_{i,•} EST indéfini :
                Σ_{i,•} ← déterminerPointDeVueMajoritaire(c)
```

ALGORITHME 2 – Détermination du point de vue majoritaire d'une communauté c.

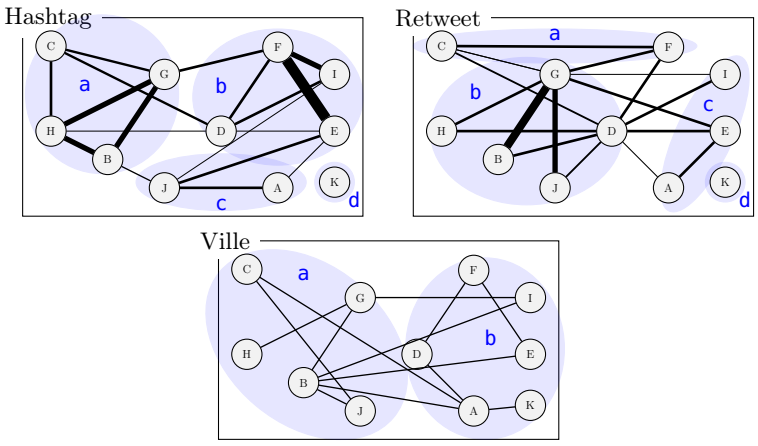
```
1  FONCTION déterminerPointDeVueMajoritaire(c) :
2      # Variable liant points de vue et fréquences d'apparition
3      pdv ← ∅
4      POUR CHAQUE v_i DANS c :
5          SI Σ_{i,•} EST défini :
6              POUR CHAQUE σ_j DANS σ :
7                  SI Σ_{i,j} ≠ 0 :
8                      pdv[Σ_{i,j}].incrémenter()
9      pdvTri ← trierParFréquenceDécroissante(pdv)
10     SI pdv ≠ ∅ ET pdvTri[1].freq ≠ pdvTri[2].freq :
11         RETOURNER pdvTri[1].pdv
12     SINON :
13         RETOURNER indéfini
```

Nous pouvons constater que notre modèle comporte donc deux étapes cruciales : (a) l'ordonnancement des proximités et (b) la sélection des profils–graines. Les sections suivantes présentent les solutions mises au point pour répondre à ces problématiques.

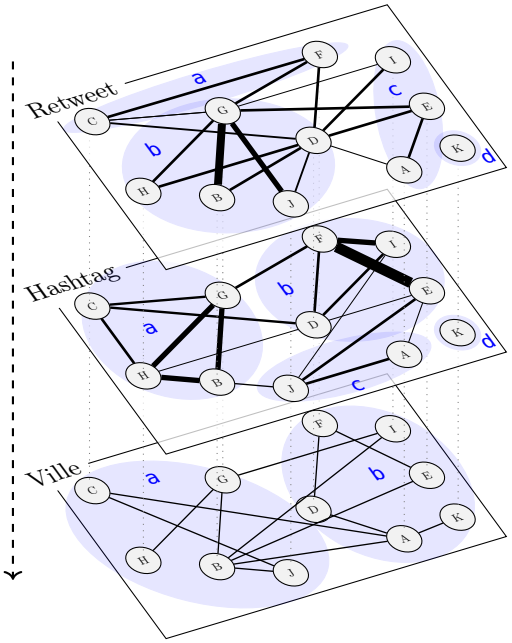
6.2.1 *Ordonnancement des proximités*

Dans le modèle `SCSD`, le point de vue d'un profil ne peut pas être modifié après son assignation à un profil. Par conséquent, l'ordre dans lequel les proximités sont considérées est donc crucial. La fonction d'ordonnancement $\boldsymbol{\omega}$ détermine la séquence ordonnée X^{ord} de proximités à utiliser durant le reste du modèle. La stratégie idéale consiste à placer en premier les proximités donnant les communautés les plus homogènes en termes de point de vue. Étant donné que cette information n'est pas connue a priori dans la majorité des cas, nous présentons ici deux options pour concevoir cette fonction d'ordonnancement $\boldsymbol{\omega}$: une version manuelle et une version automatique.

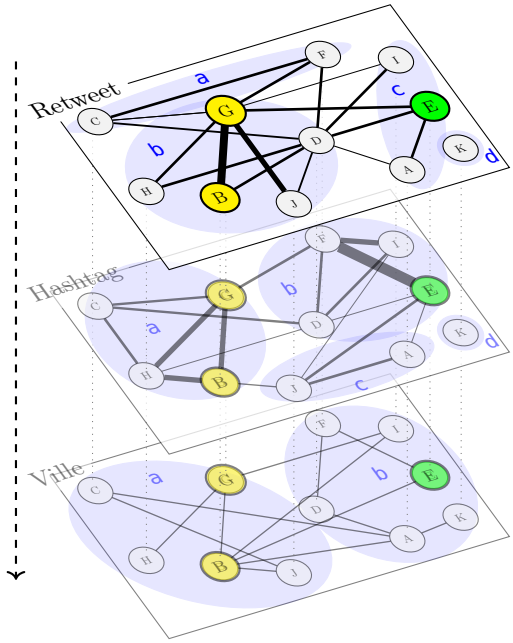
FONCTION D'ORDONNANCEMENT MANUELLE. Ce mode est utile si l'utilisateur·trice du modèle possède une expertise sur le jeu de données considéré et qu'il / elle veut choisir personnellement l'ordre dans lequel les proximités doivent être utilisées.
Dans la Figure 6.1b, nous avons décidé d'ordonner manuellement les proximités et de placer *Retweet* en premier car, compte-tenu de la littérature existante, nous supposons qu'il s'agit de la proximité qui nous donnera les communautés les plus homogènes, suivie par *Hashtag*, afin de tenter de capter des utilisations spécifiques de mots-clés en fonction des points de vue, puis par la proximité géographique *Ville*.



(a) Création des graphes de proximités et détection des communautés pour les proximités $X = (Hashtag, Retweet, Ville)$.

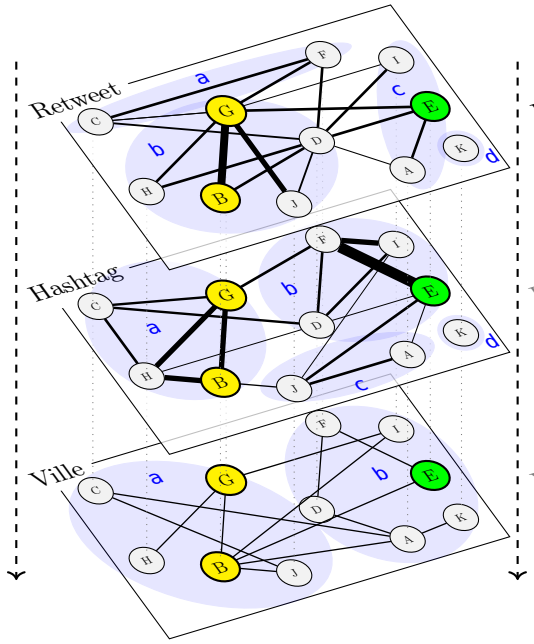


(b) Ordonnancement des proximités : $X^{ord} = (Retweet, Hashtag, Ville)$. Les proximités seront exploitées itérativement de la couche supérieure à la couche inférieure.

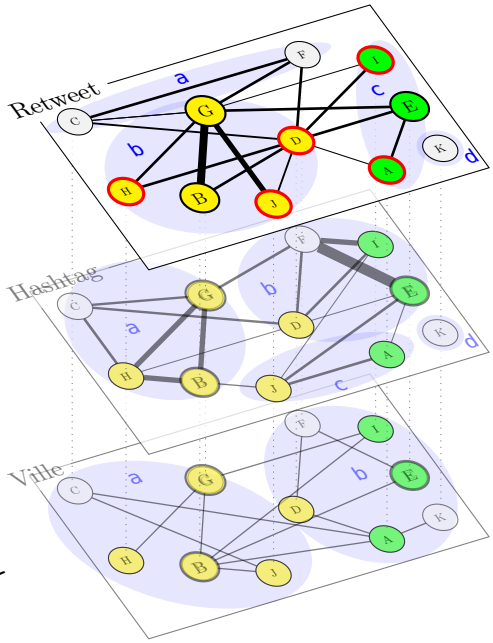


(c) Sélection des profils-graines S en fonction des communautés de la couche supérieure, représentant la première proximité exploitée.

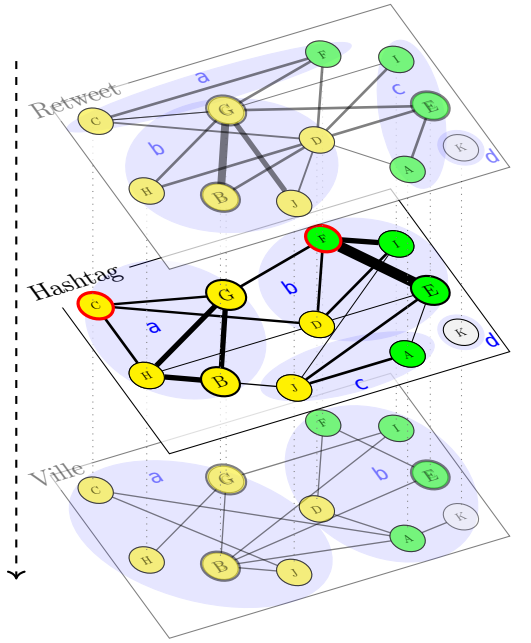
FIGURE 6.1 – Illustration des étapes d'initiation de `SCSD`.
Nous prenons ici l'exemple d'un jeu de données Twitter sur lequel nous voulons détecter deux points de vue : $\sigma = (\texttt{Jaune}, \texttt{Vert})$. Pour cela, nous exploitons une proximité textuelle (*Hashtag*), une proximité sociale (*Retweet*) et une proximité géographique (*Ville*). Nous utilisons une fonction d'ordonnancement manuelle et une sélection manuelle des profils-graines.



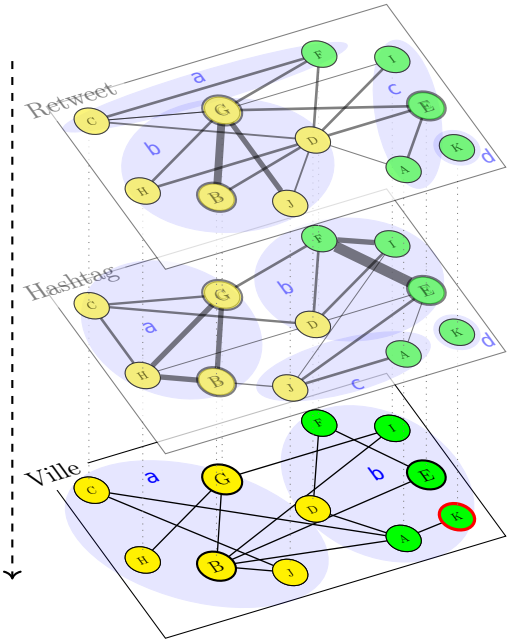
(a) Début du processus itératif : seuls les profils-graines ont un point de vue connu.



(b) Utilisation de la première proximité, *Retweet* : le point de vue majoritaire de chaque communauté est propagé aux profils de la communauté dont le point de vue connu n'est pas encore connu.



(c) Utilisation de la deuxième proximité, *Hashtag* : les points de vue continuent à être propagés, comme décrit dans la Figure 6.2b. Notez que les profils dont le point de vue a été assigné à l'étape précédente ne changent pas.



(d) Utilisation de la troisième proximité, *Ville* : un point de vue est assigné au dernier profil restant.

FIGURE 6.2 – Illustration du processus itératif d'assignation des points de vue de `SCSD`, s'appuyant sur l'exemple introduit dans la Figure 6.1.

FONCTION D'ORDONNANCEMENT AUTOMATIQUE. Le modèle peut aussi automatiquement organiser les proximités selon un critère défini au préalable. Quelques critères d'ordonnancement possibles sont :

— par modularité décroissante ;
— par nombre croissant de communautés détectées ;
— par nombre décroissant de communautés détectées ;
— par nombre croissant de profils actifs au vu de la proximité considérée ;
— par nombre décroissant de profils actifs au vu de la proximité considérée.

Une comparaison de ces fonctions est présentée dans la Section 6.3.

6.2.2 *Sélection des profils-graines*

Nous détaillons à présent notre stratégie pour la sélection des profils-graines. En effet, l'annotation manuelle de données est une tâche longue et coûteuse. `SCSD` est conçu pour être efficace avec un nombre très restreint de profils-graines (c'est-à-dire moins de 5 % du nombre total de profils à catégoriser) et la sélection des profils-graines est dépendante du coût global d'annotation s que l'utilisateur·trice du modèle souhaite y consacrer : s représente le nombre total de profils-graines à annoter manuellement.

Afin de détecter efficacement les points de vue $\boldsymbol{\sigma}$ présents dans le jeu de données étudié, nous considérons au moins s_{com} communautés-graines, c'est-à-dire des communautés contenant des profils-graines, avec $s_{com} \geq |\boldsymbol{\sigma}|$. Ces communautés-graines sont les plus grosses communautés détectées dans la partition de la première proximité considérée x_1^{ord}, contenant en tout au moins s profils.
Une fois les communautés-graines établies, nous devons déterminer le nombre de profils-graines à sélectionner dans chacune d'entre elles. Dans `SCSD`, le nombre de profils-graines présents dans chaque communauté-graine est proportionnel à la taille de cette dernière, tout en tenant compte du fait que chaque communauté-graine doit contenir au moins s_{min} profils-graines. Enfin, après avoir calculé le nombre de profils-graines nécessaire dans chaque communauté-graine, nous sélectionnons ceux-ci grâce à la fonction d'importance $\boldsymbol{\varphi}$.

Les fonctions d'importance traduisent une plus grande importance par une valeur numérique plus élevée et peuvent être déterministe ou non – la fonction aléatoire par exemple. Elles sont catégorisables en deux familles. La première est fondée sur des informations brutes provenant des profils, par exemple :

— pour les jeux de données Twitter : le nombre d'*abonnés*, le nombre d'*amis*, le nombre de *publications*, le nombre de *retweets* des publications du profil, la *séniorité*[1] ;
— pour le jeu de données CreateDebate : l'*efficacité*, le nombre de *débats*, le nombre de *publications*, le nombre de *relations*, la *séniorité*[1].

1. Calculée à l'aide de la date de création du profil.

La seconde famille se fonde sur les graphes de proximité : le *degré*, la *force*, le *PageRank* (PAGE et al., 1999), la *centralité de proximité* (BAVELAS, 1950) et la *centralité de vecteur propre* (NEWMAN, 2010).

La Section 6.3 présente une comparaison de ces fonctions d'importance et l'Algorithme 3 résume le processus que nous venons de décrire ci-dessus.

ALGORITHME 3 – Sélection des profils-graines.
s est le nombre de profils-graines souhaités, s_{com} le nombre minimum de communautés-graines à considérer, s_{min} le nombre minimum de profils-graines devant être présent dans chaque communauté-graine et φ la fonction d'importance permettant de sélectionner les profils-graines.

FONCTION *sélectionnerProfilsGraine*(x_1^{ord}, s, s_{com}, s_{min}, φ) :
  $profilsGraine \leftarrow \varnothing$
  $comsGraine \leftarrow \varnothing$
  $nbProfilsGrainePotentiels \leftarrow 0$
  $comsGrainesPotentielles \leftarrow$ trierParTailleDécroissante$\left(C_{x_1^{ord}}\right)$
  $c \leftarrow comsGrainesPotentielles[1]$
  TANT QUE $|comsGraine| < s_{com}$ OU $nbProfilsGrainePotentiels < s$:
    SI $|c| \geq s_{min}$:
      $comsGraine \leftarrow comsGraine \cup c$
      $nbProfilsGrainePotentiels \leftarrow nbProfilsGrainePotentiels + |c|$
    $c \leftarrow comsGrainesPotentielles$.suivant()
  POUR CHAQUE c DANS $comsGraine$:
    $nbProfils \leftarrow \min\left(|c|, \max\left(s_{min}, \dfrac{|c| \times s}{nbProfilsGrainePotentiels}\right)\right)$
    $nouvellesGraines \leftarrow$ sélectionnerProfilsGraine(c, $nbProfils$, φ)
    $profilsGraine \leftarrow profilsGraine \cup nouvellesGraines$
  RETOURNER $profilsGraine$

Si nous nous penchons sur l'exemple présenté en Figure 6.1c, nous voulons détecter deux points de vue ($|\sigma| = 2$) donc nous savons que $s_{com} \geq 2$. Nous devons par conséquent sélectionner nos profils-graines dans au moins deux communautés. Au vu de la taille minime de notre exemple, nous considérons bien entendu un nombre de profils-graines très petit : nous allons sélectionner uniquement $s = 3$ profils-graines, avec $s_{min} = 1$ seulement. Comme expliqué précédemment, nous nous appuyons sur les communautés de la première proximité utilisée, soit dans notre cas *Retweet*, en sélectionnant en priorité les communautés les plus grandes, c'est à dire `b` et `c`. La communauté-graine `b` étant de taille plus importante, nous sélectionnons de façon manuelle dans celle-ci deux profils-graines, contre un seul dans la communauté-graine `c` : nous considérons que l'utilisatrice finale avait déjà connaissance des points de vue des profils G, B et E, qui deviennent donc les profils-graines qui sont utilisés par la suite pour la propagation des points de vue illustrée dans la Figure 6.2.

6.3 EXPÉRIMENTATIONS

Nous examinons dans cette section l'efficacité du modèle `SCSD`. Nous utilisons des métriques classiques dans le cadre des tâches de classification, à savoir la *précision*, le *rappel* et le *score F1*, macro-moyennées sur l'ensemble des classes et calculées sur le sous-ensemble de profils $V_T \subset V$ défini dans la Section 5.3. Nos expérimentations sont conçues pour mesurer l'impact des différents composants du modèle `SCSD` et peuvent être résumées par les questions suivantes :

Q1. Quelle est l'influence de la fonction d'importance φ, utilisée dans l'étape de sélection des profils-graines, sur les performances du modèle ?

Q2. Quelle est la contribution de chaque proximité ?

Q3. Quelles performances obtiennent les fonctions d'ordonnancement automatiques comparées à un ordonnancement manuel des proximités ?

Q4. `SCSD` est-il robuste aux variations dans le nombre de profils-graines ?

Q5. Quelles sont les performances de `SCSD` comparé aux modèles de base ?

6.3.1 *Influence de la fonction de sélection des profils-graines φ*

6.3.1.1 *Relations entre les fonctions d'importance considérées*

Au vu des fonctions d'importance considérées, présentées en Section 6.2, il paraît logique de penser que certaines risquent d'être trop semblables pour qu'il soit justifié de toutes les étudier. La Figure 6.3 présente les corrélations de Spearman (Fieller, Hartley et Pearson, 1957) entre les rangs des profils restitués par les différentes fonctions d'importance. Comme supposé, de nombreuses fonctions semblent redondantes. Les corrélations importantes entre *degré*, *PageRank* et *force* (souvent proches ou supérieures à 0,70) indiquent que ces fonctions retournent des listes de profils très similaires, ce qui n'est pas surprenant au vu de leurs définitions et de la littérature existante (Fortunato et al., 2008). Les *centralités de proximité* et de *vecteur propre* sont également très corrélées (en moyenne 0,61, allant jusqu'à 0,92), indiquant que leurs évaluations de l'importance d'un profil sont proches. Le nombre de *publications* est souvent corrélé avec le nombre de *retweets* et d'*abonnés* pour Twitter et avec le nombre de *débats* et de *relations* pour CreateDebate. Sur les jeux de données Twitter, considérer les nombres d'*abonnés* et d'*amis* comme mesures d'importance résultent en des listes de profils en grande partie identiques. Ceci est peut-être dû aux profils atypiques (c'est-à-dire les figures publiques) présents dans nos jeux de données ou au fait que les profils actifs s'abonnent à de nombreux profils et en retour attirent de nombreux abonnés. D'après ces observations, nous pouvons omettre certaines fonctions. Nous considérerons donc uniquement le *degré*, la *centralité de proximité*, le nombre de *publications* et la séniorité pour la suite de nos analyses.

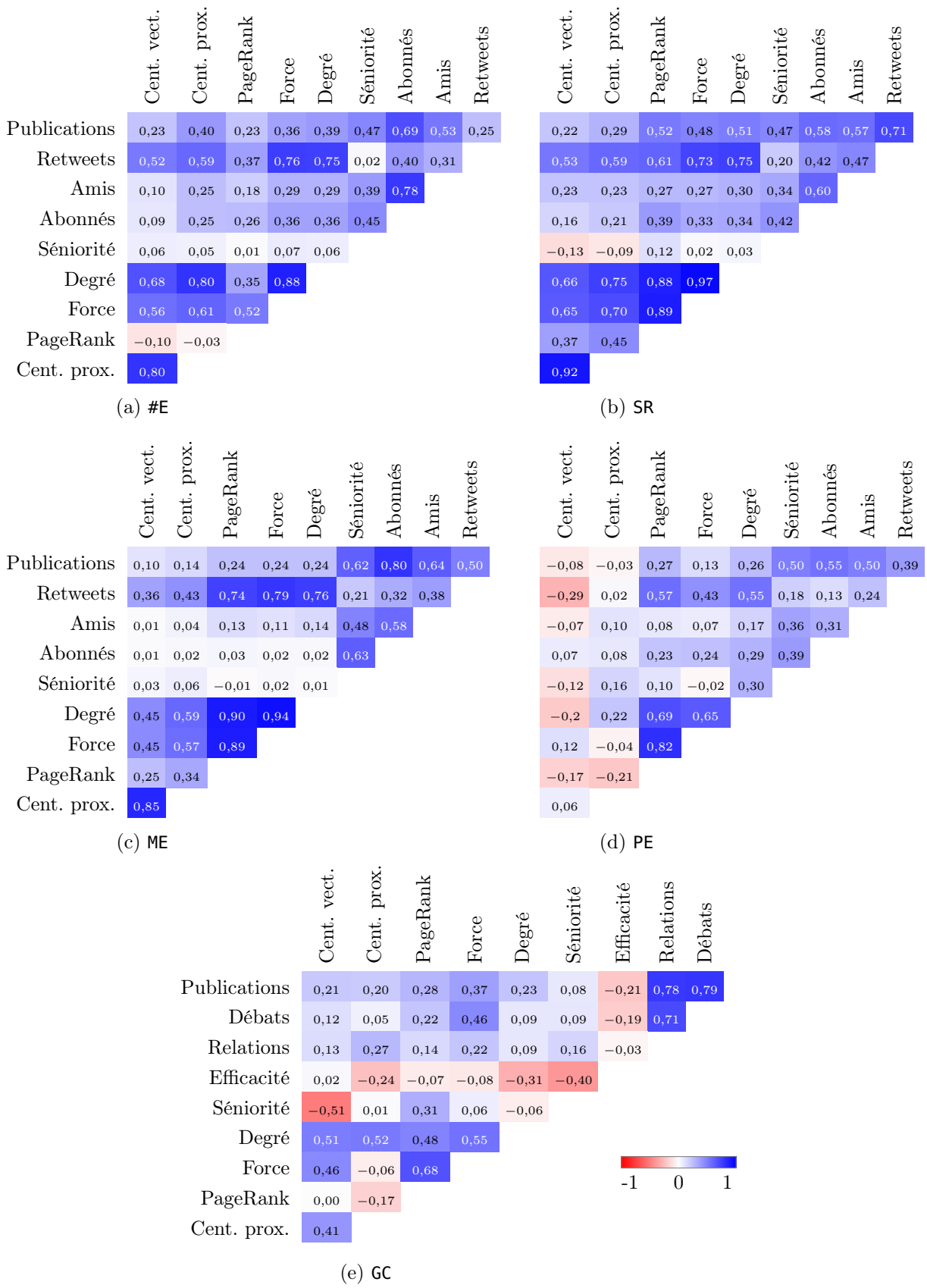


FIGURE 6.3 – Corrélation de Spearman entre les rangs des profils retournés par différentes fonctions d'importance φ.

TABLEAU 6.1 – Scores du modèle SCSD en utilisant différentes fonctions d'importance $\boldsymbol{\varphi}$.
$(s, s_{com}, s_{min}) = (3\,\% \times |V_T|, 3 \times |\boldsymbol{\sigma}|, 3)$ et $\boldsymbol{\omega} = \textit{Manual}$.
Ces scores sont identiques quelle que soit la fonction d'importance $\boldsymbol{\varphi}$ utilisée parmi : *degré*, *centralité de proximité*, nombre de *publications*, *séniorité* et *aléatoire* (moyennes sur 10 exécutions).

	#E	SR	ME	PE	GC
s	560	33	47	27	24
Précision	0.90	0.95	0.95	0.99	0.58
Rappel	0.86	0.95	0.95	0.98	0.52
Score F1	0.87	0.95	0.95	0.98	0.45

6.3.1.2 *Influence sur les performances*

Afin de comparer les fonctions d'importance dans un cadre optimal, nous avons décidé d'ordonner manuellement les proximités. Nous construisons une séquence optimale par jeu de données en ordonnant les proximités par pureté décroissante des communautés (voir Tableau 5.6), avant de prédire les points de vue avec SCSD en utilisant $(s, s_{com}, s_{min}) = (3\,\% \times |V_T|, 3 \times |\boldsymbol{\sigma}|, 3)$ (voir Tableau 6.1). La première observation que nous pouvons faire est qu'avec une séquence de proximités correctement ordonnées, le point de vue attribué aux profils est correct dans la grande majorité des cas (90 % ou plus pour les jeux de données Twitter et presque 60 % pour le jeu de données CreateDebate). La seconde est plus inattendue : les scores ne varient pas quelle que soit la fonction d'importance utilisée pour sélectionner les profils-graines. Il s'agit probablement d'une conséquence du fait que les communautés-graines sont tellement homogènes en termes de points de vue que la méthode de sélection des profils au sein de celles-ci ne change rien au résultat final.

Ce résultat a une application pratique extrêmement intéressante : une fois les communautés-graines sélectionnées et la quantité de profils-graines à annoter dans chacune calculée, il est possible de choisir des profils dont le point de vue est déjà connu. Ceci est particulièrement utile lorsque le modèle est utilisé par un·e expert·e, car il est fort probable qu'elle / il connaisse déjà les points de vue d'une partie des profils du jeu de données étudié.

6.3.2 *Contribution de chaque proximité*

Pour mesurer la contribution de chaque proximité x_i, nous avons mesuré dans le Tableau 6.2 les scores de SCSD-Basic, une version simplifié de SCSD avec une seule proximité, donc $X = (x_i)$, et en fixant $(s, s_{com}, s_{min}, \boldsymbol{\varphi}) = (3\% \times |V_T|, 3 \times |\boldsymbol{\sigma}|, 3, \textit{Degr})$. En effet, la fonction d'importance $\boldsymbol{\varphi}$ utilisée pour la sélection des profils-graines n'ayant pas d'influence sur les résultats (voir Tableau 6.1) nous utilisons simplement pour la suite des expérimentations $\boldsymbol{\varphi} = \textit{Degr}$. De plus,

TABLEAU 6.2 – Scores du modèle `SCSD-Basic`.
$(s, s_{com}, s_{min}, \boldsymbol{\varphi}) = (3\,\% \times |V_T|, 3 \times |\boldsymbol{\sigma}|, 3, Degr)$.

		ref	*kw*	$cite_{all}$	$cite_{rec}$	$call_{all}$	$call_{rec}$	$asso_{all}$	$asso_{rec}$	$asso_{\overline{rec}}$	*socio*	*beliefs*	*city*	*region*	*country*
	P	0,48	0,05	0,91	**0,96**	0,90	0,84	0,89	0,89	0,55			0,41	0,37	
`#E`	R	0,51	0,13	0,84	0,22	**0,88**	0,27	0,74	0,70	0,39			0,06	0,10	
	F1	0,45	0,07	0,87	0,35	**0,89**	0,39	0,79	0,78	0,31			0,08	0,12	
	P	0,51	0,25	0,91	**1,00**	0,25	0,78	0,97	0,99	0,25			0,60	0,56	
`SR`	R	0,37	0,46	0,71	0,34	0,42	0,23	0,84	**0,85**	0,38			0,08	0,10	
	F1	0,33	0,32	0,78	0,50	0,31	0,35	0,90	**0,91**	0,30			0,14	0,16	
	P	0,87	0,57	0,97	**1,00**	0,72	0,94	0,97	0,25	0,67			0,64	0,56	
`ME`	R	0,15	0,34	0,29	0,03	0,02	0,03	**0,81**	0,43	0,63			0,05	0,14	
	F1	0,25	0,38	0,37	0,06	0,04	0,06	**0,88**	0,31	0,63			0,09	0,22	
	P	0,74	0,26	0,99	**1,00**	0,74	0,94	0,98	0,97	0,26			0,65	0,58	
`PE`	R	0,03	0,48	0,54	0,29	0,21	0,13	**0,86**	0,81	0,39			0,09	0,18	
	F1	0,06	0,33	0,69	0,45	0,32	0,23	**0,92**	0,88	0,31			0,15	0,26	
	P	0,38	0,31	0,58	**0,83**	0,60	0,72		0,59		0,39	0,56			0,64
`GC`	R	0,02	**0,50**	0,04	0,02	0,05	0,03		0,17		0,21	0,21			0,23
	F1	0,03	**0,39**	0,08	0,04	0,09	0,05		0,24		0,25	0,27			0,25

X ne contenant dans tous les cas qu'une proximité avec `SCSD-Basic`, nous n'utilisons aucune fonction d'ordonnancement $\boldsymbol{\omega}$.

Malgré le faible nombre de profils-graines utilisés, certaines proximités permettent des prédictions extrêmement précises, mais elles ne concernent généralement que peu de profils, menant à un faible rappel (entre 3 % et 34 % de rappel pour les proximités nous donnant 100 % de précision). Ce phénomène est légèrement moins présent dans `#E`, sur lequel nous obtenons d'excellents résultats pour la détection des cinq points de vue. Les faibles scores obtenus par $asso_{rec}$ sur `ME` sont surprenants au vu des résultats sur les autres jeux de données : en effet, `ME` obtient seulement 25 % de précision, comparé à des précisions allant de 59 % à 99 %. Une analyse plus poussée nous a permis de découvrir que cette chute de performance était due au fait que sur ce jeu de données, les profils dans V_T sont fortement connectés entre eux avec cette proximité. Lors de la détection de communauté, la majorité des profils de V_T sont donc assignés à la même communauté, causant donc l'assignation d'une grande partie des profils au mauvais point de vue. Bien que le contrôle des armes à feu soit souvent présenté comme un débat entre « droite » et « gauche » ou comme un conflit entre générations, les proximités *socio* et *beliefs* n'obtiennent pas de bonnes performances (seulement 39 % et 56 % de précision, et pas plus de 21 % de rappel). Cela est possiblement dû à la petite taille du jeu de données (voir Tableau 5.3) et au caractère très basique des descripteurs sociologiques dont nous disposons.

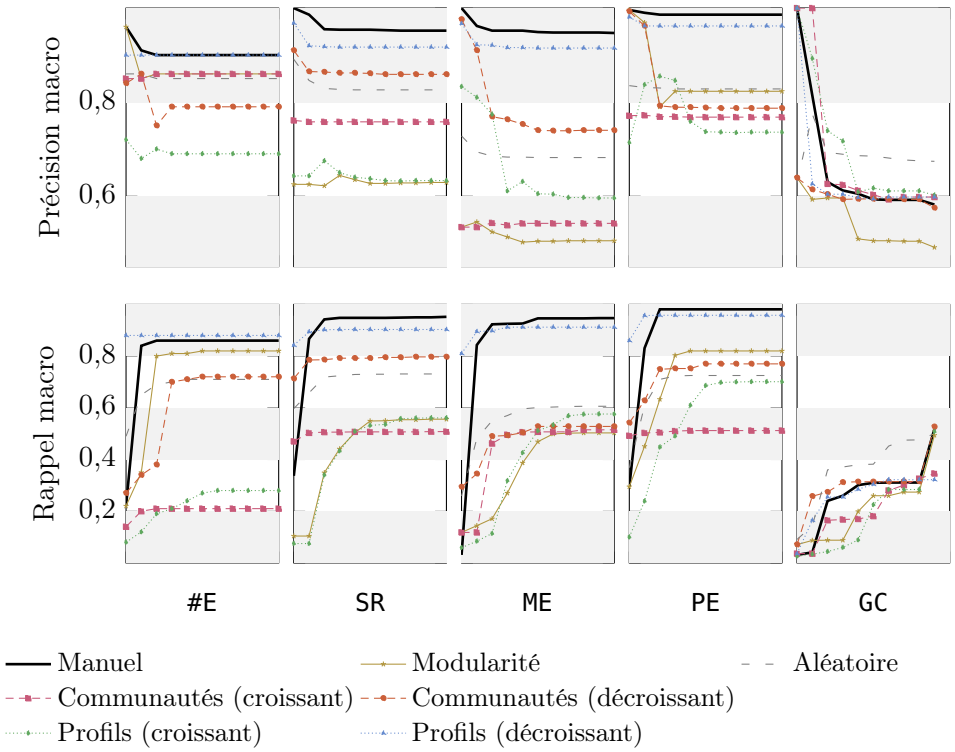


FIGURE 6.4 – Influence des fonctions d'ordonnancement $\boldsymbol{\omega}$ sur la précision et le rappel du modèle SCSD.
$(s, s_{com}, s_{min}, \boldsymbol{\varphi}) = (3\,\%\times|V_T|, 3\times|\boldsymbol{\sigma}|, 3, Degr)$.
Chaque point représente l'utilisation d'une proximité x_i^{ord} dans X^{ord}.
Les scores présentés pour $\boldsymbol{\omega} = Random$ sont les scores moyens sur 10 exécutions.

6.3.3 *Influence de la fonction d'ordonnancement $\boldsymbol{\omega}$*

L'ordre des proximités est un élément clé du modèle. En effet, celles produisant les communautés les plus homogènes doivent être utilisées en premier afin d'obtenir les meilleurs résultats. Nous avons comparé les performances du modèle SCSD en utilisant différentes séquences de proximités fournies par les fonction d'ordonnancement $\boldsymbol{\omega}$ présentées dans la Section 6.2, avec $(s, s_{com}, s_{min}, \boldsymbol{\varphi}) = (3\,\%\times|V_T|, 3\times|\boldsymbol{\sigma}|, 3, Degr)$. Nous considérons également un ordonnancement aléatoire en moyennant les scores sur 10 exécutions. La Figure 6.4 présente l'évolution de la précision et du rappel au cours du processus itératif d'attribution des points de vue et le Tableau 6.3 les scores F1 finaux. Pour Twitter, ordonner les proximités par nombre décroissant de profils est une bonne alternative à l'ordonnancement manuel. Pour #E, SR et PE le nombre décroissant de communautés pourrait être une alternative acceptable, probablement car un grand nombre de communautés implique que celles-ci soient plus petites et plus homogènes sur ces jeux de données. Pour GC, la modularité et le nombre décroissant de communautés donnent des résultats légèrement

TABLEAU 6.3 – Influence des fonctions d'ordonnancement $\boldsymbol{\omega}$ sur les scores F1 du modèle SCSD.
$(s, s_{com}, s_{min}, \boldsymbol{\varphi}) = (3\,\% \times |V_T|, 3 \times |\boldsymbol{\sigma}|, 3, Degr)$.
Les scores présentés pour $\boldsymbol{\omega} = Alatoire$ sont les scores moyens sur 10 exécutions.

FONCTION D'ORDONNANCEMENT $\boldsymbol{\omega}$	#E	SR	ME	PE	GC
Aléatoire	0,69	0,67	0,53	0,67	0,43
Manuel	0,87	**0,95**	**0,95**	**0,98**	0,45
Modularité décroissante	0,83	0,49	0,50	0,82	0,47
Nombre croissant de communautés	0,11	0,38	0,43	0,38	0,43
Nombre décroissant de communautés	0,74	0,79	0,40	0,77	**0,48**
Nombre croissant de profils présents	0,21	0,50	0,55	0,70	0,41
Nombre décroissant de profils présents	**0,89**	0,90	0,91	0,96	0,34

meilleurs que l'ordonnancement manuel et la seule fonction $\boldsymbol{\omega}$ produisant des résultats significativement affaiblis est le nombre décroissant de profils. Nous avions, avant ces expérimentations, l'intuition que la modularité pourrait être une fonction efficace pour l'ordonnancement des proximités. Malheureusement les résultats indiquent que cela varie énormément en fonction du jeu de données. Elle donne de bons résultats sur #E, PE et GC mais fait fortement chuter les scores pour SR et ME.

6.3.4 *Influence du nombre de profils-graines considérés*

Nous avons ensuite examiné l'influence du nombre de profils-graines et de communautés-graines sur les performances. Le Tableau 6.4 présente les scores F1 du modèle SCSD lorsque s et s_{com} varient, avec $(s_{min}, \boldsymbol{\varphi}) = (3, Degr)$ et $\boldsymbol{\omega} = Manual$. Avec nos cas de test, les scores de SR et ME ne varient pas, mais en revanche #E, PE et GC illustrent l'influence de s et s_{com}. SCSD obtient des performances dégradées sur GC en utilisant le nombre minimal de profils-graines et de communautés-graines, à savoir $(s, s_{com}) = (1\% \times |V_T|, |\boldsymbol{\sigma}|)$. Pour #E, s_{com} semble avoir plus d'importance que s étant donné que l'on observe une chute drastique des scores F1 en utilisant $s_{com} = |\boldsymbol{\sigma}|$ (le score F1 passant d'environ 90 % à environ 50 %), qui n'est pas compensée par l'augmentation de s. En revanche, les résultats restent excellents avec $s_{com} = 2 \times |\boldsymbol{\sigma}|$ même en ne sélectionnant que 1 % de profils-graines. Ce phénomène est également présent pour PE bien que de façon beaucoup moins notable, le score F1 ne diminuant que deux points. Ces résultats confirment que le modèle SCSD est construit pour être efficace avec un nombre de profils-graines extrêmement faible, le fait de prendre plus de 3 % de profils-graines n'apportant aucun bénéfice au niveau des performances.

TABLEAU 6.4 – Influence de s et s_{com} sur les scores F1 de SCSD, avec $(s_{min}, \boldsymbol{\varphi}) = (3, Degree)$ et $\boldsymbol{\omega} = Manual$.
Les valeurs manquantes caractérisent les cas pour lesquels la sélection des profils-graines est impossible car $s < s_{com} \times s_{min}$.

s	$1\,\% \times \lvert V_T \rvert$		$3\,\% \times \lvert V_T \rvert$		$5\,\% \times \lvert V_T \rvert$		$7\,\% \times \lvert V_T \rvert$	
s_{com}	$\lvert\sigma\rvert$	$2 \times \lvert\sigma\rvert$	$\lvert\sigma\rvert$	$2 \times \lvert\sigma\rvert$	$\lvert\sigma\rvert$	$2 \times \lvert\sigma\rvert$	$\lvert\sigma\rvert$	$2 \times \lvert\sigma\rvert$
#E	0,47	0,87	0,47	0,87	0,47	0,88	0,48	0,88
SR	0,95	-	0,95	0,95	0,95	0,95	0,95	0,95
ME	0,95	0,95	0,95	0,95	0,95	0,95	0,95	0,95
PE	0,88	-	0,88	0,90	0,88	0,90	0,88	0,90
GC	0,29	-	0,45	0,45	0,45	0,45	0,45	0,45

6.3.5 *Comparaison avec les modèles de base*

Au vu du manque de modèles génériques dans la littérature, nous avons choisi de comparer les performances du modèle SCSD à plusieurs modèles classiques reconnus pour la détection de points de vue, plutôt que d'adapter de façon imparfaite des modèles conçus sur les spécificités d'une plateforme donnée. Nous utilisons à la fois le contenu textuel des publications et les interactions sociales des profils, étant donné qu'il s'agit des deux catégories d'information communément exploitées. Pour l'aspect social, nous nous concentrons sur les deux proximités majoritairement utilisées pour cette tâche dans la littérature, à savoir $cite_{all}$ et $asso_{rec}$ [2]. Nous comparons nos résultats à la fois à des modèles supervisés, utilisant une quantité de données annotées significativement supérieure à SCSD, ainsi qu'à des modèles semi-supervisés fonctionnant, eux, avec la même portion de données annotées que notre modèle.

Les modèles de base supervisés que nous considérons sont :

SVM Un modèle SVM fondé sur la concaténation des publications de chaque profil, avec un vocabulaire constitué des 10 000 mots les plus significatifs d'après une mesure de χ^2.

RF_{cite} Un classifieur Random Forest fondé sur $cite_{all}$: chaque profil est représenté par ses similarités aux autres profils selon cette proximité.

RF_{asso} Identique à RF_{cite} mais utilisant la proximité $asso_{rec}$.

Nous utilisons pour ces modèles une validation croisée à cinq échantillons. Cependant, étant donnée la quantité de données annotées utilisée par ces modèles, la comparaison avec SCSD est quelque peu déséquilibrée.

2. Nous utilisons $asso_{rec}$ car $asso_{all}$ n'est pas défini pour GC, l'interaction *allié* étant forcément réciproque.

TABLEAU 6.5 – Scores des modèles de base et du modèle SCSD avec la configuration optimale par collection.

		SVM	RF_{cite}	RF_{asso}	LP_{cite}	LP_{asso}	SCSD
		Annotations : 80%			Annotations : 3%		
#E	P	0,75	**0,91**	0,87	0,21	0,11	0,90
	R	0,76	**0,89**	0,82	0,19	0,09	0,88
	F1	0,75	**0,89**	0,83	0,20	0,10	**0,89**
SR	P	0,92	0,94	**0,95**	0,89	**0,95**	**0,95**
	R	0,92	0,91	0,94	0,76	0,93	**0,95**
	F1	0,92	0,90	0,94	0,78	0,94	**0,95**
ME	P	0,88	0,93	**0,96**	0,20	0,11	0,95
	R	0,87	0,92	0,93	0,33	0,08	**0,95**
	F1	0,87	0,92	0,93	0,24	0,07	**0,95**
PE	P	0,89	0,94	0,98	0,12	0,91	**0,99**
	R	0,89	0,93	0,97	0,11	0,93	**0,98**
	F1	0,89	0,92	0,97	0,09	0,91	**0,98**
GC	P	0,43	**0,62**	0,55	0,45	0,40	0,57
	R	0,37	**0,54**	0,51	0,04	0,17	0,53
	F1	0,37	**0,51**	0,50	0,08	0,19	0,48

Pour contrer cela, nous comparons également SCSD à des modèles semi-supervisés utilisant la même quantité de données annotées :

$\texttt{LP}_{cite}$ Un processus de propagation de labels semi-supervisés utilisant le graphe de la proximité $cite_{all}$. La sélection des profils-graines utilisant comme fonction d'importance $\varphi = Random$, nous présentons les scores moyens sur 10 exécutions.

$\texttt{LP}_{asso}$ Identique à $\texttt{LP}_{cite}$ mais utilisant la proximité $asso_{rec}$.

Le Tableau 6.5 présente les résultats de SCSD dans sa configuration optimale par collection, déduite des expérimentations précédentes, comparés aux résultats des modèles de base. SCSD obtient, avec 30 fois moins de données annotées, des scores supérieurs ou proches des modèles de base supervisés. Comparé aux modèles de base semi-supervisés utilisant la même quantité de données annotées, il gagne en moyenne 47 points de score F1 (allant de 1 et 88 points de pourcentage). Pour ME, les faibles scores obtenus par $\texttt{LP}_{asso}$ sont probablement liés à la surabondance de liens observée dans la Section 6.3.2. Cela illustre le problème principal lorsque le modèle repose sur une proximité unique : si celle-ci n'est pas adaptée au jeu de données étudié, rien ne permet de rectifier la situation et les prédictions en résultant ne sont pas pertinentes. L'homogénéité des communautés données par $cite_{all}$ et $asso_{rec}$ étant aussi élevée pour #E que pour SR, les différences observées dans les scores sont probablement dues aux multiples points de vue de #E qui sont visiblement bien plus compliqués à gérer pour la propagation de labels que la problématique bipartite simple représentée par SR. Pour tous les modèles, les profils de GC sont significativement plus compliqués à catégoriser, probablement à cause de la petite taille du jeu de données : nous avons approximativement en moyenne 200 interactions par profil dans V_T pour chaque proximité pour GC, contre 600 pour #E, 2 500 pour SR, 500 pour ME et 5 900 pour PE (voir Tableau 5.3 et Tableau 5.2).

6.4 CONCLUSION

Nous avons présenté dans ce chapitre notre modèle séquentiel de détection des points de vue SCSD. Nos expérimentations, réalisées sur cinq jeux de données provenant de deux plateformes différentes et traitant de problématiques politiques diverses, ont montré à quel point les profils partageant les mêmes points de vue étaient proches sur les médias sociaux, particulièrement lorsque les proximités employées pouvaient être considérées comme des *liens forts* entre les profils.

Le modèle communautaire séquentiel de détection des points de vue a confirmé qu'il était possible, à partir d'une quantité de profils-graines extrêmement limitée, de détecter avec succès les points de vue inconnus en s'aidant des communautés. La force de ce modèle est de s'appuyer en priorité sur les liens inter-profils les plus forts, puis de se propager itérativement via des liens de plus en plus faibles. Malheureusement cette démarche itérative constitue également l'une de ses limites étant donné que l'ordonnancement des proximités est crucial

à son bon fonctionnement. Nous avons donc proposé un procédé d'ordonnancement automatique des proximités, bien que l'ordonnancement par expert·e reste la meilleure garantie de qualité des résultats. Nous pourrions également considérer une prise en compte plus fine des signaux contradictoires permettant de mettre à jour le point de vue assigné à un profil si des faisceaux d'indices suffisamment convaincants se présente par la suite. L'autre limite de ce modèle est qu'il ne permet pas à un profil d'exprimer plusieurs points de vue, limite que nous avons tentée d'adresser dans notre modèle suivant, présenté dans le Chapitre 7.

7

ÉVOLUTION TEMPORELLE DES POINTS DE VUE

Nous avons présenté dans le chapitre précédent le modèle `SCSD`, s'appuyant sur les communautés détectées sur un ensemble de proximités pour catégoriser les profils par points de vue. L'une de ses limites est cependant son incapacité à gérer le fait qu'un profil peut changer d'avis, et donc la dérive conceptuelle inhérente aux points de vue, ces derniers évoluant en fonction des événements. Nous présentons ici un modèle permettant d'observer l'évolution des points de vue au cours d'un événement avec seulement un profil-graine par point de vue. Nous commençons par introduire ses principes fondateurs, avant de détailler sa formalisation puis de l'évaluer sur le jeu de données #Élysée2017fr.

7.1 INTRODUCTION AU MODÈLE

Comme souligné dans la Section 2.3 (p. 31), la plupart des modèles de détection de points de vue existants se concentrent sur la détection de points de vue à un moment donné et ignorent l'évolution possible de ce point de vue au cours du temps. Or, lors d'un événement long et suscitant beaucoup de débats, tel qu'une campagne présidentielle, il peut être essentiel de prendre en compte la dérive de ceux-ci. Nous proposons donc ici un modèle d'observation des points de vue lors d'un événement long fondé sur les principes présentés ci-après.

L'idée majeure sur laquelle repose le modèle `SCSD` (Chapitre 6) est l'importance du voisinage des profils, qui produit des communautés de points de vue homogènes. Cependant, l'ordonnancement des proximités et l'utilisation des communautés non recouvrantes rendent le modèle peu flexible pour une utilisation dynamique. En effet, l'utilisation de `SCSD` impliquerait, pour chaque période considérée, de détecter de nouvelles communautés et donc de sélectionner de nouveaux profils-graines étant donné que ceux-ci dépendent des communautés découvertes. Or, dans notre cas, il semble légitime de vouloir considérer les mêmes profils-graines tout au long de l'événement. De plus, le fait de considérer une multitude de périodes augmente l'aspect épars des liens dans nos graphes de proximités, ce qui peut potentiellement affaiblir la pertinence des communautés détectées et augmenter le nombre de profils difficilement atteignables par propagation. Les performances du modèle seraient donc d'autant plus compromises que la granularité choisie pour étudier les points de vue est fine. Nous allons donc nous concentrer sur un modèle extrêmement souple exploitant directement les similarités inter-profils et fonctionnant avec la plus petite graine possible, à savoir *un seul* profil connu par point de vue, afin de rester dans le cadre de chercheurs·ses n'ayant pas les ressources pour annoter de larges quantités de données.

7.2 PRÉSENTATION FORMELLE

Comme présenté dans la Section 5.2, soient $V = \{v_1, v_2, \ldots\}$ un ensemble de profils, $\boldsymbol{\sigma}$ l'ensemble des points de vue distincts exprimés, Π l'ensemble des périodes sur lesquelles nous étudions les points de vue, X la liste des proximités utilisées et $\boldsymbol{\Sigma}$ le tenseur d'ordre 3 de dimension $|V| \times |\boldsymbol{\sigma}| \times |\Pi|$ notant les points de vue exprimés sur l'ensemble des périodes. Nous considérons donc que $\boldsymbol{\Sigma}_{i,j,p} \in [0, 1]$ représente le score d'expression du point de vue $\boldsymbol{\sigma}_j$ par le profil v_i pendant la période p. Nous présentons dans les sections suivantes comment calculer ces scores et prédire les points de vue des profils à partir de ces derniers à l'aide d'une variante d'un modèle de Rocchio (Rocchio, 1971).

La Figure 7.1 présente un exemple sur lequel nous nous appuierons tout au long de cette section afin d'expliquer le fonctionnement de notre modèle. Nous considérons dans cet exemple les points de vue $\boldsymbol{\sigma} = (\texttt{Jaune}, \texttt{Vert})$ avec $|V| = 40$, $|X| = 2$ et $|\Pi| = 5$.

7.2.1 *Construction des matrices de voisinage*

Notre modèle considère de façon simultanée toutes les proximités. Pour cela, à chaque période $p \in \Pi$, un graphe G_i^p est construit pour chaque proximité $x_i \in X$, puis les matrices d'adjacence de tous les graphes de la période sont concaténées. La matrice résultant de cette concaténation est appelée *matrice de voisinage* des profils pour la période p et notée $\mathcal{M}^p$. Par construction, elle contient $|V|$ lignes correspondant aux $|V|$ profils et $|V| \times |X|$ colonnes représentant leurs liens avec les autres profils. $\mathcal{M}^p_{i,\bullet}$ représente le vecteur caractérisant le profil v_i pour la période p en tenant compte de *toutes les proximités considérées*.

La Figure 7.2 présente un extrait de la matrice de voisinage de la première période considérée pour l'exemple introduit dans la Figure 7.1, illustrant comment la matrice de voisinage peut intuitivement traduire des similarités inter-profils. En effet, bien que nous ne voyions pas l'intégralité de la matrice de dimension $(40, 80)$, le profil B a l'air de partager des similarités avec le profil-graine `Jaune` alors que le profil C, lui, est plutôt similaire au profil-graine `Vert` au vu des informations visibles ici. Le vecteur caractérisant le profil A est nul car celui-ci était inactif sur cette période, comme le montre son isolement dans les graphes de proximité de la Figure 7.1.

7.2.2 *Mesure des similarités entre profils et points de vue*

Comme indiqué dans la Section 7.1, nous ne considérons qu'un seul profil-graine par point de vue. Nous considérons donc que $\mathcal{M}^p_{\boldsymbol{\sigma}_j,\bullet}$ est le vecteur représentant le profil-graine du point de vue $\boldsymbol{\sigma}_j$. Le score d'expression du point de vue $\boldsymbol{\sigma}_j$ par le profil v_i est mesuré comme étant la similarité cosinus entre le

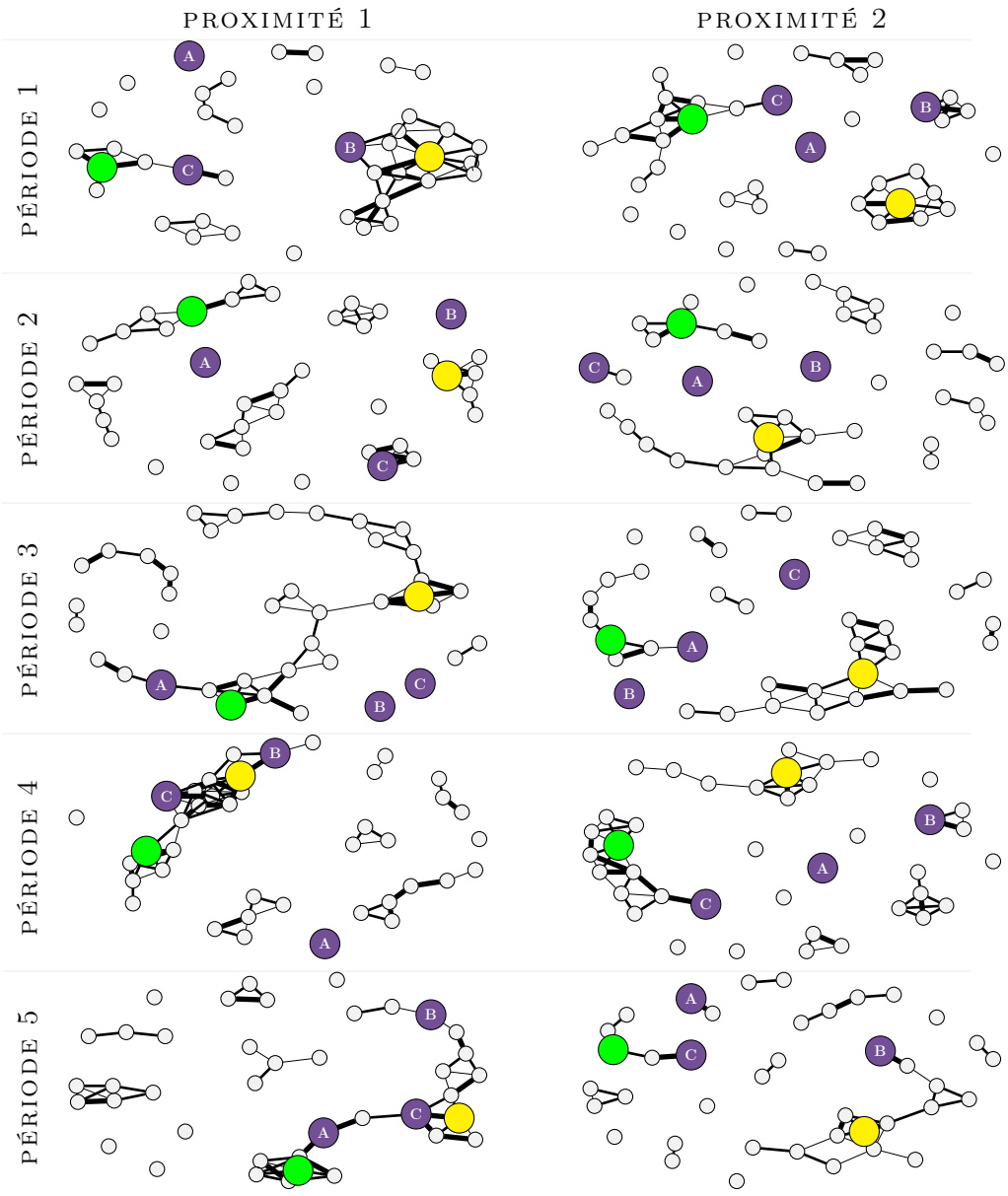


FIGURE 7.1 – Exemple d'évolution des graphes de proximités.
Nous considérons dans cet exemple 40 profils, qui apparaissent donc dans chacun des graphes de proximités, et deux points de vue : `Jaune` et `Vert`. Le nœud représentant le profil-graine de chaque point de vue est coloré de façon correspondante. Nous mettons également en évidence trois profils – A, B et C – sur lesquels nous nous concentrerons par la suite.

$$
\begin{array}{lcl}
 & & \overbrace{}^{\text{Proximité 1 (40 colonnes)}} \quad \overbrace{}^{\text{Proximité 2 (40 colonnes)}} \\
\mathcal{M}^1_{\texttt{Jaune},\bullet} & = & (\ 0\ 2\ 3\ 0\ 1\ \ldots\ 2\ 0\ \ 0\ 1\ \ldots\ 0\ 2\ 2\ 0\ 1\ ) \\
\mathcal{M}^1_{\texttt{Vert},\bullet} & = & (\ 1\ 0\ 0\ 3\ 0\ \ldots\ 0\ 1\ \ 2\ 1\ \ldots\ 1\ 1\ 0\ 4\ 0\ ) \\
\mathcal{M}^1_{\mathrm{A},\bullet} & = & (\ 0\ 0\ 0\ 0\ 0\ \ldots\ 0\ 0\ \ 0\ 0\ \ldots\ 0\ 0\ 0\ 0\ 0\ ) \\
\mathcal{M}^1_{\mathrm{B},\bullet} & = & (\ 0\ 1\ 1\ 0\ 2\ \ldots\ 1\ 0\ \ 0\ 3\ \ldots\ 0\ 3\ 0\ 0\ 1\ ) \\
\mathcal{M}^1_{\mathrm{C},\bullet} & = & (\ 2\ 0\ 1\ 2\ 0\ \ldots\ 0\ 1\ \ 1\ 0\ \ldots\ 1\ 0\ 0\ 1\ 0\ )
\end{array}
$$

FIGURE 7.2 – Extrait de la matrice de voisinage de dimension $(40, 80)$ de la première période de l'exemple présenté dans la Figure 7.1.

vecteur caractérisant le profil v_i et le vecteur caractérisant le profil-graine du point de vue $\boldsymbol{\sigma}_j$ durant la période p :

$$\boldsymbol{\Sigma}_{i,j,p} = \cos(\mathcal{M}^p_{i,\bullet}, \mathcal{M}^p_{\boldsymbol{\sigma}_j,\bullet})$$

Cette mesure présente l'avantage de ne pas se concentrer sur les liens directs entre profils mais sur les *actions communes* : si A et B partagent tous deux les publications de C mais ne partagent pas les publications l'un de l'autre, ou discutent avec C sans interagir directement entre eux, notre mesure considérera néanmoins qu'ils sont proches.

La Figure 7.3a présente les scores d'expressions des points de vue dans notre exemple pour les trois profils sur lesquels nous nous concentrons. Les scores de la première période valident les observations réalisées dans la Section 7.2.1 sur la Figure 7.2 : le profil B a un score d'expression plus important pour le point de vue `Vert` que pour le point de vue `Jaune`, alors que c'est l'inverse pour le profil C. Les scores du profil A sons tous nuls, confirmant l'inactivité du profil sur cette période.

7.2.3 *Prédiction des points de vue des profils*

Une fois les scores des différents points de vue calculés pour chaque période, nous considérons que le point de vue d'un profil est donné par le plus grand score d'expression :

$$\text{Prédiction}^p(v_i) = \begin{cases} \arg\max_j \left(\boldsymbol{\Sigma}_{i,j,p}\right) & \text{si } \boldsymbol{\Sigma}_{i,j,p} \neq \vec{0} \\ \textit{Indéterminé} & \text{sinon.} \end{cases}$$

Si plusieurs points de vue sont ex-æquo avec le score le plus important, ils sont tous exprimés par le profil.

$$\boldsymbol{\Sigma}_{(\mathrm{A,\ B,\ C}),\bullet,1} = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cc} \texttt{J} & \texttt{V} \end{array} \\ \begin{pmatrix} 0{,}0 & 0{,}0 \\ 0{,}4 & 0{,}0 \\ 0{,}1 & 0{,}3 \end{pmatrix} \end{array} \qquad \boldsymbol{\Sigma}_{(\mathrm{A,\ B,\ C}),\bullet,1} = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cc} \texttt{J} & \texttt{V} \end{array} \\ \begin{pmatrix} 0{,}0 & 0{,}0 \\ 0{,}4 & 0{,}0 \\ 0{,}1 & 0{,}3 \end{pmatrix} \end{array}$$

$$\boldsymbol{\Sigma}_{(\mathrm{A,\ B,\ C}),\bullet,2} = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cc} \texttt{J} & \texttt{V} \end{array} \\ \begin{pmatrix} 0{,}0 & 0{,}0 \\ 0{,}0 & 0{,}0 \\ 0{,}3 & 0{,}2 \end{pmatrix} \end{array} \qquad \boldsymbol{\Sigma}_{(\mathrm{A,\ B,\ C}),\bullet,2} = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cc} \texttt{J} & \texttt{V} \end{array} \\ \begin{pmatrix} 0{,}0 & 0{,}0 \\ 0{,}4 & 0{,}0 \\ 0{,}3 & 0{,}2 \end{pmatrix} \end{array}$$

$$\boldsymbol{\Sigma}_{(\mathrm{A,\ B,\ C}),\bullet,3} = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cc} \texttt{J} & \texttt{V} \end{array} \\ \begin{pmatrix} 0{,}0 & 0{,}7 \\ 0{,}0 & 0{,}0 \\ 0{,}0 & 0{,}0 \end{pmatrix} \end{array} \qquad \boldsymbol{\Sigma}_{(\mathrm{A,\ B,\ C}),\bullet,3} = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cc} \texttt{J} & \texttt{V} \end{array} \\ \begin{pmatrix} 0{,}0 & 0{,}7 \\ 0{,}4 & 0{,}0 \\ 0{,}3 & 0{,}2 \end{pmatrix} \end{array}$$

$$\boldsymbol{\Sigma}_{(\mathrm{A,\ B,\ C}),\bullet,4} = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cc} \texttt{J} & \texttt{V} \end{array} \\ \begin{pmatrix} 0{,}0 & 0{,}0 \\ 0{,}4 & 0{,}0 \\ 0{,}3 & 0{,}3 \end{pmatrix} \end{array} \qquad \boldsymbol{\Sigma}_{(\mathrm{A,\ B,\ C}),\bullet,4} = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cc} \texttt{J} & \texttt{V} \end{array} \\ \begin{pmatrix} 0{,}0 & 0{,}7 \\ 0{,}5 & 0{,}0 \\ 0{,}3 & 0{,}3 \end{pmatrix} \end{array}$$

$$\boldsymbol{\Sigma}_{(\mathrm{A,\ B,\ C}),\bullet,5} = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cc} \texttt{J} & \texttt{V} \end{array} \\ \begin{pmatrix} 0{,}0 & 0{,}5 \\ 0{,}3 & 0{,}0 \\ 0{,}3 & 0{,}3 \end{pmatrix} \end{array} \qquad \boldsymbol{\Sigma}_{(\mathrm{A,\ B,\ C}),\bullet,5} = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \begin{array}{c} \begin{array}{cc} \texttt{J} & \texttt{V} \end{array} \\ \begin{pmatrix} 0{,}0 & 0{,}5 \\ 0{,}3 & 0{,}0 \\ 0{,}3 & 0{,}3 \end{pmatrix} \end{array}$$

(a) Sans prise en compte de l'absence d'activité.

(b) En prenant en compte l'absence d'activité comme décrit dans la Section 7.2.4.

FIGURE 7.3 – Scores d'expression des point de vue `Jaune` (`J`) et `Vert` (`V`) par les profils A, B et C pendant les cinq périodes considérées pour l'exemple présenté dans la Figure 7.1.

7.2.4 *Gestion de l'absence d'activité*

Nos prédictions pour chaque période se fondent sur les traces laissées par les profils sur la plateforme. Il paraît probable que la présence de certains profils sur la plateforme soit sporadique et qu'ils ne soient pas actifs pour chaque période considérée. Dans ce cas là, l'absence d'activité d'un profil ne traduit pas nécessairement une perte d'intérêt pour l'événement et les points de vue qu'il engendre. Or avec notre implémentation actuelle, si un profil est inactif sur une période, le vecteur le caractérisant dans la matrice de voisinage sera

TABLEAU 7.1 – Points de vue prédits pour les profils A, B et C de notre exemple (Figure 7.1) sur les cinq périodes considérées au vu des scores présentés dans la Figure 7.3b.

	PÉRIODES				
PROFIL	1	2	3	4	5
A	Indéterminé	Indéterminé	Vert	Vert	Vert
B	Jaune	Jaune	Jaune	Jaune	Jaune
C	Vert	Jaune	Jaune	Jaune / Vert	Jaune / Vert

nul, donnant des scores d'expression égaux à 0 pour tous les points de vue et donc un point de vue indéterminé [1].

Afin de corriger ceci, nous allons considérer une fonction de similarité *récursive*, qui n'est mise à jour que lorsque de nouveaux indices sur le positionnement du profil sont disponibles :

$$\mathbf{\Sigma}_{i,j,p} = \begin{cases} \cos(\mathcal{M}^p_{i,\bullet}, \mathcal{M}^p_{\boldsymbol{\sigma}_j,\bullet}) & \text{si Activité}^p(v_i) > 0, \\ \mathbf{\Sigma}_{i,j,p-1} & \text{si Activité}^p(v_i) = 0 \text{ et } p > 1, \\ 0 & \text{sinon.} \end{cases}$$

avec $\text{Activité}^p(v_i)$ calculée comme suit :

$$\text{Activité}^p(v_i) = \sum_{j=1}^{|V|\times|X|} \mathcal{M}^p_{i,j}$$

Cette mesure traduit l'intensité de l'activité du profil v_i pendant la période p, et sera d'autant plus importante que le profil a été présent sur la plateforme.

Avec cette implémentation, les profils peuvent changer de points de vue si de nouveaux indices apparaissent mais, dans le cas contraire, conservent les derniers points de vue prédits pour eux. La limite découlant directement de ce fonctionnement est qu'une fois un point de vue déterminé pour un profil, nous considérons qu'il ne peut pas redevenir *indéterminé*.

Le Tableau 7.1 présente les points de vue prédits dans le cadre de notre exemple. Bien que certains profils aient été inactifs durant certaines périodes, comme illustré dans la Figure 7.1 et la Figure 7.3, nous constatons que les seuls points de vue *indéterminés* apparaissant sont dus à l'inactivité du profil A durant les deux premières périodes. Le profil B, visiblement inactif durant les périodes 2 et 3 car isolé dans les graphes de proximités, a néanmoins un point de vue prédit pour ces périodes. Nous pouvons aussi constater que le profil C exprime à la fois les points de vue `Jaune` et `Vert` durant les deux dernières

1. L'exception étant si nous incluons dans X certaines proximités qui ne dépendent pas directement de l'activité du profil, telles que la localisation géographique par exemple.

TABLEAU 7.2 – Effectifs des profils utilisés par point(s) de vue dans le jeu de données #Élysée2017fr (#E).
Les profils apparaissant sur la diagonale ne disposent que d'un seul point de vue.

POINTS DE VUE σ	FI	PS	EM	LR	FN	*Total*
France Insoumise (FI)	5 113	228	33	4	22	
Parti Socialiste (PS)		1 832	151	2	3	
En Marche ! (EM)			3 962	148	1	
Les Républicains (LR)				4 366	211	
Front National (FN)					3 376	
						19 452

périodes étant donné que ces points de vue ont le même score d'expression sur ces périodes (voir Figure 7.3b).

7.3 EXPÉRIMENTATIONS

Nous examinons dans cette section l'efficacité de notre modèle d'observation des points de vue, au travers d'évaluations qualitatives et quantitatives.

7.3.1 *Cadre expérimental*

Nous utilisons pour évaluer notre modèle le jeu de données #Élysée2017fr (#E), introduit dans le Chapitre 4, car il présente plusieurs avantages.

— Bien qu'ils ne soient pas datés, il s'agit du seul jeu de données présentant des points de vue multiples pour certains profils. De plus, les profils n'étant pas explicitement identifiés comme étant politicien·ne·s ou activistes, contrairement aux autres jeux de données, nous avons plus de chances d'observer des changements de points de vue.

— Il présente cinq points de vue, ce qui nous permettra de mobiliser des analyses plus fines qu'un jeu de données dual traditionnel.

— L'événement suivi étant la campagne présidentielle française de 2017, nous sommes bien plus à même de réaliser des évaluations qualitatives de qualité, le contexte étant familier.

Au vu de leur grande variabilité, les profils ayant un point de vue indéfini ne sont pas utilisés ici, il y a donc 19 452 profils dont les effectifs détaillés sont

présentés dans le Tableau 7.2 [2]. Comme présenté dans la Section 7.1, nous utilisons un profil-graine par point de vue. Nous utilisons ici deux ensembles de profils-graines différents :

— les comptes officiels des partis politiques
S_{partis} = (`@LePG`, `@partisocialiste`, `@enmarchefr`, `@lesRepublicains`, `@FN_officiel` [3]) ;

— les comptes officiels des candidat·e·s
$S_{\text{candidat·e·s}}$ = (`@JLMelenchon`, `@benoithamon`, `@EmmanuelMacron`, `@FrancoisFillon`, `@MLP_officiel`).

Nous nous concentrons ici sur la proximité *cite* analysée avec une périodicité journalière.

7.3.2 *Validation de l'hypothèse du voisinage*

7.3.2.1 *Évaluation des points de vue majoritaires*

Bien que notre modèle soit destiné à observer les changements de positionnement, la littérature actuelle indique que ces changements ne concerneront qu'une faible proportion des profils, comme l'ont observé Smith et al. (2013) dans leur étude. En effet si l'on observe les prédictions faites par notre modèle sur l'ensemble des périodes étudiées, seuls 19 % des profils expriment plus d'un point de vue au cours de la campagne avec S_{partis}, et 29 % avec $S_{\text{candidat·e·s}}$. Afin de valider notre hypothèse de similarité du voisinage, nous observons donc tout d'abord la correspondance entre les annotations manuelles et le point de vue majoritaire pour chaque profil. Nous considérons donc un seul point de vue par profil : le point de vue apparaissant le plus souvent sur *l'ensemble* des périodes considérées. Par exemple, si nous considérons un profil pour lequel notre modèle a prédit le point de vue `EM` sur 70 % des périodes et le point de vue `LR` sur les périodes restantes, le point de vue majoritaire considéré dans cette analyse sera `EM`.

Le Tableau 7.3 présente les performances en matière de précision, rappel et score F1. Pour cette section, les scores sont calculés pour chaque parti en considérant que les profils aux affiliations multiples expriment les deux points de vue indifféremment – par exemple un profil ayant comme annotation manuelle « `FI` / `PS` » sera considéré comme correctement catégorisé si son point de vue prédit est « `FI` » ou « `PS` ». Au niveau global, aucune différence n'est visible entre l'utilisation de $X = (cite_{all})$ et $X = (cite_{all}, cite_{rec})$, mais par contre S_{partis} obtient une précision légèrement meilleure que $S_{\text{candidat·e·s}}$, alors que

2. Ces effectifs diffèrent de ceux présentés dans le Tableau 5.2 car nous conservons ici les profils ayant des affiliations multiples, ce qui n'était pas le cas dans nos expérimentations précédentes.

3. Aujourd'hui renommé `@RNational_off` suite au changement de nom du *Front National* en *Rassemblement National*.

TABLEAU 7.3 – Précision, rappel et score F1 du modèle d'évolution des points de vue pour déterminer les points de vue majoritaires sur l'ensemble des périodes

		$X = (cite_{all})$			$X = (cite_{all}, cite_{rec})$		
		P	R	F1	P	R	F1
S_{partis}	FI	0,96	0,43	0,60	0,96	0,43	0,60
	PS	0,88	0,53	0,66	0,88	0,53	0,66
	EM	0,93	0,67	0,78	0,93	0,67	0,78
	LR	0,97	0,62	0,75	0,97	0,62	0,75
	FN	0,85	0,77	0,81	0,85	0,77	0,81
	Macro	0,92	0,60	0,72	0,92	0,60	0,72
$S_{\text{candidat·e·s}}$	FI	0,91	0,65	0,76	0,91	0,65	0,76
	PS	0,87	0,53	0,66	0,87	0,53	0,66
	EM	0,95	0,53	0,68	0,95	0,53	0,68
	LR	0,95	0,58	0,72	0,95	0,58	0,72
	FN	0,80	0,78	0,79	0,80	0,78	0,79
	Macro	0,90	0,61	0,72	0,90	0,61	0,72

cette dernière a un léger avantage en matière de rappel. La précision est dans l'ensemble excellente pour tous les partis, la précision minimale étant de 80 % pour le FN en utilisant $S_{\text{candidat·e·s}}$. Les scores de rappel sont significativement inférieurs – bien que néanmoins très satisfaisants pour des graines composées d'uniquement cinq profils, soit 0,03 % du jeu de donnés – oscillant entre 43 % et 78 %. Après une analyse plus poussée, il se trouve que ces relativement faibles taux de rappel s'expliquent en partie par les 32 % (resp. 28 %) de profils n'étant similaires à aucun point de vue avec S_{partis} (resp. $S_{\text{candidat·e·s}}$). Sans surprise, ces profils présentent une activité bien plus faible que les autres (voir Figure 7.4), 1 071 d'entre eux n'ayant même aucune activité, au niveau de la proximité *cite*, durant la campagne présidentielle.

7.3.2.2 *Évolution des métriques au cours de la campagne*

Afin d'affiner les observations précédentes, nous présentons dans la Figure 7.5 les résultats de la même analyse effectuée pour chaque période considérée, soit chaque jour du 25 novembre 2016 au 12 mai 2017. Sans surprise le rappel macro commence très bas, peu de profils étant actifs au début de la campagne, mais il augmente de façon constante pour atteindre 57 % pour $S_{\text{candidat·e·s}}$ et 58 % pour S_{partis} après le second tour. La précision macro reste relativement constante, oscillant à partir du cinquième jour entre 84 % et 90 % pour $S_{\text{candidat·e·s}}$ et entre 88 % et 92 % pour S_{partis}. Son écart-type diminue avec le temps, avant d'augmenter à nouveau durant les deux dernières semaines, probablement dû aux réorganisations à partir du premier tour. Les résultats entre les différentes configurations sont proches mais l'utilisation de S_{partis} est

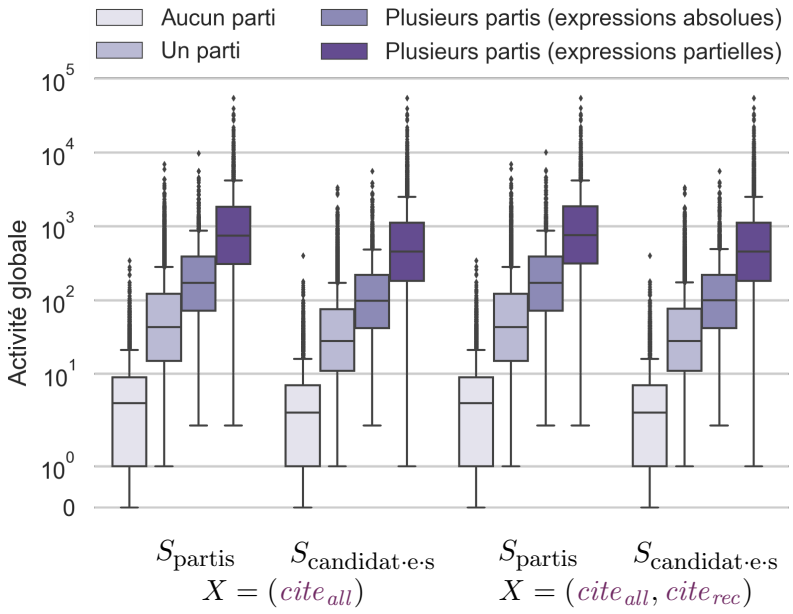


FIGURE 7.4 – Activité globale des profils en fonction du nombre de points de vue qu'ils expriment sur l'ensemble des périodes considérées et de la nature de ces expressions.
L'activité globale d'un profil v_i est la somme de son activité sur chaque période $p \in \Pi$, soit :

$$\text{Activité globale}(v_i) = \sum_{p \in \Pi} \text{Activité}^p(v_i) = \sum_{p \in \Pi} \sum \mathcal{M}^p_{i,\bullet}$$

Un profil exprimant plusieurs points de vue sur l'ensemble des périodes considérées peut exprimer un seul point de vue par période et changer de point de vue entre différentes périodes (expressions absolues), ou exprimer pendant une même période plusieurs points de vue simultanément (expressions partielles).

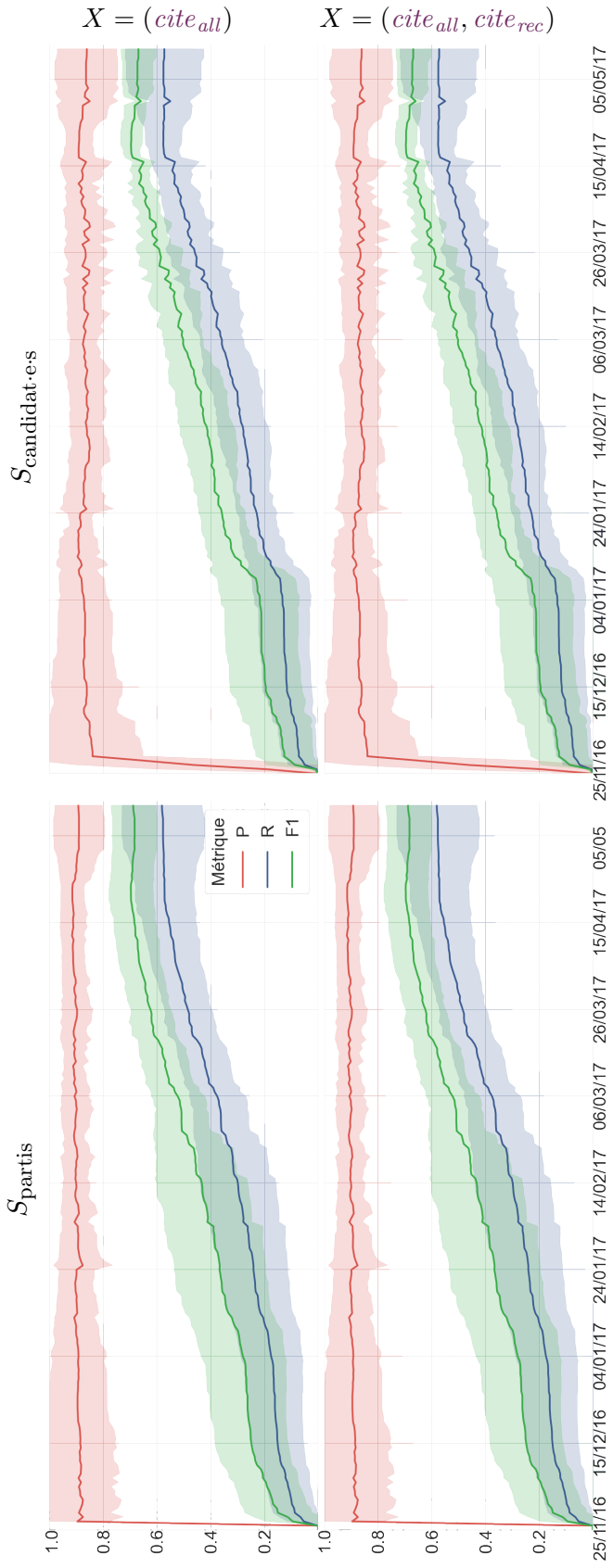


FIGURE 7.5 – Précision, rappel et score F1 du modèle d'évolution des points de vue au long de la campagne présidentielle.
Les lignes représentent la moyenne macro tous partis politiques confondus et la zone colorée l'écart-type.

légèrement meilleure, avec en moyenne 3 points de plus par jour pour toutes les métriques. En revanche, comme observé précédemment, la proximité $cite_{rec}$ n'apporte visiblement pas d'information complémentaire. Au vu de ces observations, nous nous concentrerons par la suite exclusivement sur la configuration utilisant $X = (cite_{all})$ et S_{partis}.

7.3.3 *Changements de points de vue*

Ce modèle étant principalement destiné à repérer des changements de points de vue, nous allons à présent nous concentrer sur cet aspect. Comme dans la section précédente, nous ne prenons en compte que le point de vue majoritaire pour chaque profil, ou les deux points de vue majoritaires en cas d'ex æquo.

7.3.3.1 *Évaluation qualitative*

La Figure 7.6 présente le diagramme alluvial des points de vue au cours de la campagne. Pour des raisons de lisibilité, les points de vue sont ici agrégés par semaine. Nous pouvons tout d'abord constater que les flux de changement de points de vue concernent une minorité des profils. La majorité de ces flux sont entre `LR` et `FN`, ce qui est consistant avec la proximité qu'ont entretenu ces partis dans leurs discours et thématiques. Les transferts de `LR` vers `FN` en semaines 17 et 18 sont particulièrement notables et s'expliquent aisément par les résultats du premier tour de l'élection ayant eu lieu le 23 avril, soit en fin de semaine 16. En effet, le candidat `LR`, François Fillon, ayant été éliminé lors de ce premier tour, de nombreux sympathisants `LR` ont déclaré reporter leur vote sur la candidate `FN`, Marine Le Pen. Ce report de voix vers le `FN` lors de l'entre-deux tours est également visible, bien que de façon beaucoup moins importante, depuis `FI`, tout particulièrement en semaine 18. À partir de la publication du programme d'`EM`, le 2 mars (semaine 9), les flux les plus importants de changements de points de vue se dirigent vers ce parti qui s'était établi comme le choix « de raison » (à l'exception des flux `LR`–`FN` évoqués précédemment). Les prédictions de notre modèle semblent donc être pertinentes au vu de la dynamique de la campagne.

7.3.3.2 *Évaluation quantitative*

Afin d'évaluer de façon plus poussée l'efficacité de notre modèle pour détecter les changements de points de vue, considérons à présent les 19 % de profils exprimant plus d'un point de vue au cours des périodes observées. Nous analysons leurs biographies Twitter afin d'obtenir la vérité terrain, de nombreux profils reflétant les fluctuations de leur appartenance politique dans celle-ci. Nous considérons qu'un profil soutient un parti dès lors qu'il l'exprime explicitement dans sa biographie, un profil pouvant soutenir plusieurs partis dans une même biographie. Ce soutien peut se traduire par un soutien direct au parti lui-même, à son / sa candidat·e ou à l'un de ses membres éminents. Le fait

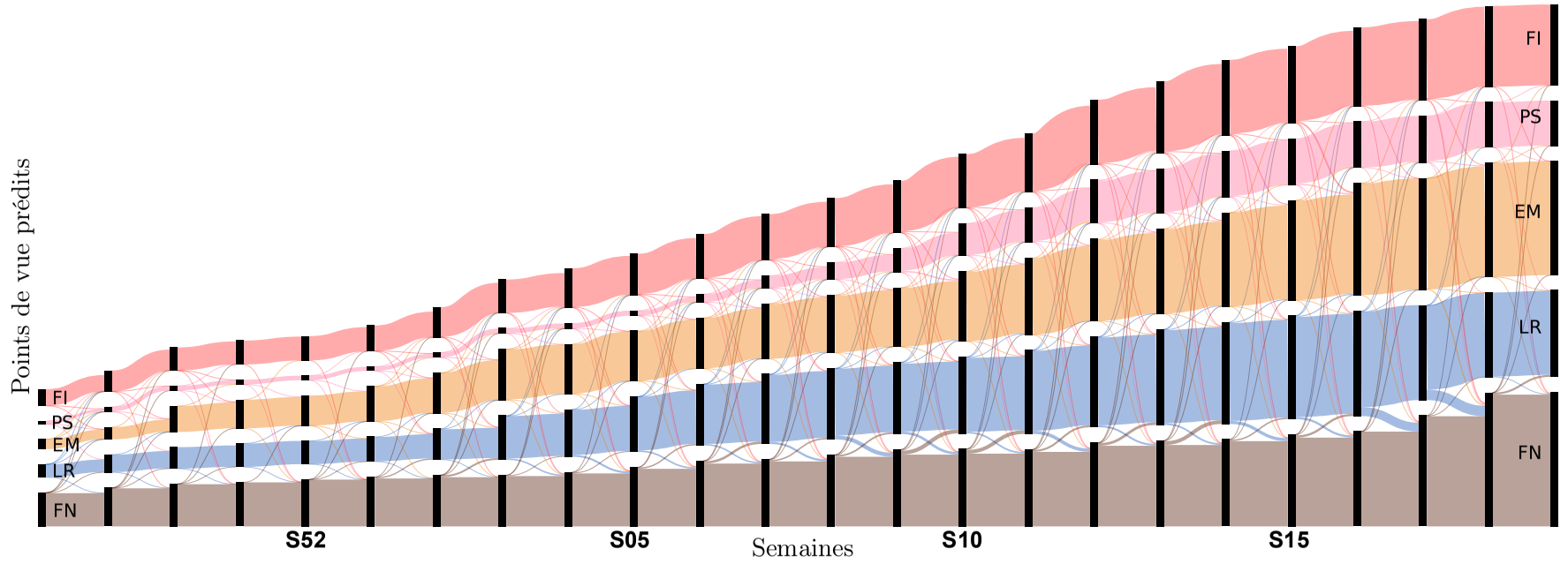


FIGURE 7.6 – Évolution des points de vue au cours de la campagne pour la graine S_{partis} et $X = (cite_{all})$.
La hauteur de chaque section représente le nombre de profil exprimant majoritairement le point de vue considéré pour une semaine donnée, et les flux entre partis représentent les profils changeant d'avis d'une semaine à l'autre.

Tableau 7.4 – Nombre de changements de point(s) de vue par profil.

NOMBRE DE CHANGEMENTS DE POINT(S) DE VUE	NOMBRE DE PROFILS
Pas de changement	2 699
1 changement	218
2 changements	51
3 changements	15
4 changements	3
5 changements	4
6 changements	1
7 changements	1
16 changements	1
TOTAL	2 993

Tableau 7.5 – Changements de points de vue déclarées par 2 993 profils dans leurs biographies Twitter.
Les effectifs indiqués en gras sont les profils n'ayant aucun changement de point de vue.

POINT DE VUE ORIGINEL	NOUVEAU POINT DE VUE																			
	FI	FI / PS	FI / PS / EM	FI / EM	FI / LR	FI / FN	PS	PS / EM	PS / LR	PS / LR / FN	PS / FN	EM	EM / LR	EM / FN	LR	LR / FN	FN
FI	**451**	6		3			2					6					3
FI / PS	2	**8**	1	1			5	2									1
FI / PS / EM	1	1															
FI / EM		1		**2**			1					1					
FI / LR																	
FI / FN	2																1
PS	11	11					**274**	71			1	31					
PS / EM							8	**5**				23					
PS / LR									**1**								
PS / LR / FN																	1
PS / FN																	1
EM	3			2			3	13				**460**	1		1		3
EM / LR												5	**8**	1	10	1	
EM / FN												1					
LR	1				1		1					20	28		**956**	45	20
LR / FN															20	**12**	9
FN	2					2	1				1	5			6	8	**536**

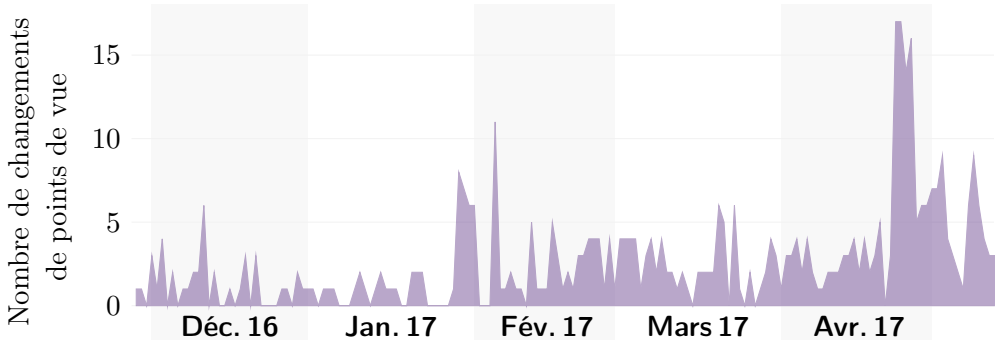


FIGURE 7.7 – Dates des changements de points de vue indiqués dans les biographies Twitter.

de nous appuyer sur les déclarations des profils mêmes, sans aucune inférence additionnelle, nous permet de réduire l'incertitude des annotations [4]. Nous obtenons ainsi les appartenances politiques déclarées de 2 993 profils, présentant au total 3 125 changements de points de vue, accompagnées de leurs dates de déclaration (voir Tableau 7.4, Tableau 7.5 et Figure 7.7). Nous pouvons déjà faire 2 observations :

— une large majorité des profils n'indique pas de changement de point de vue, suggérant que notre modèle est potentiellement trop sensible et détecte des changements de points de vue ponctuels incorrects ;

— les dates présentant le plus de changements de points de vue sont sans surprise corrélées aux événements marquants de la campagne, le pic fin janvier étant certainement causé par le Fillongate – avec notamment de nombreux supporters d'Alain Juppé passés `EM` à ce moment-là – et le pic fin avril au premier tour de l'élection.

Grâce à ces nouvelles informations, nous pouvons calculer pour chaque période les scores obtenus par notre modèle de façon similaire à l'analyse réalisée dans la Section 7.3.2.2. Il y a néanmoins un élément dont nous ne disposons pas : le délai existant entre le moment où un profil change de point de vue et le moment où il décide de mettre à jour sa biographie. Nous considérons donc différents délais dans les dates de déclarations, en simulant des dates de déclarations antérieures de 7, 15 ou 30 jours. Les résultats sont visibles dans la Figure 7.8.

La première constatation est que, comme nous le suspections, la précision de notre modèle sur ces profils est largement inférieure à la précision globale calculée lorsque nous prenons en compte la totalité des profils (voir Figure 7.5). Nous réussissons néanmoins à obtenir une précision allant de 60 % à 85 % avec uniquement cinq profils-graines. La deuxième constatation est que notre hypothèse sur le fait que les profils attendent avant de rendre public leur changement

4. Bien évidemment cela ne nous protège pas des profils exprimant un soutien par satire, mais cela représente généralement une minorité de profils

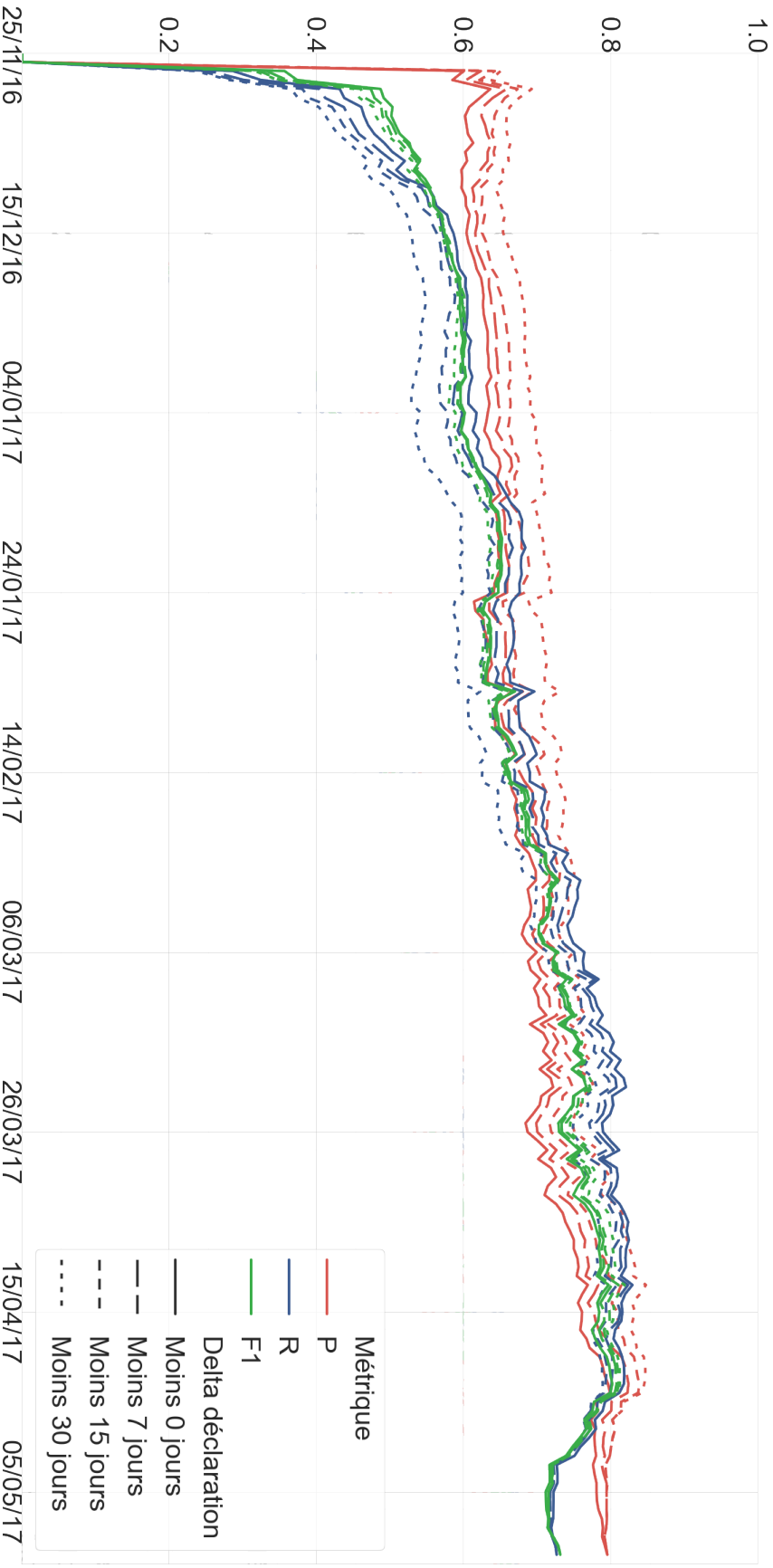


FIGURE 7.8 – Précision macro, rappel macro et score F1 macro des profils exprimant plus d'un point de vue avec le modèle d'évolution des points de vue avec S_{partis} et $X = (\textit{cite}_{all})$.

de point de vue semble se confirmer, la précision augmentant lorsque nous reculons la date théorique de déclaration des changements de points de vue des profils, la meilleure précision étant obtenue par une date théorique antérieure de 30 jours à la date réelle.

7.4 CONCLUSION

Nous avons présenté dans ce chapitre notre modèle d'observation de l'évolution temporelle des points de vue. Ce modèle est fondé sur les similarités de voisinage des profils. Cette représentation nous permet de calculer les similarités inter-profils à partir de seulement *cinq* profils-graines. Ce modèle nous a permis d'observer les points de vue des profils tout au long de la campagne présidentielle française de 2017, a notamment confirmé que la grande majorité des profils n'expriment pas de changement d'avis sur Twitter. Il nous a aussi permis d'observer que, lorsque les profils indiquent un changement de point(s) de vue dans leurs biographies, ce changement s'est déjà manifesté plusieurs jours avant dans le contenu des retweets.

Nous allons présenter dans le chapitre suivant la conclusion générale de notre mémoire abordant les perspectives à nos travaux ainsi que leurs limites.

8

CONCLUSION

8.1 SYNTHÈSE DES CONTRIBUTIONS

Nous avons présenté dans ce manuscrit deux contributions principales : un jeu de données original portant sur la campagne présidentielle française 2017 sur Twitter, présenté dans le Chapitre 4, et des modèles de détection de points de vue sur les médias sociaux fondés sur les proximités inter-profils (Partie II), dont les fondations sont introduites dans le Chapitre 5 et les détails présentés dans le Chapitre 6 et le Chapitre 7.

Pour notre première contribution, nous avons proposé un large et complexe jeu de données original de 22 853 profils Twitter impliqués dans la campagne de l'élection présidentielles française de 2017, annotés par des expert·e·s, et les 2 414 584 tweets et 7 763 931 retweets qu'ils ont publiés. Ces profils sont considérés comme ayant une orientation politique indéterminée ou sont affiliés à un ou plusieurs partis, parmi les cinq partis principaux ayant émergé durant la campagne – soient La France Insoumise (extrême gauche), le Parti Socialiste (gauche), En Marche ! (centre), les Républicains (droite) et le Front National (extrême droite). Nous fournissons également, lorsque nous pouvions le déterminer sans ambiguïté, la nature des profils (individuels ou collectifs) ainsi que le sexe de leurs propriétaires. En plus des multiples points de vue considérés, il s'agit à notre connaissance du premier jeu de données contenant un large nombre de profils et le premier à proposer des points de vue multiples pour un même profil, permettant des expérimentations plus poussées, telles que des détections de points de vue non binaires ou l'étude des électeurs·trices versatiles. Les premières analyses montrent que l'usage de Twitter varie grandement en fonction des partis, avec une large majorité de participants masculins dans les débats et de nombreux profils animés par des groupes de militant·e·s. Une grande partie des profils semblent avoir été créés spécifiquement pour la campagne, alors qu'une autre portion non négligeable avait vraisemblablement été créée pour l'élection présidentielle française précédente. Bien qu'il s'agisse d'un événement national, il a attiré une attention internationale non négligeable, particulièrement des voisins européens de la France, comme le montrent les nombreux pays et langues présents dans le jeu de données. Les réseaux de retweets et de mentions sont fortement polarisés entre partis, ce qui est confirmé par le faible nombre d'interactions inter-partis. Des analyses complémentaires sont nécessaires pour pleinement comprendre les mécanismes de la campagne, mais ce jeu de données est une ressource de valeur pour étudier le discours politique sur Twitter ou évaluer les outils automatiques sur cette plateforme.

Nous avons ensuite présenté deux modèles semi-supervisés génériques de détection de points de vue, fondés sur l'homogénéité et la complémentarité des communautés détectées sur des réseaux construits en utilisant diverses proximités textuelles, sociales et géographiques. Notre premier modèle, `SCSD`, peut aisément être personnalisé pour s'adapter à la majorité des besoins et s'appuie sur un très petit nombre de profils-graines afin de réduire les coûts d'annotation. Il procède itérativement à la propagation des points de vue dans les communautés en partant des proximités représentant les liens les plus forts entre les profils en matière de diffusion de points de vue.Les résultats de nos expérimentations suggèrent que `SCSD` réussit à fidèlement prédire les points de vue, même en n'utilisant pas plus de 1 % de données annotées et en considérant plus de deux points de vue. De plus, ils ont montré qu'étant donné que toutes les proximités diffèrent pour la diffusion de points de vue, l'utilisation de plusieurs proximités permettait de renforcer les prédictions.

L'analyse des différentes proximités est en accord avec les observations présentes dans la littérature : les profils tendent à construire leur discours sur les médias sociaux en partageant des arguments confortant leurs idées plutôt qu'en réfutant les arguments adverses. Il est intéressant de noter que même s'il arrive que les profils mentionnent leurs adversaires, ils tendent à répondre bien plus à leurs alliés. Ils tendent également à sélectionner avec soin les profils auxquels ils s'abonnent, bien qu'un autre comportement soit visible : certains profils, principalement des profils de politicien·ne·s ou d'activistes de haut rang, décident de s'abonner aussi bien à leurs adversaires qu'à leurs allié·e·s. Nous supposons qu'il s'agit d'une façon de surveiller leurs actions et leur discours, ce qui suggère que, sur Twitter, un abonnement n'est pas nécessairement une approbation du profil suivi. Ce phénomène apparaît notamment dans le jeu de données sur les élections de mi-mandat de 2014 aux États-Unis.

Notre second modèle permet d'étudier l'évolution temporelle des points de vue au cours d'un événement. Il s'appuie sur des similarités fondées sur le voisinage des profils, qui permettent de calculer les similarités des profils aux différents points de vue avec simplement un profil-graine par point de vue. Nos expérimentations ont montré que les vecteurs de voisinage se basant sur les retweets étaient très pertinents pour cette tâche et que le fait d'utiliser les profils officiels des partis politiques donnait des résultats meilleurs que l'utilisation des profils officiels de leurs candidat·e·s. Elles ont également mis en avant le fait que l'immense majorité des profils ne changeaient pas de point de vue, ou n'étaient pas assez actifs sur la plateforme pour qu'un changement de point de vue soit détectable. Parmi les profils changeant d'avis, la plupart semblent ne changer d'avis qu'une fois, les changements multiples étant rares au vu de nos observations. Un élément intéressant apparu lors de nos analyses est que les changements de points de vue sont visibles dans les retweets plusieurs jours avant que le profil ne l'annonce dans sa biographie.

Nous pensons que ces modèles pourraient être utiles aux chercheurs·ses en humanités numériques qui souhaiteraient explorer de larges jeux de données sans avoir les ressources pour les annoter manuellement.

8.2 PERSPECTIVES

8.2.1 *Utilisations possibles d'#Élysée2017fr*

Les analyses présentées dans le Chapitre 4 permettent d'avoir un aperçu des possibilités offertes par le jeu de données #Élysée2017fr. Nous présentons ici quelques idées d'investigations qui pourraient être réalisées.

— Tout d'abord, une analyse des réseaux de chaque parti pourrait déterminer quelles structures sont efficaces en matière de communication politique et améliorer la détection des profils d'*influenceurs·ses*.

— Les 2 414 584 tweets du jeu de données pourrait aussi être utilisés pour améliorer les outils de traitement automatique de la langue pour la langue française sur Twitter, étant donné que les tweets français représentent 93 % du corpus.

— Ces tweets seraient également tout à fait propices à la réalisation d'une fouille d'opinions à grande échelle : en effet, durant la campagne, les profils ont utilisé Twitter pour soutenir leur parti et leur candidat·e, mais également pour attaquer leurs adversaires. Trouver ces éléments de support et d'attaque, ainsi que leurs cibles, pourrait créer une carte intéressante de la campagne présidentielle, présentant les stratégies d'attaque et de défense des différents partis.

— Cette campagne a aussi été marquée par la propagation de rumeurs calomnieuses et de faux documents, en faisant un terrain parfait pour étudier les mécanismes et la propagation des campagnes de dénigrements. Il serait intéressant d'étudier cette propagation d'un point de vue structurel, afin d'examiner comment les rumeurs se sont propagées à travers les réseaux et quels étaient les nœuds importants, mais également d'un point de vue temporel pour voir l'influence des événements de la campagne sur la réception de ces rumeurs par les profils.

Nous ne présentons pas ici une liste exhaustive des possibilités offertes par ce jeu de données mais simplement des pistes possibles à explorer.

8.2.2 *Améliorations des modèles de détection de points de vue*

De nombreuses pistes d'améliorations pourraient être explorées afin de faire évoluer les modèles que nous avons présentés dans ce manuscrit. Nous présentons ci-après celles qui nous semblent les plus prometteuses.

PROXIMITÉS TEXTUELLES APPROFONDIES. Comme indiqué dans la Section 5.4.1, les proximités textuelles utilisées dans ce manuscrit sont limitées. L'une des premières améliorations à considérer serait probablement l'utilisation de proximités fondées sur des similarités textuelles plus fines, exploitant par exemple des plongements lexicaux ou d'autres outils permettant de capturer de façon plus efficace la sémantique des publications.

PRISE EN COMPTE DE LA NATURE DU PROFIL. Bien que nous disposions parfois des informations, nous n'avons pas utilisé la nature du profil lors de nos prédictions. Or, il paraît probable que les points de vue, ou les changements de points de vue, ne soient pas exprimés de la même façon par un profil individuel que par un profil institutionnel ou un profil de militant·e·s. Nous pourrions par exemple énoncer l'hypothèse qu'un profil géré par un groupe d'activistes a une probabilité moindre de changer de point de vue, en raison du niveau d'implication nécessaire pour être un·e militant·e actif·ve mais également pour des raisons d'image vis-à-vis de la cause ou de l'organisation défendue. Il pourrait donc être intéressant d'intégrer cette variable dans le modèle.

PRISE EN COMPTE DU SEXE. Pour des raisons similaires à la prise en compte de la nature du profil, la prise en compte du sexe des acteurs·trices pourrait apporter un éclairage nouveau lors de l'étude de certaines problématiques. En effet, de nombreux travaux ont mis en lumière des différences d'expression liées à la socialisation des individus (DERKS, FISCHER et BOS, 2008 ; KIVRAN-SWAINE et al., 2012).

INDICATEUR D'INCERTITUDE. Comme évoqué dans la Section 7.3, sur les médias sociaux, tous les profils ne sont pas suffisamment actifs pour permettre une prédiction de qualité. En effet, plus nous disposons d'informations laissées par le profil, mieux nous pouvons le catégoriser. L'indication d'un niveau d'incertitude permettrait d'indiquer à l'utilisateur·trice final·e si certaines prédictions sont fondées sur des informations limitées, ce qui lui permettrait par exemple de les filtrer s'il / elle le souhaite. Une solution serait d'associer à chaque profil un degré d'incertitude directement lié à son niveau d'activité ou encore à son degré d'appartenance aux communautés détectées, tels qu'utilisé par certains algorithmes de détection de communautés recouvrantes (AIROLDI et al., 2008 ; DUNN, 1973).

ASPECT GÉOGRAPHIQUE APPROFONDI. Nous n'avons utilisé dans ce travail que des proximités géographiques naïves mais il serait certainement pertinent d'utiliser des proximités géographiques plus poussées. L'une des difficultés principales de l'aspect géographique est le choix de granularités pertinentes vis-à-vis du sujet étudié, nécessitant souvent l'implication d'un expert du domaine et la capacité à gérer de nombreuses granularités à la fois.

LIEN TEMPOREL ENTRE LES PÉRIODES. Il peut arriver lors d'un événement qu'en raison de l'actualité, par exemple, un profil se rapproche ponctuellement d'un autre point de vue sans qu'il s'agisse d'un réel changement de point de vue. Hors, à l'heure actuelle, notre modèle d'évolution temporelle des points de vue considère chaque période de façon indépendante, détectant donc dans ces cas-là un changement de point de vue alors qu'il s'agit d'un intérêt passager. Il serait pertinent de lier les périodes afin d'assurer une certaine

constance dans le point de vue, le profil ne changeant ainsi de point d'un vue qu'à partir d'un certain seuil d'indices recueillis.

PRÉDICTION DES TRAJECTOIRES DES COMMUNAUTÉS. L'observation de l'évolution des points de vue par période nous apporte des informations intéressantes sur un événement a posteriori. Il pourrait également être intéressant d'ajouter un module de prédiction des trajectoires futures des communautés de points de vue afin d'étudier l'évolution du débat et les événements externes influençant les profils en temps réel.

UTILISATION DE PLUSIEURS PROFILS-GRAINES. Nous avons choisi dans le Chapitre 7 de ne considérer qu'un seul profil-graine par point de vue mais il serait tout à fait possible de rajouter au modèle une étape d'agrégation afin de prendre en compte plusieurs profils-graines. Cette étape permettrait de renforcer la robustesse du modèle en considérant l'ensemble des profils des représentant·e·s des points de vue étudiés. Dans le cas de la campagne présidentielle par exemple, nous pourrions prendre en compte les profils des membres éminents des partis n'étant pas candidats, comme Alain Juppé pour les Républicains ou Florian Philippot pour le Front National.

COMPARAISON AVEC DES MODÈLES ALTERNATIFS. Il serait également intéressant de comparer nos travaux avec des modélisations différentes des relations profils / partis. Nous pourrions par exemple étudier les différences avec la coupe minimum utilisée par Hasanuzzaman et al. (2016b), qui permettrait de modéliser à la fois l'appartenance aux partis et la distance entre profils, ou encore avec la méthode prétopologique mixte de Levorato (2010) qui pourrait mettre en évidence des propriétés de groupes de profils passées inaperçues auparavant. Une autre alternative serait l'utilisation de modèles de clustering multi-vues où chaque proximité pourrait être considérée comme une vue particulière du jeu de données étudié (Saha, Mitra et Kramer, 2018 ; Sublemontier, 2012).

8.3 LIMITES D'UTILISATION DE CES MODÈLES

8.3.1 *Polarisation*

Une critique potentielle à l'encontre de nos modèles concerne leur spécificité à des sujets fortement polarisés liés à la gouvernance d'un pays, au détriment de la généricité. Après tout, la majorité des travaux sur la détection de points de vue s'intéresse à des élections ou des partis politiques. Cependant, nous pensons qu'il s'agit d'une vision quelque peu réductrice. Tout d'abord, il est vrai que les sujets politiques sont des terrains pertinents à analyser avec nos modèles car ils sont très souvent polarisés. Mais si nous considérons la politique dans le sens original d'idées débattues par un groupe social, ces sujets sont loin d'être

restreints car ils englobent des thèmes extrêmement variés : désinformation, harcèlement, homéopathie, médecines alternatives, libre accès, économie, etc.

Enfin, nous pensons qu'il pourrait être plus utile d'examiner la pertinence d'utilisation de nos modèles sur un jeu de données en mesurant la controverse générée par le sujet considéré. En effet, la détection de points de vue est sans surprise proche des travaux sur la détection de controverse, la controverse naissant de la confrontation de points de vue divergeants. De nombreux travaux ont développé des mesures de controverses sur diverses plateformes. Nous pouvons notamment citer les travaux de Dori-Hacohen, Jensen et Allan (2016) et Vuong et al. (2008) fondés sur les collaborations et l'historique d'éditions pour détecter les articles Wikipedia controversiaux, Garimella et al. (2016) qui détectent les controverses sur Twitter à l'aide de marches aléatoires sur les graphes de retweets, ou encore la modélisation de la controverse de Jang, Dori-Hacohen et Allan (2017), utilisant deux axes orthogonaux : un pour le niveau de controverse et l'autre pour le niveau d'importance du sujet. L'utilisation de telles mesures de controverse pourrait justifier la pertinence de l'emploi de nos modèles ou au contraire indiquer que le sujet considéré n'est pas assez controversé pour garantir une détection fiable des points de vue sous-jacents.

8.3.2 *Biais des données collectées sur les médias sociaux*

Comme évoqué dans la Section 2.2.1, les données présentes sur les médias sociaux présentent de nombreux biais sur lesquels nous souhaitons revenir ici.

8.3.2.1 *Biais de représentativité*

Le premier biais est bien entendu celui de la représentativité des profils sur les médias sociaux. Comme l'énonce boyd (2010) :

> « Big Data presents new opportunities for understanding social practice. Of course the next statement must begin with a "but". And that "but" is simple: Just because you see traces of data doesn't mean you always know the intention or cultural logic behind them. And just because you have a big N doesn't mean that [the sample is] representative or generalizable. »

En effet, tout le monde n'utilise pas les médias sociaux. Le dernier rapport du Pew Reseach Center (Poushter, Stewart et Chwe, 2018) indique que sur les 39 pays étudiés, la médiane d'utilisation des médias sociaux est de 53 %, allant de 20 % en Tanzanie à 75 % en Jordanie. Ces proportions changent fortement en fonction des générations : les jeunes adultes sont en moyenne plus présents de 38 points sur les médias sociaux comparés aux anciennes générations de leurs pays respectifs, allant de 8 points en Jordanie à 58 points aux Philippines. Il note également que dans les pays émergents ou en voie de développement, les hommes ont plus tendance à utiliser les médias sociaux comparés aux femmes, alors que la tendance s'inverse dans les pays développés.

Il faut de plus noter que même lorsque les personnes sont présentes sur les médias sociaux, elles n'y sont pas nécessairement actives, ou ne s'expriment pas sur tous les sujets. Si nous prenons l'exemple de notre jeu de données #Élysée2017fr, COLLEONI, ROZZA et ARVIDSSON (2014) estiment qu'environ 10 % des tweets portent sur des sujets politiques. Cela signifie que nous avons capturé dans notre jeu de données une petite fraction des profils Twitter français [1], qui représentaient eux-même en 2013 environ 5 % de la population française (IPSOS, 2013). La démographie de ces profils est aussi très différente de la démographie de la population française globale : Twitter indique que la moitié de ses utilisateurs·trices français·e·s ont entre 25 et 39 ans, avec 54 % d'hommes et 34 % d'utilisateurs·trices appartenant aux classes supérieures (TWITTER MARKETING FR, 2016). En guise de comparaison, les hommes représentent 49 % de la population française globale, 18 % de la population a entre 25 et 39 ans et les classes supérieures ne représentent que 15 % de la population (INSEE, 2010, 2017).

8.3.2.2 *Influence des acteurs·trices*

MADLBERGER et ALMANSOUR (2014) ont également montré que de nombreux protagonistes jouaient un rôle dans la création et la collecte des données provenant des médias sociaux et pouvaient donc influencer les résultats des modèles prédictifs utilisant les données collectées :

1. Les profils, en décidant de publier ou non du contenu. Or, de nombreuses caractéristiques sociales ou environnementales peuvent faire varier la propension d'un profil à s'exprimer sur un média social. De plus, les profils peuvent tout à fait publier du contenu qu'ils savent incorrect ou parodique.
2. La plateforme en elle-même. En effet la plateforme évolue avec le temps, faisant de ce fait également évoluer l'expérience et l'approche des profils. Les publications disponibles à la collecte, et leur quantité, changent aussi selon la méthode considérée (sur Twitter par exemple, les méthodes de collectes peuvent être l'API Twitter Streaming, l'API Twitter REST, les fournisseurs externes, ...). Les même critères de recherche utilisés avec différentes méthodes de collecte résulteront en des jeux de données différents. Les plateformes de médias sociaux étant gérées par des entreprises privées, elles peuvent également à tout moment décider de changer les conditions d'accès à leurs données. Certains projets de recherche courent ainsi le risque de perdre du jour au lendemain l'accès aux données qui ont permis leur validation, parfois même avant d'avoir eu le temps de démarrer.
3. Les chercheurs·ses, à travers les nombreux choix faits lors de la collecte et du nettoyage du jeu de données. Si la collecte se fonde sur des mots-clés,

1. #Élysée2017fr contient aussi des profils étrangers mais ils représentent une extrême minorité.

comment être certain que les mots-clés choisis sont les plus pertinents ? Si le critère est géographique, la question de la pertinence se pose toujours, en plus de restreindre la collecte aux seules publications géotaggées (pour Twitter, cela représente environ 1,5 % des tweets). Enfin le choix de la période de collecte est un autre facteur de variabilité du jeu de données, et certains travaux ont montré à quel point un changement des dates de collecte pouvait modifier les résultats d'un modèle de prédictions (Jungherr, Jürgens et Schoen, 2012).

4. Les annotateurs·trices, lorsque les données collectées sont destinées à être utilisées par un modèle statistique supervisé. Les annotations servant à entraîner le modèle, elles influencent nécessairement les résultats finaux. Or, les annotations peuvent différer significativement d'une personne à l'autre, et la gestion de ces désaccords est donc un choix sensible.
5. D'autres acteurs peuvent encore interférer de façon ponctuelle, Madlberger et Almansour (2014) citent notamment l'exemple du blocage national de Twitter en mars 2014 par le gouvernement turc, qui a pu potentiellement impacter certaines collectes ayant lieu à ce moment-là.

Le travail de Madlberger et Almansour repose sur des données Twitter mais les rôles identifiés peuvent aisément se retrouver sur les autres plateformes de médias sociaux.

Notre but ici n'est bien évidemment pas de discréditer les travaux réalisés à l'aide de données collectées sur les médias sociaux, mais simplement de rappeler que de nombreux biais existent et qu'il est important de les garder à l'esprit et d'être prudent·e dans les observations qui en résultent (Boyadjian, Olivesi et Velcin, 2017 ; Shadowen, 2017). Il est notamment à notre avis imprudent d'extrapoler un résultat obtenu sur Twitter ou Facebook à l'ensemble de la population d'un pays : comme nous venons de le montrer, les différences sont bien trop nombreuses entre les deux populations, et les biais possibles trop nombreux, pour réaliser des comparaisons pertinentes. Ces données sont, par contre, un très bon matériau pour étudier les différences d'usages entre médias sociaux, ou examiner sur une même plateforme les différences entre plusieurs événements de nature semblable.

ANNEXES

A

PANORAMA DES MÉDIAS SOCIAUX

A.1 THE CONVERSATION PRISM

Lancée en 2008, The Conversation Prism [1] est une visualisation catégorisant les principaux médias sociaux en fonction de leur utilisation (voir Figure A.1.1 et Figure A.1.2).

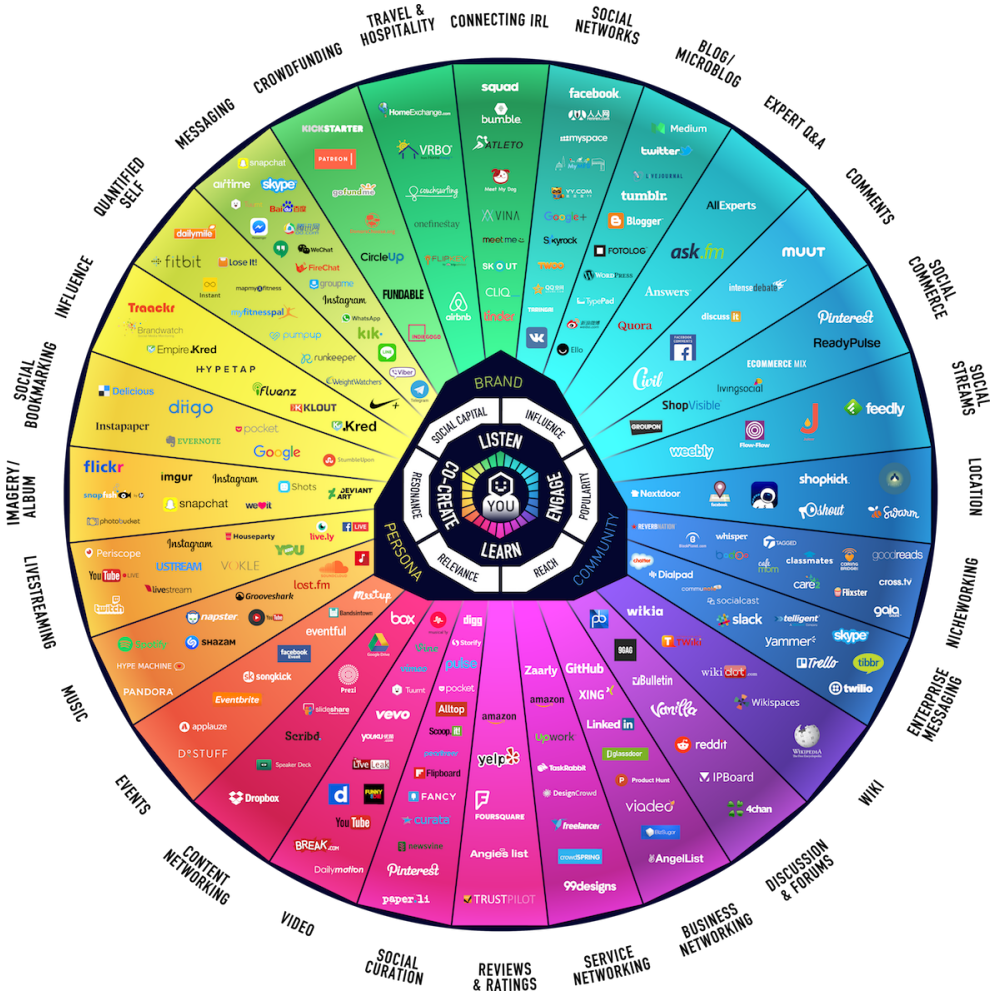


FIGURE A.1.1 – Carte du paysage des médias sociaux.

1. https://conversationprism.com

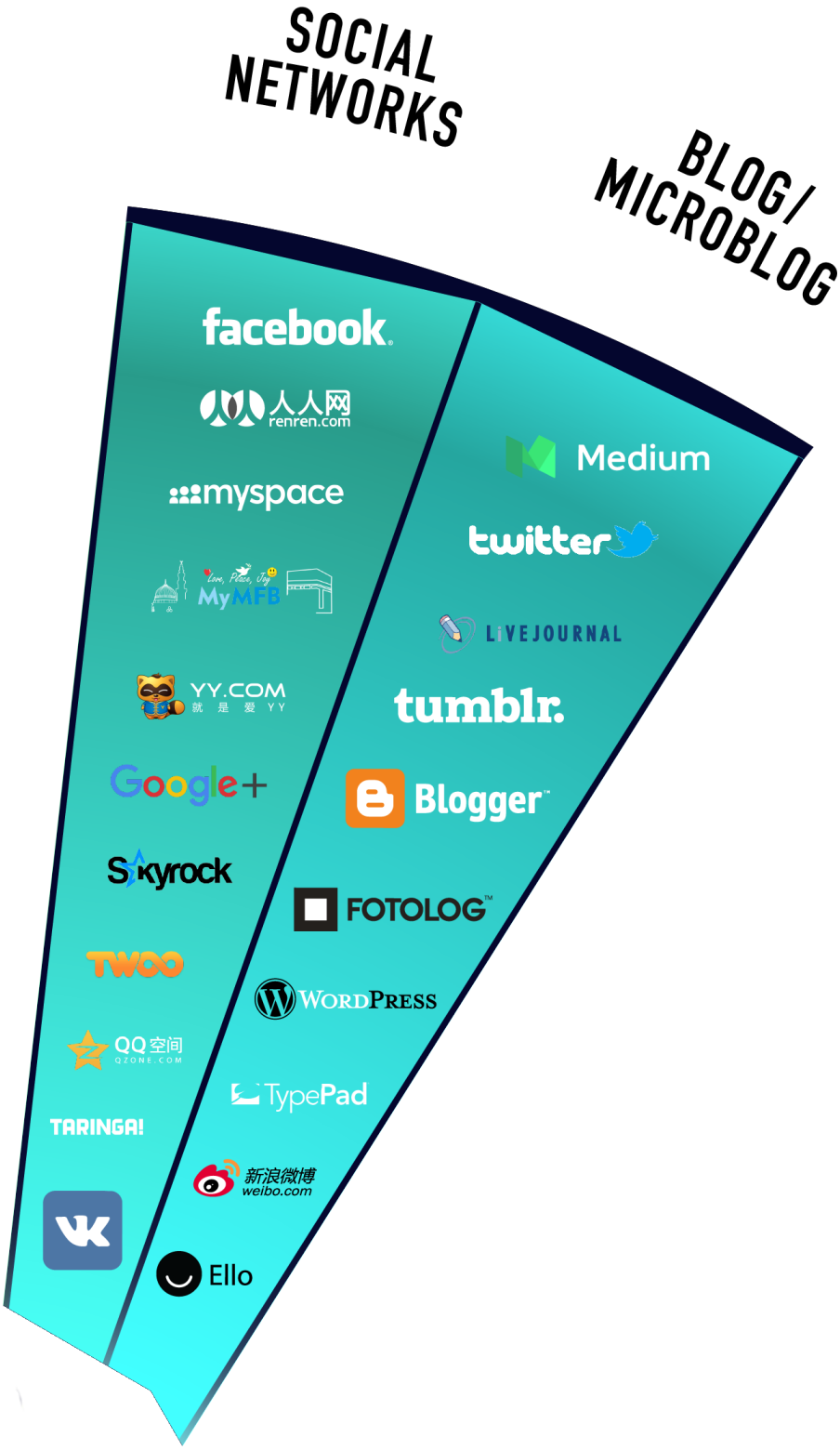


FIGURE A.1.2 – Zoom sur les réseaux sociaux et les services de blogs / micro-blogs.

A.2 AUDIENCE DES MÉDIAS SOCIAUX

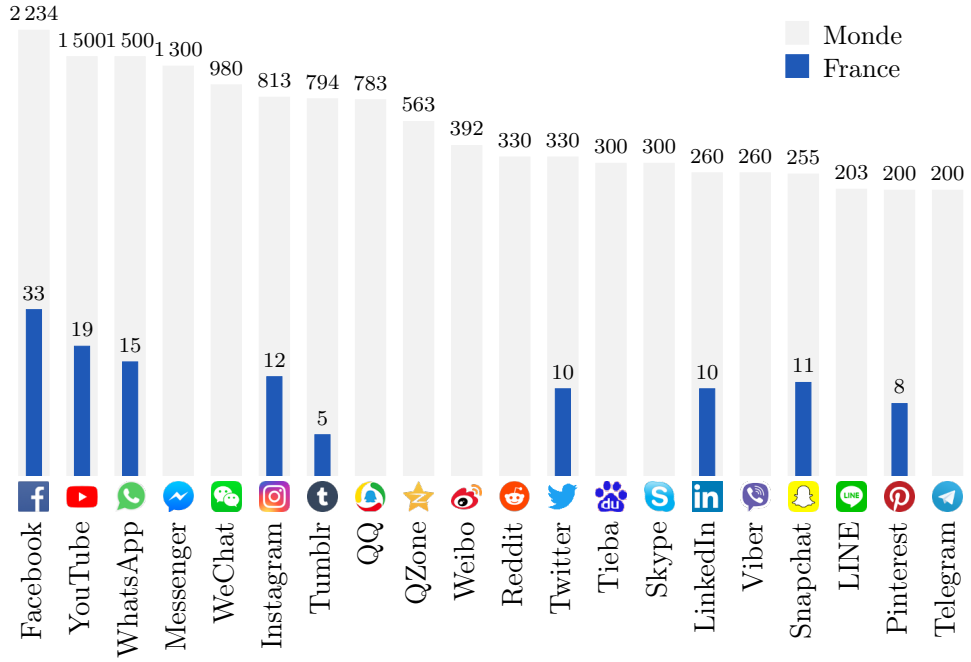


FIGURE A.2.1 – Nombres de profils actifs déclarés par les plateformes au niveau mondial et au niveau français (en millions et échelle logarithmique).
L'absence d'indication sur le nombre de profils français ne signifie pas nécessairement que la plateforme n'est pas présente en France mais que les données ne sont pas disponibles.

Il est très difficile d'obtenir des chiffres fiables sur la fréquentation des médias sociaux. Ces chiffres proviennent de deux types de sources :

— Ils peuvent être communiqués directement par la plateforme considérée, la plupart du temps sous forme de « nombre de profils actifs » ou de « nombre de visiteurs uniques par période ».

— Ils peuvent également être estimés par des entreprises de marketing numérique mesurant le traffic de nombreux sites web, dont la méthode de mesure est généralement opaque.

Les entités communiquant les chiffres d'audience ont donc toujours des intérêts économiques dépendant de ceux-ci. Il s'agit néanmoins des seules données dont nous disposons pour estimer l'audience globale des plateformes, les sondages ne permettant de cibler qu'une petite quantité d'utilisateurs.

La Figure A.2.1 présente les nombres de profils actifs déclarés par les plateformes au niveau mondial [2] et au niveau français [3]. Le Tableau A.2.1 présente les sites web les plus fréquentés au niveau mondial en matière de traffic, et le Tableau A.2.2 les sites web les plus fréquentés en France, selon les classements publiés par SimilarWeb [4] et par Alexa [5]. Il est aisé de constater que les sites possédant des fonctionnalités sociales y sont extrêmement présents.

TABLEAU A.2.1 – Sites web les plus fréquentés au niveau mondial en terme de traffic selon SimilarWeb et Alexa.
Les sites présentant des fonctionnalités sociales sont indiqués par un fond gris.

	RANG MONDIAL	
SITE	SimilarWeb	Alexa
google.com	1	1
youtube.com	2	2
facebook.com	3	3
baidu.com	4	4
yahoo.com	5	6
instagram.com	6	12
twitter.com	7	11
xnxx.com	8	
vk.com	9	17
wikipedia.org	10	5
xvideos.com	11	43
yandex.ru	12	25
pornhub.com	13	26
amazon.com	14	8
google.com.br	15	30
live.com	16	15

2. https://web.archive.org/web/20180710232150/https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

3. http://web.archive.org/web/20180723180847/http://www.alexitauzin.com/2013/04/combien-dutilisateurs-de-facebook.html

4. SimilarWeb est une entreprise israëlienne de marketing numérique fondée en 2007.
Classement mondial : https://www.similarweb.com/fr/top-websites
Classement français : https://www.similarweb.com/top-websites/france

5. Alexa est une entreprise américaine de marketing numérique du groupe Amazon, fondée en 1996. Son classement a notamment été utilisé par THELWALL (2009) pour étudier l'audience des médias sociaux en 2008.
Classement mondial : http://web.archive.org/web/20180723000004/https://www.alexa.com/topsites
Classement français : http://web.archive.org/web/20180510051234/https://www.alexa.com/topsites/countries/FR

	RANG MONDIAL	
SITE	SimilarWeb	Alexa
google.co.in	17	13
google.co.uk	18	29
xhamster.com	19	
ok.ru	20	
mail.ru	21	46
google.co.jp	22	21
googleweblight.com	23	
yahoo.co.jp	24	33
reddit.com	25	18
google.de	26	40
netflix.com	27	27
qq.com	28	7
google.fr	29	37
google.ru	30	28
ampproject.org	31	
ebay.com	32	39
google.com.mx	33	
google.com.tr	34	
google.it	35	
google.es	36	
bing.com	37	44
google.co.id	38	
sogou.com	39	
google.pl	40	
linkedin.com	41	32
google.ca	42	
msn.com	43	45
naver.com	44	48
pinterest.com	45	
taobao.com	46	9
whatsapp.com	47	
sm.cn	48	
yidianzixun.com	49	
tumblr.com	50	
tmall.com		10
sohu.com		14
jd.com		16

	RANG MONDIAL	
SITE	SimilarWeb	Alexa
sina.com.cn		19
weibo.com		20
360.cn		22
login.tmall.com		23
blogspot.com		24
google.com.hk		31
csdn.net		34
t.co		35
twitch.tv		36
alipay.com		38
pages.tmall.com		41
microsoft.com		42
wikia.com		47
aliexpress.com		49
imdb.com		50

TABLEAU A.2.2 – Sites web les plus fréquentés en France en terme de traffic selon SimilarWeb et Alexa.
Les sites présentant des fonctionnalités sociales sont indiqués par un fond gris.

	RANG FRANÇAIS	
SITE	SimilarWeb	Alexa
google.fr	1	1
google.com	2	2
facebook.com	3	4
youtube.com	4	3
leboncoin.fr	5	7
orange.fr	6	10
amazon.fr	7	6
live.com	8	8
wikipedia.org	9	5
yahoo.com	10	9
twitter.com	11	11
pornhub.com	12	22
free.fr	13	12
xnxx.com	14	
instagram.com	15	13
xhamster.com	16	27
xvideos.com	17	41
cdiscount.com	18	17
sfr.fr	19	28
lequipe.fr	20	34
lemonde.fr	21	23
lefigaro.fr	22	24
credit-agricole.fr	23	37
labanquepostale.fr	24	19
ebay.fr	25	20
pole-emploi.fr	26	26
netflix.com	27	21
jeuxvideo.com	28	31
20minutes.fr	29	
meteofrance.com	30	39
voirfilms.ws	31	36
msn.com	32	29

	RANG FRANÇAIS	
SITE	SimilarWeb	Alexa
francetvinfo.fr	33	47
news.google.com	34	
linkedin.com	35	15
programme-tv.net	36	44
pinterest.fr	37	48
vente-privee.com	38	
leparisien.fr	39	
impots.gouv.fr	40	25
qwant.com	41	
allocine.fr	42	30
pagesjaunes.fr	43	
purepeople.com	44	
laposte.net	45	
youporn.com	46	
tripadvisor.fr	47	
mappy.com	48	
bouyguestelecom.fr	49	
caf.fr	50	49
reddit.com		14
livejasmin.com		16
vk.com		18
zone-telechargement1.com		32
twitch.tv		33
t.co		35
blogspot.fr		38
aliexpress.com		40
bongacams.com		42
bing.com		43
microsoft.com		45
laposte.fr		46
wordpress.com		50

B

CORPUS #ÉLYSÉE2017FR

B.1 STRUCTURE DU JEU DE DONNÉES

Le jeu de données #Élysée2017fr est disponible à l'adresse suivante : `https://dataverse.mpi-sws.org/dataverse/icwsm18`.
Les annotations manuelles partagées dans `profiles_annotations.csv` sont détaillées dans le Tableau B.1.1. Les fichiers `posts_ids_*` contiennent les identifiants des tweets et des retweets, divisés en fonction du parti d'affiliations de leur auteur·e pour plus de flexibilité. Ces publications représentent une base de données relationnelle de 12Go une fois récupérés avec l'API Twitter. Les fichiers `networks_*` contiennent les réseaux de retweets et de mentions décrits dans la Section 4.5.4, au format NCOL et au format GraphML. Plus de détails sont disponibles dans le fichier `README`.

TABLEAU B.1.1 – Contenu du fichier `profiles_annotations.csv` pour chaque profil annoté.

Colonne	Contenu
`FROM_USER_ID`	L'identifiant Twitter du profil.
`PROFILE_NATURE`	*individual* si le profil est géré par une personne, sinon *non individual*. La catégorie *non individual* peut être *political* pour les partis ou associations politiques et les groupes de militant·e·s, *media* pour les entités médiatiques comme les journaux, ou *other*.
`PARTY`	Le(s) parti(s) politique(s) d'affiliation du profil (voir Section 4.4.2.2), separé par une ligne oblique (ex : « ps/fi »).
`MEDIA_PROFESSIONAL`	Indique si le profil s'identifie comme un·e professionel·le des medias (uniquement pour les profils gérés par une personne).
`SEX`	Indique le sexe du / de la propriétaire du profil : *m*, *f*, ou rien (uniquement pour les profils gérés par une personne).

B.2 DONNÉES DÉTAILLÉES

TABLEAU B.2.1 – Mots-clés utilisés lors de la collecte initiale des données

	MOT-CLÉ	DATE DE DÉBUT	DATE DE FIN
1	#6èmeRépublique	14/02/17	23/02/17
2	#AvenirEnCommun	14/02/17	23/02/17
3	#BenoitHamon2017	14/02/17	23/02/17
4	#FillonPresident	14/02/17	23/02/17
5	#FranceInsoumise	14/02/17	23/02/17
6	#Hamon	14/02/17	23/02/17
7	#Hamon2017	14/02/17	23/02/17
8	#JLM	14/02/17	23/02/17
9	#JLM2017	14/02/17	23/02/17
10	#JLMCHIFFRAGE	20/02/17	23/02/17
11	#Obsinsoumis	14/02/17	23/02/17
12	#PG	14/02/17	23/02/17
13	#avecFrancoisFillon	14/02/17	23/02/17
14	#avecHamon	14/02/17	23/02/17
15	#avecjadot	14/02/17	23/02/17
16	#fairebattrelecoeurdelafrance	14/02/17	23/02/17
17	#fillonistes	14/02/17	23/02/17
18	#insoumis	14/02/17	23/02/17
19	#jeanlucmelenchon	14/02/17	23/02/17
20	#lafrancenmarche	14/02/17	23/02/17
21	#lavenirencommun	14/02/17	23/02/17
22	#lfi	14/02/17	23/02/17
23	#macronlyon	14/02/17	23/02/17
24	#macronprésident	14/02/17	23/02/17
25	#mobilisationfillon	14/02/17	23/02/17
26	#teammacron	14/02/17	23/02/17
27	6èmerépublique	23/02/17	12/05/17
28	AuNomDuPeuple	25/11/16	23/02/17
29	AvenirEnCommun	23/02/17	12/05/17
30	LaPrimaire	25/11/16	12/05/17
31	Obsinsoumis	23/02/17	12/05/17
32	PG	23/02/17	12/05/17
33	PS election	25/11/16	12/05/17

	MOT-CLÉ	DATE DE DÉBUT	DATE DE FIN
34	PS elections	25/11/16	12/05/17
35	PS presidentelle	25/11/16	12/05/17
36	PS presidentielles	25/11/16	12/05/17
37	Primaire	25/11/16	12/05/17
38	PrimaireDeGauche	25/11/16	12/05/17
39	alleznathalie	25/11/16	12/05/17
40	alliot-marie election	25/11/16	12/05/17
41	alliot-marie elections	25/11/16	12/05/17
42	alliot-marie presidentielle	25/11/16	12/05/17
43	alliot-marie presidentielles	25/11/16	12/05/17
44	arthaud election	25/11/16	12/05/17
45	arthaud elections	25/11/16	12/05/17
46	arthaud presidentielle	25/11/16	12/05/17
47	arthaud presidentielles	25/11/16	12/05/17
48	arthaud2017	25/11/16	12/05/17
49	avecfillon	23/02/17	12/05/17
50	avecfrancoisfillon	23/02/17	12/05/17
51	avechamon	23/02/17	12/05/17
52	avecjadot	23/02/17	12/05/17
53	avecjlm	23/02/17	12/05/17
54	avecmacron	23/02/17	12/05/17
55	avecmarine	20/03/17	12/05/17
56	bayrou election	25/11/16	12/05/17
57	bayrou elections	25/11/16	12/05/17
58	bayrou presidentielle	25/11/16	12/05/17
59	bayrou presidentielles	25/11/16	12/05/17
60	benoithamon	23/02/17	12/05/17
61	benoithamon2017	23/02/17	12/05/17
62	bh2017	23/02/17	12/05/17
63	cheminade election	25/11/16	12/05/17
64	cheminade elections	25/11/16	12/05/17
65	cheminade presidentielle	25/11/16	12/05/17
66	cheminade presidentielles	25/11/16	12/05/17
67	cheminade2017	25/11/16	12/05/17
68	dupont-aignan	25/11/16	12/05/17
69	eelv	25/11/16	12/05/17
70	em2017	25/11/16	12/05/17
71	enbleumarine	23/02/17	12/05/17

	MOT-CLÉ	DATE DE DÉBUT	DATE DE FIN
72	enmarche	25/11/16	12/05/17
73	etsicetaitm	25/11/16	12/05/17
74	fairebattrelecoeur	23/02/17	12/05/17
75	fairebattrelecoeurdelafrance	23/02/17	12/05/17
76	fillon election	25/11/16	12/05/17
77	fillon elections	25/11/16	12/05/17
78	fillon presidentielle	25/11/16	12/05/17
79	fillon presidentielles	25/11/16	12/05/17
80	fillon	25/11/16	12/05/17
81	fillon2017	25/11/16	12/05/17
82	fillonistes	23/02/17	12/05/17
83	fillonpresident	23/02/17	12/05/17
84	fn election	25/11/16	12/05/17
85	fn elections	25/11/16	12/05/17
86	fn presidentielle	25/11/16	12/05/17
87	fn presidentielles	25/11/16	12/05/17
88	francebleumarine	23/02/17	12/05/17
89	franceenmarche	23/02/17	12/05/17
90	franceenordre	23/02/17	12/05/17
91	franceinsoumise	23/02/17	12/05/17
92	francois bayrou	25/11/16	12/05/17
93	futurdesirable	23/02/17	12/05/17
94	gaino2017	25/11/16	12/05/17
95	guaino election	25/11/16	12/05/17
96	guaino elections	25/11/16	12/05/17
97	guaino presidentielle	25/11/16	12/05/17
98	guaino presidentielles	25/11/16	12/05/17
99	hamon	23/02/17	12/05/17
100	hamon2017	23/02/17	12/05/17
101	hg2017	25/11/16	12/05/17
102	humaindabord	23/02/17	12/05/17
103	insoumis	23/02/17	12/05/17
104	jadot	25/11/16	12/05/17
105	jadot2017	25/11/16	12/05/17
106	jc2017	25/11/16	12/05/17
107	jean lassale	25/11/16	12/05/17
108	jeanlasalle2017	25/11/16	12/05/17
109	jeanlucmelenchon	23/02/17	12/05/17

	MOT-CLÉ	DATE DE DÉBUT	DATE DE FIN
110	jeunesavecfillon	23/02/17	12/05/17
111	jevotefillon	23/02/17	12/05/17
112	jevotehamon	23/02/17	12/05/17
113	jevotejadot	23/02/17	12/05/17
114	jevotemarine	23/02/17	12/05/17
115	jevotemelenchon	23/02/17	12/05/17
116	jlm	23/02/17	12/05/17
117	jlm2017	25/11/16	14/02/17
118	jlmchiffrage	23/02/17	12/05/17
119	juppe election	25/11/16	12/05/17
120	juppe elections	25/11/16	12/05/17
121	juppe presidentielle	25/11/16	12/05/17
122	juppe presidentielles	25/11/16	12/05/17
123	juppe	25/11/16	12/05/17
124	juppe2017	25/11/16	12/05/17
125	lafrancenmarche	23/02/17	12/05/17
126	lasalle election	25/11/16	12/05/17
127	lasalle elections	25/11/16	12/05/17
128	lasalle presidentielle	25/11/16	12/05/17
129	lasalle presidentielles	25/11/16	12/05/17
130	lavenirencommun	23/02/17	12/05/17
131	lepionmacron	23/02/17	12/05/17
132	lesrepublicains	25/11/16	12/05/17
133	lfi	23/02/17	12/05/17
134	macron	25/11/16	12/05/17
135	macron2017	25/11/16	12/05/17
136	macrondegage	23/02/17	12/05/17
137	macronlyon	23/02/17	12/05/17
138	macronprésident	23/02/17	12/05/17
139	mam2017	25/11/16	12/05/17
140	marine2017	25/11/16	12/05/17
141	melenchon	25/11/16	12/05/17
142	mlp	20/03/17	12/05/17
143	mlp2017	20/03/17	12/05/17
144	mobilisationfillon	23/02/17	12/05/17
145	nadot election	25/11/16	12/05/17
146	nadot elections	25/11/16	12/05/17
147	nadot presidentielle	25/11/16	12/05/17

	MOT-CLÉ	DATE DE DÉBUT	DATE DE FIN
148	nadot presidentielles	25/11/16	12/05/17
149	nadot	25/11/16	12/05/17
150	nda2017	25/11/16	12/05/17
151	patriote	20/03/17	12/05/17
152	planb	01/03/17	12/05/17
153	poutou2017	25/11/16	12/05/17
154	presidentielle2017	25/11/16	12/05/17
155	presidentielles2017	25/11/16	12/05/17
156	primaireLeDebat	25/11/16	12/05/17
157	primairesPS	25/11/16	12/05/17
158	teammacron	23/02/17	12/05/17
159	tousavecseb	25/11/16	12/05/17

TABLEAU B.2.2 – Pays présents dans #Élysée2017fr par parti.

PAYS	FI	PS	EM	LR	FN	Reste
France	2689	1090	2704	2848	1723	1899
États-Unis	26	8	35	21	69	35
Royaume-Uni	30	6	40	18	22	25
Belgique	23	1	15	19	14	27
Canada	16	4	15	10	13	13
Espagne	12	1	13	6	13	25
Suisse	11	4	10	10	6	10
Italie	2	1	10	6	13	13
Allemagne	8	1	6	3	3	8
Maroc	6	4	4	1	2	23
Pays-Bas	1		3	1	10	3
Russie	4	1			7	2
Brésil	4		2	3	3	1
Irlande	2	2	3	3	1	4
Australie	3		4	2	2	2
Liban	1	1	1	4	3	8
Suède	5	1	2		1	
Japon	5	1	2		1	1
Chine			5	3		
Grèce	2	1	2	1	1	15
Portugal	2		3		2	2
Vietnam				6		1
Luxembourg	1		1	4		2
Norvège		2	1		2	
Tunisie	1	2	1	1		10
Côte d'Ivoire			3	2		1
Turquie	2	1	2			4
Pologne	1			2	2	1
Mexique		2		2		1
République Tchèque			1	2	1	1
Singapour			1	2	1	
Slovénie				1	3	
Sénégal	2		2			2
Égypte	2			1	1	2
Argentine	2		2			1
Thaïlande	2			1	1	1

PAYS	FI	PS	EM	LR	FN	Reste
Roumanie	2			2		
Chili	1			1	2	
Israël	1		1		2	10
Burkina Faso	1		2			
Irak	1	1		1		
Algérie	2					17
Mayotte	2					
Cameroun			1	1		1
Indonésie			1		1	1
Danemark			2			1
Kenya			2			
Nouvelle-Zélande					2	
Zimbabwe	1					
Népal	1					1
Lituanie	1					1
Madagascar			1			
Guinée		1				
Chypre			1			1
Mongolie				1		
Mali	1					
Birmanie					1	
Rwanda		1				
Corée du Sud			1			1
Malaisie	1					
Haïti	1					1
Pérou			1			2
Inde			1			3
Nigéria			1			1
Libye					1	
Bulgarie			1			
Albanie	1					
Cuba	1					
Finlande					1	1
Croatie	1					
Colombie	1					1
Autriche		1				
Slovaquie					1	
Émirats Arabes Unis			1			8

PAYS	FI	PS	EM	LR	FN	Reste
Philippines				1		1
Afrique du Sud			1			1
Géorgie				1		
Vénézuéla				1		
Arabie Saoudite					1	1
Islande	1					
Syrie						1
Yémen						1
Équateur						1
République Dominicaine						2
Iran						1
Koweït						1
Qatar						2
Malte						1
Pakistan						3

TABLEAU B.2.3 – Langues principales d'#Élysée2017fr.

LANGUE PRINCIPALE	Nombre de profils
Français	22238
Anglais	525
Espagnol	76
Italien	32
Grec	22
Allemand	14
Arabe	11
Néerlandais	10
Portugais	10
Polonais	9
Suédois	5
Tchèque	4
Japonais	4
Roumain	3
Slovène	3
Turc	3
Danois	2
Persan	2
Russe	2
Finnois	1
Hongrois	1

TABLEAU B.2.4 – Langues secondaires d'#Élysée2017fr.

LANGUE SECONDAIRE	Nombre de profils
Français	9754
Anglais	3934
Espagnol	561
Italien	229
Portugais	127
Allemand	107
Estonien	67
Roumain	50
Danois	43
Néerlandais	42
Hongrois	41
Gallois	26
Suédois	26
Turc	26
Arabe	22
Tchèque	20
Grec	18
Letton	17
Polonais	16
Lituanien	15
Finnois	12
Norvégien	12
Japonais	6
Russe	5
Slovène	4
Persan	3
Hindi	3
Chinois	2
Bulgare	1
Islandais	1
Coréen	1
Ourdou	1
Vietnamien	1

TABLEAU B.2.5 – Partis d'affiliation des profils d'#Élysée2017fr dont la langue principale n'est pas le français.

LANGUE PRINCIPALE	FI	PS	EM	LR	FN	Reste
Anglais	50	9	94	10	161	201
Espagnol	8	1	7	5	17	38
Italien	1		3	1	19	8
Grec	3		1		1	17
Allemand	1		2		5	6
Arabe						11
Portugais	1		1	1	5	2
Néerlandais			1		4	5
Polonais				1	4	4
Suédois		1			2	2
Japonais	1				1	2
Tchèque			1			3
Slovène					3	
Roumain		1		1		1
Turc						3
Russe					2	
Danois			1			1
Persan	1					1
Finnois						1
Hongrois						1

Liste des figures

FIGURE 1.1 Nombre de publications mentionnant les termes « Social media », « Social network », « Facebook » ou « Twitter » référencés dans la bibliothèque numérique de l'Association for Computing Machinery (ACM) 2
FIGURE 2.1 Exemple d'un profil Twitter. 12
FIGURE 2.2 Exemple d'un extrait de fil Twitter. 13
FIGURE 2.3 Évolution des communautés politiques durant la campagne présidentielle française de 2017 (GAUMONT, PANAHI et CHAVALARIAS, 2017). 32
FIGURE 3.1 Exemples de graphes. 36
FIGURE 3.2 Exemples de matrices d'adjacences. 36
FIGURE 3.3 Exemples de communautés. 36
FIGURE 4.1 Exemple de présentation d'un profil Twitter lors de la tâche d'annotation. 48
FIGURE 4.2 Formulaire d'annotation d'un profil Twitter. 49
FIGURE 4.3 Part de profils, tweets et retweets par parti. 54
FIGURE 4.4 Distribution des natures de profils par parti. 54
FIGURE 4.5 Nombre de profils créés par mois. 55
FIGURE 4.6 Évolution du nombre de tweets et de retweets publiés . 56
FIGURE 4.7 Réseau des retweets et des mentions. 61
FIGURE 5.1 Information mutuelle normalisée entre les communautés des proximités. 79
FIGURE 6.1 Illustration des étapes d'initiation de `SCSD`. 86
FIGURE 6.2 Illustration du processus itératif d'assignation des points de vue de `SCSD`. 87
FIGURE 6.3 Corrélation de Spearman entre les rangs des profils retournés par différentes fonctions d'importance φ. . . . 91
FIGURE 6.4 Influence des fonctions d'ordonnancement ω sur la précision et le rappel du modèle `SCSD`. 94
FIGURE 7.1 Exemple d'évolution des graphes de proximités. 103
FIGURE 7.2 Extrait de la matrice de voisinage pour la période $p = 1$. 104
FIGURE 7.3 Scores d'expression des point de vue `Jaune` (`J`) et `Vert` (`V`) par les profils A, B et C pendant les cinq périodes considérées pour l'exemple présenté dans la Figure 7.1. 105
FIGURE 7.4 Activité globale des profils en fonction du nombre de points de vue qu'ils expriment sur l'ensemble des périodes considérées et de la nature de ces expressions. . 110

FIGURE 7.5 Précision, rappel et score F1 du modèle d'évolution des points de vue au long de la campagne présidentielle. . . 111
FIGURE 7.6 Évolution des points de vue au cours de la campagne . 113
FIGURE 7.7 Dates des changements de points de vue indiqués dans les biographies Twitter. 115
FIGURE 7.8 Précision macro, rappel macro et score F1 macro des profils exprimant plus d'un point de vue avec le modèle d'évolution des points de vue avec S_{partis} et $X = (cite_{all})$.116
FIGURE A.1.1 Carte du paysage des médias sociaux. iii
FIGURE A.1.2 Zoom sur les réseaux sociaux et les services de blogs / microblogs. iv
FIGURE A.2.1 Nombres de profils actifs déclarés par les plateformes au niveau mondial et au niveau français. v

Liste des tableaux

TABLEAU 2.1 Recouvrements lexicaux entre les termes « fouille de points de vue », « fouille d'opinions », « analyse de sentiments » et « analyse d'émotions ». 16
TABLEAU 2.2 Médias sociaux utilisés pour évaluer les modèles de détection de points de vue. 19
TABLEAU 2.3 Thèmes utilisés pour évaluer les modèles de détection de points de vue 20
TABLEAU 2.4 Synthèse de la littérature existant sur la détection des points de vue statiques 30
TABLEAU 2.5 Jeux de données existant sur le thème des points de vue politiques au niveau profil sur Twitter 34
TABLEAU 3.1 Caractéristiques des algorithmes de détection de communautés non recouvrantes, avec $n = |V|$ et $m = |E|$. . 39
TABLEAU 4.1 Exemples de profils Twitter présentant des affiliations politiques multiples 48
TABLEAU 4.2 Nombre de profils par parti politique. 53
TABLEAU 4.3 Nombre médian et moyen de publications par profil et par parti. 54
TABLEAU 4.4 Pays présents dans #Élysée2017fr par parti. 57
TABLEAU 4.5 Langues principales et secondaires utilisées par les profils d'#Élysée2017fr. 58
TABLEAU 4.6 Partis d'affiliation des profils dont la langue principale n'est pas le français. 59
TABLEAU 4.7 Caractéristiques globales des réseaux de retweets et de mentions. 59
TABLEAU 4.8 Détails sur les nœuds et les arêtes des réseaux de retweets et de mentions. 59
TABLEAU 4.9 Assortativité des réseaux de retweets et de mentions sur le critère du parti politique 60
TABLEAU 4.10 Nombre moyen de retweets / mentions par profil en fonction des affiliations politiques des profils impliqués. 62
TABLEAU 5.1 Notations utilisées. 69
TABLEAU 5.2 Jeux de données utilisés lors des expérimentations . . . 70
TABLEAU 5.3 Nombre d'arêtes $|E_i|$ dans G_i 74
TABLEAU 5.4 Transitivité des graphes de proximité. 76
TABLEAU 5.5 Pureté moyenne des communautés contenant au moins trois profils dans V_T 78

Tableau 5.6 Pureté moyenne des 10 plus grosses communautés contenant au moins cinq profils dans V_T 78
Tableau 6.1 Scores du modèle `SCSD` en utilisant différentes fonctions d'importance φ 92
Tableau 6.2 Scores du modèle `SCSD-Basic` 93
Tableau 6.3 Influence des fonctions d'ordonnancement ω sur le score F1 du modèle `SCSD`. 95
Tableau 6.4 Influence de s et s_{com} sur les scores F1 de `SCSD`. 96
Tableau 6.5 Scores des modèles de base et du modèle `SCSD` avec la configuration optimale par collection. 97
Tableau 7.1 Points de vue prédits pour les profils A, B et C de notre exemple (Figure 7.1) sur les cinq périodes considérées au vu des scores présentés dans la Figure 7.3b. 106
Tableau 7.2 Effectifs des profils utilisés par point(s) de vue dans le jeu de données #Élysée2017fr 107
Tableau 7.3 Précision, rappel et score F1 du modèle d'évolution des points de vue pour déterminer les points de vue majoritaires sur l'ensemble des périodes 109
Tableau 7.4 Nombre de changements de point(s) de vue par profil. . 114
Tableau 7.5 Changements de points de vue déclarées par 2 993 profils dans leurs biographies Twitter 114
Tableau A.2.1 Sites web les plus fréquentés au niveau mondial en terme de traffic, selon SimilarWeb et Alexa vi
Tableau A.2.2 Sites web les plus fréquentés en France en terme de traffic, selon SimilarWeb et Alexa ix
Tableau B.1.1 Contenu du fichier `profiles_annotations.csv` pour chaque profil annoté. xi
Tableau B.2.1 Mots-clés utilisés lors de la collecte initiale des données xii
Tableau B.2.2 Pays présents dans #Élysée2017fr par parti. xvii
Tableau B.2.3 Langues principales d'#Élysée2017fr. xx
Tableau B.2.4 Langues secondaires d'#Élysée2017fr. xxi
Tableau B.2.5 Partis d'affiliation des profils d'#Élysée2017fr dont la langue principale n'est pas le français. xxii

Liste des algorithmes

Algorithme 1 Fonctionnement global de SCSD. 84
Algorithme 2 Détermination du point de vue majoritaire. 85
Algorithme 3 Sélection des profils-graines 89

BIBLIOGRAPHIE

Abbasi, Mohammad Ali, Reza Zafarani, Jiliang Tang et Huan Liu (2014). « Am i More Similar to My Followers or Followees ? : Analyzing Homophily Effect in Directed Social Networks ». In : Proceedings of the 25th ACM Conference on Hypertext and Social Media. ACM Press, p. 200-205. doi : 10.1145/2631775.2631828 (cf. p. 19, 20, 25, 30).

Abbott, Rob, Brian Ecker, Pranav Anand et Marilyn Walker (2016). « Internet Argument Corpus 2.0 : An SQL Schema for Dialogic Social Media and the Corpora to Go with It ». In : Proceedings of the Tenth International Conference on Language Resources and Evaluation. ELRA, p. 23-28. isbn : 978-2-9517408-9-1 (cf. p. 70, 71).

Abdaoui, Amine, Jérôme Azé, Sandra Bringay et Pascal Poncelet (2017). « FEEL : A French Expanded Emotion Lexicon ». In : Language Resources and Evaluation 51.3, p. 833-855. doi : 10.1007/s10579-016-9364-5 (cf. p. 15).

Adamic, Lada A. et Natalie Glance (2005). « The Political Blogosphere and the 2004 U.S. Election : Divided They Blog ». In : Proceedings of the 3rd International Workshop on Link Discovery. ACM Press, p. 36-43. doi : 10.1145/1134271.1134277 (cf. p. 67).

Airoldi, Edoardo M., David M. Blei, Stephen E. Fienberg et Eric P. Xing (2008). « Mixed Membership Stochastic Blockmodels ». In : J. Mach. Learn. Res. 9, p. 1981-2014. issn : 1532-4435 (cf. p. 40, 122).

Akoglu, Leman (2014). « Quantifying Political Polarity Based on Bipartite Opinion Networks ». In : Proceedings of the 8th International AAAI Conference on Weblogs and Social Media, p. 2-11 (cf. p. 19, 20, 22, 30).

Allport, Floyd H (1937). « Toward a Science of Public Opinion ». In : Public Opinion Quarterly 1.1, p. 7-23. doi : 10.1086/265034 (cf. p. 14, 17).

Artstein, Ron et Massimo Poesio (2008). « Inter-Coder Agreement for Computational Linguistics ». In : Computational Linguistics 34.4, p. 555-596. doi : 10.1162/coli.07-034-R2 (cf. p. 50).

Attal, Jean-Philippe, Maria Malek et Marc Zolghadri (2016). « Overlapping Community Detection Using Core Label Propagation and Belonging Function ». In : Neural Information Processing. Springer, p. 165-174. isbn : 978-3-319-46675-0 (cf. p. 41).

Baccianella, Stefano, Andrea Esuli et Fabrizio Sebastiani (2010). « SentiWordNet 3.0 : An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining ». In : Proceedings of the International Conference on Language Resources and Evaluation (cf. p. 15).

Barberá, Pablo (2015). « Birds of the Same Feather Tweet Together : Bayesian Ideal Point Estimation Using Twitter Data ». In : Political Analysis 23.01, p. 76-91. doi : 10.1093/pan/mpu011 (cf. p. 19, 25, 30).

Barberá, Pablo et Gonzalo Rivero (2015). « Understanding the Political Representativeness of Twitter Users ». In : Social Science Computer Review 33.6, p. 712-729. doi : 10.1177/0894439314558836 (cf. p. 20, 34).

Barberá, Pablo, John T. Jost, Jonathan Nagler, Joshua A. Tucker et Richard Bonneau (2015). « Tweeting From Left to Right : Is Online Political Communication More Than an Echo Chamber ? » In : Psychological Science 26.10, p. 1531-1542. doi : 10.1177/0956797615594620 (cf. p. 24, 67).

Barrat, A., M. Barthelemy, R. Pastor-Satorras et A. Vespignani (2004). « The Architecture of Complex Weighted Networks ». In : Proceedings of the National Academy of Sciences 101.11, p. 3747-3752. doi : 10.1073/pnas.0400087101 (cf. p. 75).

Barrelet, Carl Julien, Sebnem Sahin Kuzulugil et Ayşe Başar Bener (2016). « The Twitter Bullishness Index : A Social Media Analytics Indicator for the Stock Market ». In : Proceedings of the 20th International Database Engineering & Applications Symposium. ACM Press, p. 394-395. doi : 10.1145/2938503.2938508 (cf. p. 1).

Bavelas, Alex (1950). « Communication Patterns in Task-Oriented Groups ». In : The Journal of the Acoustical Society of America 22.6, p. 725-730. doi : 10.1121/1.1906679 (cf. p. 89).

Bedi, Punam et Chhavi Sharma (2016). « Community Detection in Social Networks : Community Detection in Social Networks ». In : Wiley Interdisciplinary Reviews : Data Mining and Knowledge Discovery 6.3, p. 115-135. doi : 10.1002/widm.1178 (cf. p. 41).

Bermingham, Adam et Alan F. Smeaton (2011). « On Using Twitter to Monitor Political Sentiment and Predict Election Results. » In : Sentiment Analysis Where AI Meets Psychology (SAAIP) Workshop at the International Joint Conference for Natural Language Processing (cf. p. 23).

Bhowmick, Plaban Kr., Pabitra Mitra et Anupam Basu (2008). « An Agreement Measure for Determining Inter-Annotator Reliability of Human Judgements on Affective Text ». In : Proceedings of the Workshop on Human Judgements in Computational Linguistics. isbn : 978-1-905593-49-1 (cf. p. 51).

Blaz, Cássio Castaldi Araujo et Karin Becker (2016). « Sentiment Analysis in Tickets for IT Support ». In : Proceedings of the 13th International Conference on Mining Software Repositories. ACM Press, p. 235-246. doi : 10.1145/2901739.2901781 (cf. p. 15).

Blondel, Vincent D, Jean-Loup Guillaume, Renaud Lambiotte et Etienne Lefebvre (2008). « Fast Unfolding of Communities in Large Networks ». In : Journal of Statistical Mechanics : Theory and Experiment 2008.10, P10008. doi : 10.1088/1742-5468/2008/10/P10008 (cf. p. 38).

Boireau, Michaël (2014). « Determining Political Stances from Twitter Timelines : The Belgian Parliament Case ». In : Proceedings of the 2014 Conference on Electronic Governance and Open Society : Challenges in Eurasia. ACM Press, p. 145-151. doi : 10.1145/2729104.2729114 (cf. p. 19-21, 30).

Borra, Erik et Bernhard Rieder (2014). « Programmed Method : Developing a Toolset for Capturing and Analyzing Tweets ». In : Aslib Journal of Infor-

mation Management 66.3, p. 262-278. DOI : 10.1108/AJIM-09-2013-0094 (cf. p. 45).

BOULLIER, Dominique et Audrey LOHARD (2012). Opinion mining et sentiment analysis. OpenEdition Press. ISBN : 978-2-8218-1226-0 978-2-8218-1227-7 978-2-8218-1228-4 (cf. p. 67).

BOURDIEU, Pierre (1973). « L'opinion Publique n'existe Pas ». In : Les temps modernes 318, p. 1292-1309 (cf. p. 4, 14).

BOYADJIAN, Julien, Aurélie OLIVESI et Julien VELCIN (2017). « Le web politique au prisme de la science des données : Des croisements disciplinaires aux renouvellements épistémologiques ». In : Réseaux 204.4, p. 9. DOI : 10.3917/res.204.0009 (cf. p. 126).

BRIGADIR, Igor, Derek GREENE et Pádraig CUNNINGHAM (2015). « Analyzing Discourse Communities with Distributional Semantic Models ». In : Proceedings of the ACM Web Science Conference. ACM Press, p. 1-10. DOI : 10.1145/2786451.2786470 (cf. p. 1, 3, 34, 70).

BURGESS, Jean et Ariadna MATAMOROS-FERNÁNDEZ (2016). « Mapping Sociocultural Controversies across Digital Media Platforms : One Week of #gamergate on Twitter, YouTube, and Tumblr ». In : Communication Research and Practice 2.1, p. 79-96. DOI : 10.1080/22041451.2016.1155338 (cf. p. 2).

CASTRO, Rodrigo, Leonardo KUFFO et Carmen VACA (2017). « Back to #6D : Predicting Venezuelan States Political Election Results through Twitter ». In : 2017 Fourth International Conference on eDemocracy eGovernment. IEEE, p. 148-153. DOI : 10.1109/ICEDEG.2017.7962525 (cf. p. 23).

CASTRO, Rodrigo et Carmen VACA (2017). « National Leaders' Twitter Speech to Infer Political Leaning and Election Results in 2015 Venezuelan Parliamentary Elections ». In : 2017 IEEE International Conference on Data Mining Workshops. IEEE, p. 866-871. DOI : 10.1109/ICDMW.2017.118 (cf. p. 23).

CAVALLARI, Sandro, Vincent W. ZHENG, Hongyun CAI, Kevin Chen-Chuan CHANG et Erik CAMBRIA (2017). « Learning Community Embedding with Community Detection and Node Embedding on Graphs ». In : Proceedings of the 2017 ACM on Conference on Information and Knowledge Management - CIKM '17. ACM Press, p. 377-386. DOI : 10.1145/3132847.3132925 (cf. p. 41).

CERON, Andrea (2017). « Intra-Party Politics in 140 Characters ». In : Party Politics 23.1, p. 7-17. DOI : 10.1177/1354068816654325 (cf. p. 19-21, 30).

CHEREPNALKOSKI, Darko et Igor MOZETIC (2015). « A Retweet Network Analysis of the European Parliament ». In : 2015 11th International Conference on Signal-Image Technology Internet-Based Systems. IEEE, p. 350-357. DOI : 10.1109/SITIS.2015.8 (cf. p. 19, 20, 25).

CHEUNG, Ming, Xiaopeng LI et James SHE (2017). « An Efficient Computation Framework for Connection Discovery Using Shared Images ». In : ACM Transactions on Multimedia Computing, Communications, and Applications 13.4, p. 1-21. DOI : 10.1145/3115951 (cf. p. 1).

Cohen, Jacob (1960). « A Coefficient of Agreement for Nominal Scales ». In : Educational and Psychological Measurement 20.1, p. 37-46. DOI : 10.1177/001316446002000104 (cf. p. 50, 51).

Colleoni, Elanor, Alessandro Rozza et Adam Arvidsson (2014). « Echo Chamber or Public Sphere ? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data : Political Homophily on Twitter ». In : Journal of Communication 64.2, p. 317-332. DOI : 10.1111/jcom.12084 (cf. p. 72, 125).

Conover, M. D., J. Ratkiewicz, M. Francisco, B. Gonçalves, A. Flammini et F. Menczer (2011a). « Political Polarization on Twitter ». In : Proceedings of the 5th International AAAI Conference on Weblogs and Social Media. AAAI Press, p. 89-96 (cf. p. 20, 24, 33, 67, 77).

Conover, Michael D., Bruno Goncalves, Jacob Ratkiewicz, Alessandro Flammini et Filippo Menczer (2011b). « Predicting the Political Alignment of Twitter Users ». In : 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing. IEEE, p. 192-199. DOI : 10.1109/PASSAT/SocialCom.2011.34 (cf. p. 19, 25, 30, 33, 77).

Danon, Leon, Albert Díaz-Guilera, Jordi Duch et Alex Arenas (2005). « Comparing Community Structure Identification ». In : Journal of Statistical Mechanics : Theory and Experiment 2005.09, P09008-P09008. DOI : 10.1088/1742-5468/2005/09/P09008 (cf. p. 80).

David, Esther, Maayan Zhitomirsky-Geffet, Moshe Koppel et Hodaya Uzan (2016). « Utilizing Facebook Pages of the Political Parties to Automatically Predict the Political Orientation of Facebook Users ». In : Online Information Review 40.5, p. 610-623. DOI : 10.1108/OIR-09-2015-0308 (cf. p. 19-22, 30).

Derks, Daantje, Agneta H. Fischer et Arjan E.R. Bos (2008). « The Role of Emotion in Computer-Mediated Communication : A Review ». In : Computers in Human Behavior 24.3, p. 766-785. DOI : 10.1016/j.chb.2007.04.004 (cf. p. 122).

Dong, Rui, Yizhou Sun, Lu Wang, Yupeng Gu et Yuan Zhong (2017). « Weakly-Guided User Stance Prediction via Joint Modeling of Content and Social Interaction ». In : Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. ACM Press, p. 1249-1258. DOI : 10.1145/3132847.3133020 (cf. p. 19, 20, 27, 30).

Dori-Hacohen, Shiri, David Jensen et James Allan (2016). « Controversy Detection in Wikipedia Using Collective Classification ». In : Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM Press, p. 797-800. DOI : 10.1145/2911451.2914745 (cf. p. 124).

Dunn, J. C. (1973). « A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters ». In : Journal of Cybernetics 3.3, p. 32-57. DOI : 10.1080/01969727308546046 (cf. p. 40, 122).

EISENSTEIN, Jacob (2013). « What to Do about Bad Language on the Internet ». In : Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies. ACL, p. 359-369 (cf. p. 18).

EKMAN, Paul, Wallace V. FRIESEN, Maureen O'SULLIVAN, Anthony CHAN, Irene DIACOYANNI-TARLATZIS, Karl HEIDER, Rainer KRAUSE, William Ayhan LECOMPTE, Tom PITCAIRN, Pio E. RICCI-BITTI, Klaus SCHERER, Masatoshi TOMITA et Athanase TZAVARAS (1987). « Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion. » In : Journal of Personality and Social Psychology 53.4, p. 712-717. DOI : 10.1037/0022-3514.53.4.712 (cf. p. 15, 16).

FANG, Anjie, Iadh OUNIS, Philip HABEL, Craig MACDONALD et Nut LIMSOPATHAM (2015). « Topic-Centric Classification of Twitter User's Political Orientation ». In : Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM Press, p. 791-794. DOI : 10.1145/2766462.2767833 (cf. p. 19-21, 30).

FARKAS, Illés, Dániel ÁBEL, Gergely PALLA et Tamás VICSEK (2007). « Weighted Network Modules ». In : New Journal of Physics 9.6, p. 180-180. DOI : 10.1088/1367-2630/9/6/180 (cf. p. 40).

FERRARA, Emilio (2017). « Disinformation and Social Bot Operations in the Run up to the 2017 French Presidential Election ». In : First Monday 22.8. DOI : 10.5210/fm.v22i8.8005 (cf. p. 58).

FIELLER, E. C., H. O. HARTLEY et E. S. PEARSON (1957). « Tests for Rank Correlation Coefficients. I ». In : Biometrika 44.3-4, p. 470-481. DOI : 10.1093/biomet/44.3-4.470 (cf. p. 90).

FLEISS, Joseph L. (1971). « Measuring Nominal Scale Agreement among Many Raters. » In : Psychological Bulletin 76.5, p. 378-382. DOI : 10.1037/h0031619 (cf. p. 51).

FORTUNATO, Santo (2010). « Community Detection in Graphs ». In : Physics Reports 486.3-5, p. 75-174. DOI : 10.1016/j.physrep.2009.11.002 (cf. p. 37).

FORTUNATO, Santo, Marián BOGUÑÁ, Alessandro FLAMMINI et Filippo MENCZER (2008). « Approximating PageRank from In-Degree ». In : Algorithms and Models for the Web-Graph. T. 4936. Springer, p. 59-71. ISBN : 978-3-540-78807-2 978-3-540-78808-9 (cf. p. 90).

FRAISIER, Ophélie, Guillaume CABANAC, Yoann PITARCH, Romaric BESANCON et Mohand BOUGHANEM (2017). « Uncovering Like-minded Political Communities on Twitter ». In : Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval. ACM Press, p. 261-264. DOI : 10.1145/3121050.3121091 (cf. p. 33, 67).

FRAISIER, Ophélie, Guillaume CABANAC, Yoann PITARCH, Romaric BESANCON et Mohand BOUGHANEM (2018). « #Élysée2017fr : The 2017 French Presidential Campaign on Twitter ». In : Proceedings of the 12th International AAAI Conference on Weblogs and Social Media. AAAI Press (cf. p. 43).

Garimella, Kiran, Gianmarco De Francisci Morales, Aristides Gionis et Michael Mathioudakis (2016). « Quantifying Controversy in Social Media ». In : Proceedings of the Ninth ACM International Conference on Web Search and Data Mining. ACM Press, p. 33-42. doi : 10.1145/2835776.2835792 (cf. p. 124).

Garimella, Kiran, Aristides Gionis, Nikos Parotsidis et Nikolaj Tatti (2017). « Balancing Information Exposure in Social Networks ». In : Advances in Neural Information Processing Systems 30. Curran Associates, Inc., p. 4666-4674 (cf. p. 34, 70).

Gaumont, Noé, Maziyar Panahi et David Chavalarias (2017). « Identification et reconfiguration des communautés politiques durant la campagne présidentielle française 2017 : Analyse socio-sémantiques des réseaux de militants politiques sur Twitter ». In : url : https://hal.archives-ouvertes.fr/hal-01575456. Preprint sur HAL (cf. p. 31, 32).

Gayo-Avello, Daniel (2011). « Don't Turn Social Media into Another 'Literary Digest' Poll ». In : Communications of the ACM 54.10, p. 121. doi : 10.1145/2001269.2001297 (cf. p. 17, 23).

Gayo-Avello, Daniel (2012). « No, You Cannot Predict Elections with Twitter ». In : IEEE Internet Computing 16.6, p. 91-94. doi : 10.1109/MIC.2012.137 (cf. p. 1, 33).

Gayo-Avello, Daniel (2013). « A Meta-Analysis of State-of-the-Art Electoral Prediction From Twitter Data ». In : Social Science Computer Review 31.6, p. 649-679. doi : 10.1177/0894439313493979 (cf. p. 23).

Gayo-Avello, Daniel (2015). « Political Opinion ». In : Twitter : A Digital Socioscope. Cambridge University Press, p. 52-74. doi : 10.1017/CBO9781316182635.004 (cf. p. 14).

Giachanou, Anastasia et Fabio Crestani (2016). « Like It or Not : A Survey of Twitter Sentiment Analysis Methods ». In : ACM Computing Surveys 49.2, p. 1-41. doi : 10.1145/2938640 (cf. p. 15).

Gindl, Stefan, Albert Weichselbraun et Arno Scharl (2013). « Rule-Based Opinion Target and Aspect Extraction to Acquire Affective Knowledge ». In : Proceedings of the 22nd International Conference on World Wide Web. ACM Press, p. 557-564. doi : 10.1145/2487788.2487994 (cf. p. 2).

Girvan, M. et M. E. J. Newman (2002). « Community Structure in Social and Biological Networks ». In : Proceedings of the National Academy of Sciences 99.12, p. 7821-7826. doi : 10.1073/pnas.122653799 (cf. p. 77).

Gkontzis, Andreas F., Christoforos V. Karachristos, Chris T. Panagiotakopoulos, Elias C. Stavropoulos et Vassilios S. Verykios (2017). « Sentiment Analysis to Track Emotion and Polarity in Student Fora ». In : Proceedings of the 21st Pan-Hellenic Conference on Informatics. ACM Press, 39 :1-39 :6. doi : 10.1145/3139367.3139389 (cf. p. 15).

Golder, Scott et Michael Macy (2015). « Opportunities and Challenges for Online Social Research ». In : Twitter : A Digital Socioscope. Cambridge University Press, p. 1-20 (cf. p. 1).

Guerrero-Solé, Frederic (2017). « Community Detection in Political Discussions on Twitter : An Application of the Retweet Overlap Network Method to the Catalan Process Toward Independence ». In : Social Science Computer Review 35.2, p. 244-261. doi : 10.1177/0894439315617254 (cf. p. 19, 20, 25).

Gunes, Omer (2016). « Aspect Term and Opinion Target Extraction from Web Product Reviews Using Semi-Markov Conditional Random Fields with Word Embeddings As Features ». In : Proceedings of the 6th International Conference on Web Intelligence, Mining and Semantics. ACM Press, 6 :1-6 :5. doi : 10.1145/2912845.2936809 (cf. p. 15).

Hasan, Kazi Saidul et Vincent Ng (2013). « Stance Classification of Ideological Debates : Data, Models, Features, and Constraints ». In : Proceedings of the Sixth International Joint Conference on Natural Language Processing, p. 1348-1356 (cf. p. 19, 20, 22, 30).

Hasanuzzaman, Mohammed, Gaël Dias et Andy Way (2017). « Demographic Word Embeddings for Racism Detection on Twitter ». In : Proceedings of the Eighth International Joint Conference on Natural Language Processing, p. 926-936 (cf. p. 2).

Hasanuzzaman, Mohammed, Wai Leung Sze, Mahammad Parvez Salim et Gaël Dias (2016a). « Collective Future Orientation and Stock Markets ». In : Frontiers in Artificial Intelligence and Applications, p. 1616-1617. doi : 10.3233/978-1-61499-672-9-1616 (cf. p. 1).

Hasanuzzaman, Mohammed, Gaël Dias, Stéphane Ferrari, Yann Mathet et Andy Way (2016b). « Identifying Temporal Orientation of Word Senses ». In : Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning. ACL, p. 22-30. doi : 10.18653/v1/K16-1003 (cf. p. 123).

INSEE (2010). Des Spécificités Socioprofessionnelles Régionales (cf. p. 125).

INSEE (2017). Population Totale Par Sexe et Âge Au 1er Janvier 2017, France Métropolitaine (cf. p. 125).

IPSOS (2013). Usages et Pratiques de Twitter En France (cf. p. 125).

Iyengar, Shanto et Sean J. Westwood (2015). « Fear and Loathing across Party Lines : New Evidence on Group Polarization ». In : American Journal of Political Science 59.3, p. 690-707. doi : 10.1111/ajps.12152 (cf. p. 4, 67).

Jaccard, Paul (1901). « Étude Comparative de La Distribution Florale Dans Une Portion Des Alpes et Des Jura ». In : Bulletin de la Société Vaudoise des Sciences Naturelles 37, p. 547-579 (cf. p. 72).

Jang, Myungha, Shiri Dori-Hacohen et James Allan (2017). « Modeling Controversy within Populations ». In : Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval. ACM Press, p. 141-149. doi : 10.1145/3121050.3121067 (cf. p. 124).

Jungherr, Andreas (2013). « Tweets and Votes, a Special Relationship : The 2009 Federal Election in Germany ». In : Proceedings of the 2nd Workshop on Politics, Elections and Data. ACM Press, p. 5-14. doi : 10.1145/2508436.2508437 (cf. p. 34).

Junghherr, Andreas, Pascal Jürgens et Harald Schoen (2012). « Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions : A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welpe, I. M. "Predicting Elections With Twitter : What 140 Characters Reveal About Political Sentiment" ». In : Social Science Computer Review 30.2, p. 229-234. doi : 10.1177/0894439311404119 (cf. p. 23, 33, 126).

Kaplan, Andreas M. et Michael Haenlein (2010). « Users of the World, Unite ! The Challenges and Opportunities of Social Media ». In : Business Horizons 53.1, p. 59-68. doi : 10.1016/j.bushor.2009.09.003 (cf. p. 12).

Kivran-Swaine, Funda, Sam Brody, Nicholas Diakopoulos et Mor Naaman (2012). « Of Joy and Gender : Emotional Expression in Online Social Networks ». In : Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion. ACM Press, p. 139. doi : 10.1145/2141512.2141562 (cf. p. 122).

Koumpouri, Athanasia, Iosif Mporas et Vasileios Megalooikonomou (2015). « Feature Selection for Improving Opinion Identification from Web Authors' Posts ». In : Proceedings of the 19th Panhellenic Conference on Informatics. ACM Press, p. 117-122. doi : 10.1145/2801948.2802026 (cf. p. 2).

Kratzke, Nane (2017). « The #BTW17 Twitter Dataset–Recorded Tweets of the Federal Election Campaigns of 2017 for the 19th German Bundestag ». In : Data 2.4, p. 34. doi : 10.3390/data2040034 (cf. p. 3, 34).

Lancichinetti, Andrea, Filippo Radicchi, José J. Ramasco et Santo Fortunato (2011). « Finding Statistically Significant Communities in Networks ». In : PLoS ONE 6.4, e18961. doi : 10.1371/journal.pone.0018961 (cf. p. 40).

Landis, J. Richard et Gary G. Koch (1977). « The Measurement of Observer Agreement for Categorical Data ». In : Biometrics 33.1, p. 159. doi : 10.2307/2529310 (cf. p. 52).

Levorato, Vincent (2010). « Une méthode mixte d'analyse d'un réseau social : classification prétopologique et centralité d'intermédiarité ». In : EGC, A5-77-88 (cf. p. 123).

Lingras, Pawan et Chad West (2004). « Interval Set Clustering of Web Users with Rough K-Means ». In : Journal of Intelligent Information Systems 23.1, p. 5-16. doi : 10.1023/B:JIIS.0000029668.88665.1a (cf. p. 40).

Liu, Bing (2012). « Sentiment Analysis and Opinion Mining ». In : Synthesis Lectures on Human Language Technologies 5.1, p. 1-167. doi : 10.2200/S00416ED1V01Y201204HLT016 (cf. p. 2, 15, 63).

Livne, Avishay, Matthew P. Simmons, Eytan Adar et Lada A. Adamic (2011). « The Party Is Over Here : Structure and Content in the 2010 Election ». In : Proceedings of the International AAAI Conference on Weblogs and Social Media (cf. p. 23).

Lo, James, Sven-Oliver Proksch et Jonathan B. Slapin (2016). « Ideological Clarity in Multiparty Competition : A New Measure and Test Using Election Manifestos ». In : British Journal of Political Science 46.03, p. 591-610. doi : 10.1017/S0007123414000192 (cf. p. 21).

Lu, Haokai, James Caverlee et Wei Niu (2015). « BiasWatch : A Lightweight System for Discovering and Tracking Topic-Sensitive Opinion Bias in Social Media ». In : Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. ACM Press, p. 213-222. doi : 10.1145/2806416.2806573 (cf. p. 34).

MacQueen, J. (1967). « Some Methods for Classification and Analysis of Multivariate Observations ». In : Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1 : Statistics. University of California Press, p. 281-297 (cf. p. 40).

Madlberger, Lisa et Amai Almansour (2014). « Predictions Based on Twitter &#x2014 ; A Critical View on the Research Process ». In : IEEE, p. 1-6. doi : 10.1109/ICODSE.2014.7062667 (cf. p. 17, 125, 126).

Magdy, Walid, Kareem Darwish, Norah Abokhodair, Afshin Rahimi et Timothy Baldwin (2016). « #ISISisNotIslam or #DeportAllMuslims ? : Predicting Unspoken Views ». In : Proceedings of the 8th ACM Conference on Web Science. ACM Press, p. 95-106. doi : 10.1145/2908131.2908150 (cf. p. 19, 20, 27, 30).

Mahendiran, Aravindan, Wei Wang, Jaime Arredondo Sanchez Lira, Bert Huang, Lise Getoor, David Mares et Naren Ramakrishnan (2014). « Discovering Evolving Political Vocabulary in Social Media ». In : 2014 International Conference on Behavioral, Economic, and Socio-Cultural Computing. IEEE, p. 1-7. doi : 10.1109/BESC.2014.7059504 (cf. p. 18).

Makazhanov, Aibek, Davood Rafiei et Muhammad Waqar (2014). « Predicting Political Preference of Twitter Users ». In : Social Network Analysis and Mining 4.1. doi : 10.1007/s13278-014-0193-5 (cf. p. 19, 20, 26, 30, 33).

Martin, Shawn, W. Michael Brown, Richard Klavans et Kevin W. Boyack (2011). « OpenOrd : An Open-Source Toolbox for Large Graph Layout ». In : Proc.SPIE, p. 786806. doi : 10.1117/12.871402 (cf. p. 60).

McLachlan, G. J. et K. E. Basford (1988). Mixture Models. Inference and Applications to Clustering (cf. p. 41).

McPherson, Miller, Lynn Smith-Lovin et James M Cook (2001). « Birds of a Feather : Homophily in Social Networks ». In : Annual Review of Sociology 27.1, p. 415-444. doi : 10.1146/annurev.soc.27.1.415 (cf. p. 4, 24, 67).

Meo, Pasquale De, Katarzyna Musial-Gabrys, Domenico Rosaci, Giuseppe M. L. Sarnè et Lora Aroyo (2017). « Using Centrality Measures to Predict Helpfulness-Based Reputation in Trust Networks ». In : ACM Transactions on Internet Technology 17.1, p. 1-20. doi : 10.1145/2981545 (cf. p. 1).

Merry, Melissa (2016). « Making Friends and Enemies on Social Media : The Case of Gun Policy Organizations ». In : Online Information Review 40.5, p. 624-642. doi : 10.1108/OIR-10-2015-0333 (cf. p. 24, 77).

Metaxas, Panagiotis T., Eni Mustafaraj et Daniel Gayo-Avello (2011). « How (Not) to Predict Elections ». In : 2011 IEEE Third International Conference on Social Computing (cf. p. 23).

Mohammad, Saif M., Parinaz Sobhani et Svetlana Kiritchenko (2017). « Stance and Sentiment in Tweets ». In : ACM Transactions on Internet Technology 17.3, p. 1-23. doi : 10.1145/3003433 (cf. p. 19, 20, 23, 30).

Moreno, J.L. et P-H Maucorps (1955). « Fondements de la sociométrie. Traduit d'après la seconde édition américaine (Who Shall Survive ?) par H. Lesage et P.-H. Maucorps ». In : Revue française de science politique 5.3, p. 641-646. issn : 0035-2950 (cf. p. 11).

Moreno, Jacob Levy (1935). « Who Shall Survive ? A New Approach to the Problem of Human Interrelations ». In : Journal of the American Medical Association 104.12, p. 1033. doi : 10.1001/jama.1935.02760120075039 (cf. p. 11).

Musto, Cataldo, Marco de Gemmis, Giovanni Semeraro et Pasquale Lops (2017). « A Multi-Criteria Recommender System Exploiting Aspect-Based Sentiment Analysis of Users' Reviews ». In : Proceedings of the Eleventh ACM Conference on Recommender Systems. ACM Press, p. 321-325. doi : 10.1145/3109859.3109905 (cf. p. 15).

Newman, M. E. J. (2002). « Assortative Mixing in Networks ». In : Physical Review Letters 89.20. doi : 10.1103/PhysRevLett.89.208701 (cf. p. 60).

Newman, M. E. J. (2003). « Mixing Patterns in Networks ». In : Physical Review E 67.2. doi : 10.1103/PhysRevE.67.026126 (cf. p. 60).

Newman, M. E. J. (2006). « Finding Community Structure in Networks Using the Eigenvectors of Matrices ». In : Physical Review E 74.3. doi : 10.1103/PhysRevE.74.036104 (cf. p. 38).

Newman, Mark (2010). Mathematics of Networks. Oxford University Press. doi : 10.1093/acprof:oso/9780199206650.001.0001 (cf. p. 89).

Nguyen, Thien Hai, Kiyoaki Shirai et Julien Velcin (2015). « Sentiment Analysis on Social Media for Stock Movement Prediction ». In : Expert Systems with Applications 42.24, p. 9603-9611. doi : 10.1016/j.eswa.2015.07.052 (cf. p. 1).

O'Connor, Brendan, Ramnath Balasubramanya, Bryan R. Routledge et Noah A. Smith (2010). « From Tweets to Polls : Linking Text Sentiment to Public Opinion Time Series ». In : Proceedings of the International AAAI Conference on Weblogs and Social Media (cf. p. 23).

Ohbe, Tatsuya, Tadachika Ozono et Toramatsu Shintani (2017). « A Sentiment Polarity Classifier for Regional Event Reputation Analysis ». In : Proceedings of the International Conference on Web Intelligence. ACM Press, p. 1207-1213. doi : 10.1145/3106426.3109416 (cf. p. 15).

Orman, Keziban, Vincent Labatut et Hocine Cherifi (2013). « An Empirical Study of the Relation between Community Structure and Transitivity ». In : Complex Networks. T. 424. Springer, p. 99-110. doi : 10.1007/978-3-642-30287-9_11 (cf. p. 75).

Page, Lawrence, Sergey Brin, Rajeev Motwani et Terry Winograd (1999). The PageRank Citation Ranking : Bringing Order to the Web. Rapp. tech. Stanford InfoLab (cf. p. 89).

PALLA, Gergely, Imre DERÉNYI, Illés FARKAS et Tamás VICSEK (2005). « Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society ». In : Nature 435.7043, p. 814-818. DOI : 10.1038/nature03607 (cf. p. 40).

PENNACCHIOTTI, Marco et Ana-Maria POPESCU (2011). « Democrats, Republicans and Starbucks Afficionados : User Classification in Twitter ». In : Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM Press, p. 430. DOI : 10.1145/2020408.2020477 (cf. p. 19, 20, 28, 30, 33).

PHETHEAN, Christopher, Thanassis TIROPANIS et Lisa HARRIS (2015). « Assessing the Value of Social Media for Organisations : The Case for Charitable Use ». In : Proceedings of the ACM Web Science Conference. ACM Press, 32 :1-32 :9. DOI : 10.1145/2786451.2786457 (cf. p. 1).

PONS, Pascal et Matthieu LATAPY (2005). « Computing Communities in Large Networks Using Random Walks ». In : Computer and Information Sciences - ISCIS 2005. T. 3733. Springer, p. 284-293. DOI : 10.1007/11569596_31 (cf. p. 40).

POUSHTER, Jacob, Rhonda STEWART et Hanyu CHWE (2018). Social Media Use Continues To Rise in Developing Countries, but Plateaus Across Developed Ones. Rapp. tech. Pew Research Center. URL : http://assets.pewresearch.org/wp-content/uploads/sites/2/2018/06/15135408/Pew-Research-Center_Global-Tech-Social-Media-Use_2018.06.19.pdf (cf. p. 124).

PREOŢIUC-PIETRO, Daniel, Ye LIU, Daniel HOPKINS et Lyle UNGAR (2017). « Beyond Binary Labels : Political Ideology Prediction of Twitter Users ». In : Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. ACL, p. 729-740. DOI : 10.18653/v1/P17-1068 (cf. p. 34).

PROIOS, Dimitris, Magdalini EIRINAKI et Iraklis VARLAMIS (2015). « TipMe : Personalized Advertising and Aspect-Based Opinion Mining for Users and Businesses ». In : Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. ACM Press, p. 1489-1494. DOI : 10.1145/2808797.2809324 (cf. p. 2).

RABELO, J., R. B. C. PRUDENCIO et F. BARROS (2012). « Collective Classification for Sentiment Analysis in Social Networks ». In : 2012 IEEE 24th International Conference on Tools with Artificial Intelligence. IEEE, p. 958-963. DOI : 10.1109/ICTAI.2012.135 (cf. p. 19, 20, 28, 30).

RAGHAVAN, Usha Nandini, Réka ALBERT et Soundar KUMARA (2007). « Near Linear Time Algorithm to Detect Community Structures in Large-Scale Networks ». In : Physical Review E 76.3, p. 036106. DOI : 10.1103/PhysRevE.76.036106 (cf. p. 39).

RAJADESINGAN, Ashwin et Huan LIU (2014). « Identifying Users with Opposing Opinions in Twitter Debates ». In : Social Computing, Behavioral-Cultural Modeling and Prediction. T. 8393. Springer, p. 153-160. DOI : 10.1007/978-3-319-05579-4_19 (cf. p. 19, 20, 28, 30).

Ratkiewicz, Jacob, Michael Conover, Mark Meiss, Bruno Gonçalves, Snehal Patil, Alessandro Flammini et Filippo Menczer (2011). « Truthy : Mapping the Spread of Astroturf in Microblog Streams ». In : Proceedings of the 20th International Conference Companion on World Wide Web. ACM Press, p. 249-252. doi : 10.1145/1963192.1963301 (cf. p. 1, 33).

Razzaq, Muhammad Asif, Ali Mustafa Qamar et Hafiz Syed Muhammad Bilal (2014). « Prediction and Analysis of Pakistan Election 2013 Based on Sentiment Analysis ». In : 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. IEEE, p. 700-703. doi : 10.1109/ASONAM.2014.6921662 (cf. p. 23).

Rizos, Georgios, Symeon Papadopoulos et Yiannis Kompatsiaris (2017). « Multilabel User Classification Using the Community Structure of Online Networks ». In : PLOS ONE 12.3, e0173347. doi : 10.1371/journal.pone.0173347 (cf. p. 19, 20, 25, 30).

Rocchio, J. J. (1971). « Relevance Feedback in Information Retrieval ». In : The SMART Retrieval System – Experiments in Automatic Document Processing. Prentice-Hall, Inc., p. 313-323 (cf. p. 102).

Rosvall, M., D. Axelsson et C. T. Bergstrom (2009). « The Map Equation ». In : The European Physical Journal Special Topics 178.1, p. 13-23. doi : 10.1140/epjst/e2010-01179-1 (cf. p. 39).

Saez-Trumper, Diego (2014). « Fake Tweet Buster : A Webtool to Identify Users Promoting Fake News Ontwitter ». In : Proceedings of the 25th ACM Conference on Hypertext and Social Media. ACM Press, p. 316-317. doi : 10.1145/2631775.2631786 (cf. p. 1, 33).

Saha, Sriparna, Sayantan Mitra et Stefan Kramer (2018). « Exploring Multiobjective Optimization for Multiview Clustering ». In : ACM Transactions on Knowledge Discovery from Data 20.2, p. 1-30. doi : 10.1145/3182181 (cf. p. 123).

Sang, Erik Tjong Kim et Johan Bos (2012). « Predicting the 2011 Dutch Senate Election Results with Twitter ». In : Proceedings of the Workshop on Semantic Analysis in Social Media. ACL, p. 53-60 (cf. p. 23, 33, 34).

Saprykin, Dmitry, Galina Kurcheeva et Maxim Bakaev (2016). « Impact of Social Media Promotion in the Information Age ». In : Proceedings of the International Conference on Electronic Governance and Open Society : Challenges in Eurasia. ACM Press, p. 229-236. doi : 10.1145/3014087.3014109 (cf. p. 1).

Shadowen, Ashley Nicole (2017). Ethics and Bias in Machine Learning : A Technical Study of What Makes Us "Good". Rapp. tech. CUNY John Jay College of Criminal Justice, p. 24 (cf. p. 126).

Shi, Jianbo et Jitendra Malik (2000). « Normalized Cuts and Image Segmentation ». In : IEEE Transactions on Pattern Analysis and Machine Intelligence 22.8, p. 18 (cf. p. 41).

Singhal, Amit (2001). « Modern Information Retrieval : A Brief Overview ». In : Bulletin of the IEEE Computer Society Technical Committee on Data Engineering 24.4, p. 35-43 (cf. p. 72).

Skoric, Marko, Nathaniel Poor, Palakorn Achananuparp, Ee-Peng Lim et Jing Jiang (2012). « Tweets and Votes : A Study of the 2011 Singapore General Election ». In : 2012 45th Hawaii International Conference on System Sciences. IEEE, p. 2583-2591. doi : 10.1109/HICSS.2012.607 (cf. p. 23).

Slapin, Jonathan B. et Sven-Oliver Proksch (2008). « A Scaling Model for Estimating Time-Series Party Positions from Texts ». In : American Journal of Political Science 52.3, p. 705-722. doi : 10.1111/j.1540-5907.2008.00338.x (cf. p. 21).

Sluban, Borut, Jasmina Smailovic, Matja Juric, Igor Mozetic et Stefano Battiston (2014). « Community Sentiment on Environmental Topics in Social Networks ». In : 2014 Tenth International Conference on Signal-Image Technology and Internet-Based Systems. IEEE, p. 376-382. doi : 10.1109/SITIS.2014.27 (cf. p. 15).

Smith, Laura M., Linhong Zhu, Kristina Lerman et Zornitsa Kozareva (2013). « The Role of Social Media in the Discussion of Controversial Topics ». In : 2013 International Conference on Social Computing. IEEE, p. 236-243. doi : 10.1109/SocialCom.2013.41 (cf. p. 24, 31, 108).

Stone, Philip J., Robert F. Bales, J. Zvi Namenwirth et Daniel M. Ogilvie (1966). « The General Inquirer : A Computer System for Content Analysis and Retrieval Based on the Sentence as a Unit of Information ». In : Behavioral Science 7.4, p. 484-498. doi : 10.1002/bs.3830070412 (cf. p. 15).

Strapparava, Carlo et Rada Mihalcea (2008). « Learning to Identify Emotions in Text ». In : Proceedings of the 2008 ACM Symposium on Applied Computing. ACM Press, p. 1556-1560. doi : 10.1145/1363686.1364052 (cf. p. 15).

Sublemontier, Jacques-Henri (2012). « Classification non supervisée : de la multiplicité des données à la multiplicité des analyses ». Thèse de doct. Université d'Orléans (cf. p. 123).

Sungsri, Teerapong et Usanad Ua-apisitwong (2017). « The Analysis and Summarizing System of Thai Hotel Reviews Using Opinion Mining Technique ». In : 2017 5th International Conference on Information and Education Technology. ACM Press, p. 167-170. doi : 10.1145/3029387.3029391 (cf. p. 15).

Sunstein, Cass R (2009). Republic. Com 2. 0. Princeton University Press. isbn : 978-0-691-14328-6 978-1-4008-2783-1 978-1-282-75459-1 (cf. p. 4, 24, 67).

Tang, Tiffany Y., Pinata Winoto, Aonan Guan et Guanxing Chen (2018). « "The Foreign Language Effect" and Movie Recommendation : A Comparative Study of Sentiment Analysis of Movie Reviews in Chinese and English ». In : Proceedings of the 2018 10th International Conference on Machine Learning and Computing. ACM Press, p. 79-84. doi : 10.1145/3195106.3195130 (cf. p. 15).

Thelwall, Mike (2009). « Social Network Sites : Users and Uses ». In : Advances in Computers 76, p. 19-73 (cf. p. vi).

Thonet, Thibaut (2017). « Modèles thématiques pour la découverte non supervisée de points de vue sur le Web ». Thèse de doctorat. Université Paul Sabatier (cf. p. 14).

Thonet, Thibaut, Guillaume Cabanac, Mohand Boughanem et Karen Pinel-Sauvagnat (2016). « VODUM : a Topic Model Unifying Viewpoint, Topic and Opinion Discovery ». In : Advances in Information Retrieval. T. 9626. Springer, p. 533-545. doi : 10.1007/978-3-319-30671-1_39 (cf. p. 18).

Thonet, Thibaut, Guillaume Cabanac, Mohand Boughanem et Karen Pinel-Sauvagnat (2017). « Users Are Known by the Company They Keep : Topic Models for Viewpoint Discovery in Social Networks ». In : Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. ACM Press, p. 87-96. doi : 10.1145/3132847.3132897 (cf. p. 19, 27, 30).

Trabelsi, Amine et Osmar Zaiane (2018). « Unsupervised Model for Topic Viewpoint Discovery in Online Debates Leveraging Author Interactions ». In : Proceedings of the 12th International Conference on Weblogs and Social Media (cf. p. 19, 20, 27).

Trice, Michael (2015). « Putting GamerGate in Context : How Group Documentation Informs Social Media Activity ». In : ACM Press, p. 1-5. doi : 10.1145/2775441.2775471 (cf. p. 2).

Tumasjan, Andranik, Timm O. Sprenger, Philipp G. Sandner et Isabell M. Welpe (2010). « Predicting Elections with Twitter : What 140 Characters Reveal about Political Sentiment ». In : Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media (cf. p. 23, 33, 34).

Twitter Marketing FR (2016). Utilisateurs de Twitter En France (cf. p. 125).

Volkova, Svitlana, Yoram Bachrach et Benjamin Van Durme (2016). « Mining User Interests to Predict Perceived Psycho-Demographic Traits on Twitter ». In : 2016 IEEE Second International Conference on Big Data Computing Service and Applications. IEEE, p. 36-43. doi : 10.1109/BigDataService.2016.28 (cf. p. 19, 20, 26, 30).

Vosoughi, Soroush, Mostafa 'Neo' Mohsenvand et Deb Roy (2017). « Rumor Gauge : Predicting the Veracity of Rumors on Twitter ». In : ACM Transactions on Knowledge Discovery from Data 11.4, p. 1-36. doi : 10.1145/3070644 (cf. p. 33).

Vu, Phong Minh, Hung Viet Pham, Tam The Nguyen et Tung Thanh Nguyen (2016). « Phrase-Based Extraction of User Opinions in Mobile App Reviews ». In : Proceedings of the 31st IEEE/ACM International Conference on Automated Software Engineering. ACM Press, p. 726-731. doi : 10.1145/2970276.2970365 (cf. p. 15).

Vuong, Ba-Quy, Ee-Peng Lim, Aixin Sun, Minh-Tam Le et Hady Wirawan Lauw (2008). « On Ranking Controversies in Wikipedia : Models and Evaluation ». In : Proceedings of the International Conference on Web Search and Web Data Mining. ACM Press, p. 171. doi : 10.1145/1341531.1341556 (cf. p. 124).

WAKITA, Ken et Toshiyuki TSURUMI (2007). « Finding Community Structure in Mega-Scale Social Networks ». In : Proceedings of the 16th International Conference on World Wide Web. ACM Press, p. 1275. DOI : 10.1145/1242572.1242805 (cf. p. 38).

WANG, Chun, Shirui PAN, Guodong LONG, Xingquan ZHU et Jing JIANG (2017). « MGAE : Marginalized Graph Autoencoder for Graph Clustering ». In : Proceedings of the 2017 ACM on Conference on Information and Knowledge Management - CIKM '17. ACM Press, p. 889-898. DOI : 10.1145/3132847.3132967 (cf. p. 41).

WEBER, Ingmar, Venkata R. Kiran GARIMELLA et Alaa BATAYNEH (2013). « Secular vs. Islamist Polarization in Egypt on Twitter ». In : Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. ACM Press, p. 290-297. DOI : 10.1145/2492517.2492557 (cf. p. 24, 33, 77).

WEBER, Ingmar, Venkata Rama Kiran GARIMELLA et Asmelash TEKA (2013). « Political Hashtag Trends ». In : Advances in Information Retrieval. T. 7814. Springer, p. 857-860. DOI : 10.1007/978-3-642-36973-5_102 (cf. p. 33).

WONG, Felix Ming Fai, Chee-Wei TAN, Soumya SEN et Mung CHIANG (2013). « Quantifying Political Leaning from Tweets and Retweets ». In : Proceedings of the 7th International Conference on Weblogs and Social Media. AAAI Press, p. 640-649 (cf. p. 19, 20, 29, 30).

WONG, Felix Ming Fai, Chee Wei TAN, Soumya SEN et Mung CHIANG (2016). « Quantifying Political Leaning from Tweets, Retweets, and Retweeters ». In : IEEE Transactions on Knowledge and Data Engineering 28.8, p. 2158-2172. DOI : 10.1109/TKDE.2016.2553667 (cf. p. 20, 29, 30).

YANG, Zaihan, Alexander KOTOV, Aravind MOHAN et Shiyong LU (2015). « Parametric and Non-Parametric User-Aware Sentiment Topic Models ». In : Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM Press, p. 413-422. DOI : 10.1145/2766462.2767758 (cf. p. 2).

ZHANG, Shaodian, Lin QIU, Frank CHEN, Weinan ZHANG, Yong YU et Noémie ELHADAD (2017). « We Make Choices We Think Are Going to Save Us : Debate and Stance Identification for Online Breast Cancer CAM Discussions ». In : Proceedings of the 26th International Conference on World Wide Web Companion. IW3C2, p. 1073-1081. DOI : 10.1145/3041021.3055134 (cf. p. 19, 20, 22, 30).

ZHI, Shuting, Xiaoge LI, Jinan ZHANG, Xian FAN, Liping DU et Zhongyang LI (2017). « Aspects Opinion Mining Based on Word Embedding and Dependency Parsing ». In : Proceedings of the International Conference on Advances in Image Processing. ACM Press, p. 210-215. DOI : 10.1145/3133264.3133305 (cf. p. 15).

ZUBIAGA, Arkaitz, Elena KOCHKINA, Maria LIAKATA, Rob PROCTER, Michal LUKASIK, Kalina BONTCHEVA, Trevor COHN et Isabelle AUGENSTEIN (2018). « Discourse-Aware Rumour Stance Classification in Social Media

Using Sequential Classifiers ». In : Information Processing & Management 54.2, p. 273-290. DOI : 10.1016/j.ipm.2017.11.009 (cf. p. 19, 20, 30, 31).

BOYD, D., S. GOLDER et G. LOTAN (2010). « Tweet, Tweet, Retweet : Conversational Aspects of Retweeting on Twitter ». In : 2010 43rd Hawaii International Conference on System Sciences. IEEE, p. 1-10. DOI : 10.1109/HICSS.2010.412 (cf. p. 24).

BOYD, Danah M. et Nicole B. ELLISON (2007). « Social Network Sites : Definition, History, and Scholarship ». In : Journal of Computer-Mediated Communication 13.1, p. 210-230. DOI : 10.1111/j.1083-6101.2007.00393.x (cf. p. 1, 11).

BOYD, Danah (2010). Big Data : Opportunities for Computational and Social Sciences. URL : http://www.zephoria.org/thoughts/archives/2010/04/17/big-data-opportunities-for-computational-and-social-sciences.html (cf. p. 124).

Ce document a été composé en utilisant la classe `classicthesis` développée par André Miede. Le style a été inspiré par le célèbre livre de Robert Bringhurst sur la typographie « Les éléments du style typographique ». `classicthesis` est disponible pour LaTeX et LyX :

https://bitbucket.org/amiede/classicthesis/

Les utilisateurs et utilisatrices satisfait·e·s de `classicthesis` envoient habituellement une carte postale à l'auteur, une collection de cartes postales reçues est présentée ici :

http://postcards.miede.de/

Bien qu'ayant apporté quelques modifications au style de base, l'auteure de cette thèse est définitivement satisfaite de `classicthesis`, et enverra donc une carte postale avec plaisir.

Version finale : 10 décembre 2018

RÉSUMÉ

De nombreux domaines ont intérêt à étudier les points de vue exprimés en ligne, que ce soit à des fins de marketing, de cybersécurité ou de recherche avec l'essor des humanités numériques. Nous proposons dans ce manuscrit deux contributions au domaine de la fouille de points de vue, axées sur la difficulté à obtenir des données annotées de qualité sur les médias sociaux. Notre première contribution est un jeu de données volumineux et complexe de 22 853 profils Twitter actifs durant la campagne présidentielle française de 2017. C'est l'un des rares jeux de données considérant plus de deux points de vue et, à notre connaissance, le premier avec un grand nombre de profils et le premier proposant des communautés politiques recouvrantes. Ce jeu de données peut être utilisé tel quel pour étudier les mécanismes de campagne sur Twitter ou pour évaluer des modèles de détection de points de vue ou des outils d'analyse de réseaux. Nous proposons ensuite deux modèles génériques semi-supervisés de détection de points de vue, utilisant une poignée de profils-graines, pour lesquels nous connaissons le point de vue, afin de catégoriser le reste des profils en exploitant différentes proximités inter-profils. En effet, les modèles actuels sont généralement fondés sur les spécificités de certaines plateformes sociales, ce qui ne permet pas l'intégration de la multitude de signaux disponibles. En construisant des proximités à partir de différents types d'éléments disponibles sur les médias sociaux, nous pouvons détecter des profils suffisamment proches pour supposer qu'ils partagent une position similaire sur un sujet donné, quelle que soit la plateforme. Notre premier modèle est un modèle ensembliste séquentiel propageant les points de vue grâce à un graphe multicouche représentant les proximités entre les profils. En utilisant des jeux de données provenant de deux plateformes, nous montrons qu'en combinant plusieurs types de proximité, nous pouvons correctement étiqueter 98 % des profils. Notre deuxième modèle nous permet d'observer l'évolution des points de vue des profils pendant un événement, avec seulement un profil-graine par point de vue. Ce modèle confirme qu'une grande majorité de profils ne changent pas de position sur les médias sociaux, ou n'expriment pas leur revirement.

Mots-clés : Fouille de points de vue, médias sociaux, expression politique, humanités numériques

ABSTRACT

Numerous domains have interests in studying the viewpoints expressed online, be it for marketing, cybersecurity, or research purposes with the rise of computational social sciences. We propose in this manuscript two contributions to the field of stance detection, focused around the difficulty of obtaining annotated data of quality on social medias. Our first contribution is a large and complex dataset of 22 853 Twitter profiles active during the French presidential campaign of 2017. This is one of the rare datasets that considers a non-binary stance classification and, to our knowledge, the first one with a large number of profiles, and the first one proposing overlapping political communities. This dataset can be used as-is to study the campaign mechanisms on Twitter, or used to test stance detection models or network analysis tools. We then propose two semi-supervised generic stance detection models using a handful of seed profiles for which we know the stance to classify the rest of the profiles by exploiting various proximities. Indeed, current stance detection models are usually grounded on the specificities of some social platforms, which is unfortunate since it does not allow the integration of the multitude of available signals. By infering proximities from differents types of elements available on social medias, we can detect profiles close enough to assume they share a similar stance on a given subject. Our first model is a sequential ensemble algorithm which propagates stances thanks to a multi-layer graph representing proximities between profiles. Using datasets from two platforms, we show that, by combining several types of proximities, we can achieve excellent results. Our second model allows us to observe the evolution of profiles' stances during an event with as little as one seed profile by stance. This model confirms that a large majority of profiles do not change their stance on social medias, or do not express their change of heart.

Keywords : Stance detection, Social media, Political expression, Computational social science