



HAL
open science

Détection des influenceurs dans des médias sociaux par une approche hybride

Kevin Deturck, Damien Nouvel, Namrata Patel, Frederique Segond

► **To cite this version:**

Kevin Deturck, Damien Nouvel, Namrata Patel, Frederique Segond. Détection des influenceurs dans des médias sociaux par une approche hybride. Traitement Automatique des Langues Naturelles (TALN 2022), Jun 2022, Avignon, France. pp.54-63. hal-03701499

HAL Id: hal-03701499

<https://hal.science/hal-03701499v1>

Submitted on 24 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Détection des influenceurs dans des médias sociaux par une approche hybride

Kévin Deturck¹ Damien Nouvel¹ Namrata Patel² Frédérique Segond^{1,3}

(1) Inalco - Ertim, 2, rue de Lille, 75007 Paris, France

(2) Université Montpellier 3 - MIAp, Rte de Mende, 34090 Montpellier, France

(3) Inria Minatec, 17, avenue des Martyrs, 38000 Grenoble, France

kevin.deturck@inalco.fr, damien.nouvel@inalco.fr,
namrata.patel@univ-montp3.fr, frederique.segond@inalco.fr

RESUME

L'influence sociale est un phénomène important dans divers domaines, tels que l'économie et la politique, qui a gagné en résonance avec la popularité des médias sociaux, notamment les réseaux sociaux et les forums. La majorité des travaux sur ce sujet propose des approches fondées sur des théories en sciences humaines (sociologie, linguistique), et des techniques d'analyse de réseau (mesures de propagation et de centralité) ou de TAL. Dans cet article, nous présentons un modèle d'influence inspiré de travaux en psychologie sociale, sur lequel nous construisons un système combinant un module de TAL pour détecter les messages reflétant les processus d'influence, associé à une analyse par centralité de la transmission de ces messages. Nos expériences sur le forum de débats *Change My View* montrent que l'approche par hybridation, comparée à la centralité seule, aide à mieux détecter les influenceurs.

ABSTRACT

Influencer detection in social media, a hybrid approach

Social influence is an important phenomenon in various fields, such as economics and politics, which has gained resonance with the popularity of social media, especially social networks and forums. Most of the works on this topic propose approaches based on theories in human sciences (sociology, linguistics) and techniques using either network analysis (propagation and centrality measures) or NLP. In this paper, we present a model of influence inspired by works in social psychology, which we implement with an NLP system to detect messages reflecting influence processes, combined with a centrality analysis of the transmission of these messages. Our experiments on the *Change my view* debate forum show that the hybridization approach, compared to the centrality analysis alone, gives significant improvements to detect influencers.

MOTS-CLES : influenceurs, réseaux sociaux, médias sociaux, TAL, graphes, centralité

KEYWORDS: influencers, social networks, social media, NLP, graphs, centrality

1 Introduction

L'*influence* est un phénomène social qui repose en grande partie sur les actions d'individus impactant les opinions, les décisions et les comportements d'autres individus. La notion d'*influenceur*, apparue avec les réseaux sociaux, se rapporte à l'activité d'influence sur Internet, en particulier pour des considérations commerciales ou politiques. Les institutions comme les marques font appel à des influenceurs qui agissent comme des leviers de communication.

La problématique motivant ce travail est de concevoir un système pour détecter automatiquement les influenceurs à partir de données issues de médias sociaux (Deturck, 2021). Nos objectifs sont les suivants : modéliser l'influence à partir de données issues de médias sociaux, identifier des traits linguistiques associés à ce modèle d'influence dans des messages et les détecter automatiquement avec du TAL, enfin, montrer que l'approche classique par centralité peut être combinée à une approche avec du TAL pour mieux détecter les influenceurs.

Nous commençons par un état de l'art en section 2. La section 3 présente la dimension TAL de notre approche, un aspect très peu pris en compte dans la plupart des travaux sur le sujet. Nous présentons l'hybridation de cette approche avec une approche classique par centralité et discutons les résultats en section 4, avant de conclure en section 5.

2 État de l'art

Le psychosociologue Alex Bavelas affirme que la position centrale d'un individu dans les interactions d'un groupe social induit de l'influence sur ce groupe (Bavelas, 1948), ce qu'il vérifiera sur des petits groupes d'individus (Bavelas, 1950). Le sociologue L.C. Freeman met au point des mesures de centralité pour les graphes sociaux, utilisant les notions d'intermédierité (Freeman, 1977) et de proximité (Freeman, 1978). Dans les années 1990, de nouvelles mesures de centralité apparaissent pour mesurer l'importance des pages web, la plus célèbre étant PageRank (Page et al., 1998). Toutes ces mesures sont utilisées directement pour détecter les influenceurs dans les réseaux sociaux en ligne (Wibisono & Ruldeviyani, 2021) ou sont adaptées à ce type d'environnement (Singh, 2022).

D'autres approches s'intéressent au contenu échangé entre les individus, caractérisant les messages des influenceurs afin de les détecter avec du TAL, que ce soit en s'attachant à décrire une stratégie de narration chez les influenceurs (Feng et al., 2021), ou en se basant sur des traits comportementaux des influenceurs, traduits en des traits discursifs (Rosenthal & McKeown, 2017). Des travaux ont intégré dans la mesure de centralité des informations issues de la sémantique des messages, comme la similarité (Katsimpras et al., 2015), l'originalité (Song et al., 2007) ou le niveau d'expertise (Li et al., 2013). Nous nous inspirons de ces travaux en y ajoutant une modélisation unifiée de l'influence qui intègre l'analyse structurelle et une dimension TAL bien plus présente et approfondie par une caractérisation des phénomènes d'influence dans des messages de médias sociaux.

3 L'influence dans les messages

Poursuivant l'objectif d'une modélisation pragmatique de l'influence, nous avons choisi d'en étudier les manifestations langagières, dans des messages de médias sociaux (Nouvel et al., 2019). Afin de concevoir une telle modélisation, notre approche empirique part de discussions incluant des influenceurs pour analyser à la fois leurs messages et ceux des autres participants. Notre objectif est d'identifier des régularités de discours des influenceurs et des influencés.

3.1 Modélisation de l'influence individuelle

Pour modéliser *l'influence individuelle* (à l'échelle d'un individu et non d'un groupe), nous avons analysé manuellement des fils de discussion du forum *Change My View*, en anglais, où le créateur d'un fil de discussion expose son point de vue et son raisonnement sur un sujet de son choix afin que d'autres participants le fassent changer d'avis. Si un participant y parvient, l'auteur initial doit citer le message clé, expliquer son changement d'avis et récompenser l'auteur du message par un symbole « *delta* », validé par la modération du forum si le message explicatif respecte les contraintes du forum.

Nos observations sur *Change My View* nous ont conduit à modéliser l'influence individuelle comme un processus en trois étapes : *stimulus*, *stimulation* et *décision*. Les composantes de *stimulus* et de *stimulation* correspondent à un cadre théorique en psychologie sociale, qui conçoit l'environnement social d'un individu comme porteur de *stimuli* pouvant induire la *stimulation* de son état psychologique (Turner & Oakes, 1986) et ainsi le conduire à une prise de *décision* (Ajzen, 1996).

En analysant manuellement des fils de discussion du forum « *Change my view* » contenant des changements d'avis et donc des actions d'influence individuelle, nous avons identifié des traits énonciatifs réguliers correspondant à chacune des trois composantes du modèle. Ne sont pas détaillés ici les traits énonciatifs de *stimulation* car ils ne sont pas l'objet d'expériences dans le présent article. Il y a trois types de *stimuli* : le *claim*, désignant un énoncé présenté comme factuel, la *pédagogie*, un type d'énoncé par lequel un énonciateur fait la leçon à son audience ou la conseille, et l'*argumentation*, un énoncé utilisé par un énonciateur pour rallier une audience à son opinion en la motivant par des arguments.

- Exemple de *Claim* : « *Salman hold talks with Putin* » ; il s'agit de la description d'un prétendu événement, vrai ou faux, il s'agit bien d'un énoncé donné comme factuel
- Exemple de *Pédagogie* : « *Turn it off so they can stay in the darkness of their misguidance* » ; il s'agit de l'expression d'un conseil directement adressé à son destinataire, avec une explication, pour le guider dans son comportement et sa pensée
- Exemple d'*argumentation* : « *Since technology is always consolidating, it's only logical that it'll continue to do so.* » ; l'auteur énonce ici son idée sur un sujet (« *it'll continue to do so* ») qu'il justifie par un argument (« *Since technology is always consolidating* »), en insistant sur le caractère logique de son argumentation (« *it's only logical* »)

Le trait énonciatif de *décision*, dans *Change My View*, est l'expression d'un changement d'avis, la finalité d'une action d'influence dans ce forum. Cela correspond à toute modification, même partielle, du positionnement intellectuel d'un individu sur un sujet.

- Exemple de *Changement d'avis* : « This definitely makes me rethink my point on this » ; l'auteur dit penser autrement suite à la lecture d'un message, ce qui est indiciateur d'un impact sur son positionnement intellectuel

Un système de détection des influenceurs, fondé sur notre modèle de l'influence individuelle doit être capable de repérer automatiquement les traits énonciatifs que nous venons de présenter. Dans la section suivante, nous présentons l'implémentation et l'évaluation des modules destinés à la détection automatique des traits *stimuli* et du trait de *décision*.

3.2 Détection automatique de l'influence individuelle

Pour le développement et l'évaluation d'un système de TAL qui détecte les messages contenant nos traits énonciatifs d'influence, nous avons constitué un corpus de référence avec une campagne d'annotation qui consistait à repérer dans des messages les segments de texte correspondant à chacun des traits. Bien que la campagne d'annotation ait porté sur l'ensemble des traits, nous n'avons pas eu le temps d'expérimenter sur la *stimulation*. Aussi, les expériences pour la détection du changement d'avis ont été décrites dans une publication antérieure (Deturck, 2018) et nous n'avons pas utilisé les données de la campagne d'annotation pour ces dernières mais le système « *dela* » du forum. Ainsi, nous focalisons ici la présentation sur la détection des traits *stimuli*.

Des deux corpus utilisés pour cette campagne, l'un est extrait du forum *Change My View*, avec un filtrage des discussions assurant un minimum d'activité de l'auteur initial et des autres participants (Tan et al. 2016). De ce corpus pré-existant, nous n'avons conservé que les discussions contenant au minimum un changement d'avis du participant initial (celui qui expose son point de vue). L'autre corpus est constitué de tweets en anglais, d'individus désignés comme étant *pro-État islamique*¹ : partant de l'hypothèse que les individus *pro-État islamique* agissent en influenceurs, l'objectif était de voir si nous retrouvions nos traits *stimuli* dans leurs tweets. Ce corpus ne présentant pas les réactions aux tweets, nous ne l'avons pas utilisé pour l'annotation des traits de *stimulation* et de *décision* puisque ces derniers concernent les réactions aux *stimuli*.

Les annotateurs étaient des étudiants en TAL dont l'anglais n'était pas la langue maternelle. Pour cadrer le travail d'annotation, nous avons rédigé un guide (Deturck, 2021) qui a été amélioré itérativement, en collaboration avec les annotateurs, lors de discussions et de réconciliations (unification des désaccords entre les annotateurs) après chaque session d'annotation. Nous avons organisé quatre sessions concernant les traits *stimuli*, une session n'a concerné que le trait *Claim* car

¹ <https://www.kaggle.com/fifthtribe/how-isis-uses-twitter>

nous n'avions alors rédigé le guide que pour ce trait-ci. Il y a eu un total de 27 annotateurs, ceux des deux premières sessions, sept duos puis cinq duos et 1 trio, étaient complètement différents, tandis que ceux des deux dernières sessions, deux duos, étaient quatre annotateurs ayant déjà travaillé lors de la deuxième session.

À chaque session, les groupes ont reçu pour l'annotation des corpus de volumes identiques et d'un seul genre textuel (tweet ou forum) pour simplifier l'annotation. Pour le trait *claim*, 1126 messages ont été annotés avec 45% de messages contenant une instance de *claim*, pour les traits *pédagogie* et *argumentation*, 716 messages ont été annotés avec 14% de messages contenant une instance de *pédagogie* et 7% une instance d'*argumentation*. Sur les sessions 2 à 4, qui portent sur tous les traits *stimuli*, les accords inter-annotateurs mesurés avec la métrique *GammaCat* (Mathet, 2017), qui compare la catégorisation des segments jugés similaires par leur position dans les textes selon un algorithme associé, a une valeur moyenne de 0,70 et affichant une progression allant de 0,53 sur la session 2 à 0,88 sur la session 4.

Pour la détection automatique des *stimuli*, conçue comme un problème de classification binaire des messages par trait selon que le trait est présent ou pas, nous avons utilisé l'ensemble des données issues de la campagne d'annotation avec un partitionnement de type 70%-30% respectivement pour le développement et pour l'évaluation.

Avec la librairie Scikit-learn², nous avons expérimenté quatre algorithmes classiques : un algorithme SVM (l'implémentation *SVC*), l'algorithme des forêts aléatoires, un autre fondé sur la régression logistique et un perceptron multicouche. Nous avons associé à ces algorithmes différents descripteurs des messages basés d'une part sur une représentation TF-IDF des messages, visant la spécificité des énoncés *stimuli* à travers les caractères, tokens, catégories morphosyntaxiques et dépendances syntaxiques en présence, et nous avons utilisé d'autre part des plongements lexicaux Bert³ et Glove⁴, en utilisant respectivement les librairies PyTorch⁵ et Transformers (Wolf et al., 2020) avec le modèle Bert d'une part et la librairie Spacy⁶ d'autre part pour passer d'une représentation lexicale à une représentation des messages. L'ensemble de ces descripteurs sont agrégés.

Nous avons optimisé les données produites par les descripteurs des messages en utilisant les algorithmes *StandardScaler* et *SelectKBest*. Nous avons aussi optimisé la configuration des algorithmes en utilisant l'algorithme *GridSearchCV* avec une validation croisée sur les données d'entraînement, sur quelques paramètres génériques, les autres étant laissés à leur valeur par défaut : *k* parmi {5, 10, 20, 30} pour *SelectKBest*, *C* parmi {0,00001, 1, 1 000} pour SVM et la régression logistique, *n_estimators* parmi {10, 100, 500} pour les forêts aléatoires, *penalty* parmi {l1, l2} et *solver* parmi {liblinear, saga} pour la régression logistique, *max_iter* parmi {1 000, 5 000, 10 000}

² <https://scikit-learn.org/stable/>

³ <https://huggingface.co/bert-base-uncased>

⁴ [spacy en vectors web lg-2.3.0](https://spacy.io/vectors)

⁵ <https://pypi.org/project/torch/>

⁶ <https://spacy.io>

pour le perceptron multicouche, ainsi que les paramètres *kernel* et *gamma* pour SVM, *criterion* et *max_features* pour les forêts aléatoires, *activation* et *solver* pour le perceptron multicouche, parmi toutes les valeurs discrètes proposées par la librairie.

Nous présentons ci-dessous les configurations ayant produit les meilleurs résultats pour la détection de chaque trait *stimulus*.

- Claim : Forêts aléatoires + Glove ; précision à 0,66, rappel à 0,69, f-mesure à 0,67
- Pédagogie : Régression logistique + Bert + Glove + TF-IDF avec morpho-syntaxe, et dépendances syntaxiques ; précision à 0,29, rappel à 0,76, f-mesure à 0,42
- Argumentation : Perceptron multicouche + Bert + TF-IDF 2-3 caractères, 1-2 tokens ; précision à 0,52, rappel à 0,62, f-mesure à 0,57

Nous évinçons la pédagogie pour l'hybridation à cause de la faible précision. Chacune des configurations sera utilisée indépendamment pour la détection de chaque trait.

4 Approche hybride à la détection des influenceurs

Nous allons voir comment nous avons combiné l'approche classique d'analyse de la structure des échanges par centralité à l'analyse des messages pour détecter l'influence et les influenceurs. Notre hypothèse est que l'hybridation de ces deux types d'information permettra une meilleure détection.

4.1 Modélisation de l'influence par hybridation

Le principe de l'hybridation est de créer des graphes qui intègrent les informations sémantiques issues des messages correspondant à nos trois composantes d'influence individuelle : le *stimulus*, la *stimulation* et la *décision*. La construction du graphe utilisé dans l'hybridation est faite par l'ajout successif de trois composantes que nous présentons ordonnées ci-dessous.

1. **Le graphe de *stimuli*** : les nœuds du graphe représentent les individus tandis que chaque arc du graphe orienté représente la circulation d'au minimum un message contenant un *stimulus* d'un participant vers un autre. Nous considérons qu'un message *stimulus* a transité d'un individu *Orig* vers un individu *Dest* dans les trois cas suivants : l'individu *Orig* a envoyé un message contenant un *stimulus* à l'individu *Dest*, l'individu *Dest* a répondu à un message de l'individu *Orig* contenant un *stimulus* (même s'il ne lui était pas explicitement destiné), l'individu *Dest* a répondu à l'individu *Orig* par un message contenant un changement d'avis.
2. **La pondération de l'influence individuelle dans le graphe de *stimuli*** : pour quantifier dans le graphe de *stimuli* l'influence individuelle entre les individus représentés, nous attribuons à chaque arc du graphe le poids moyen de tous les messages *stimuli* ayant transité sur cet arc. Le poids d'un message *stimulus* est conçu pour refléter la probabilité que ce

message influence un individu cible, ce que nous associons à une caractérisation de la réaction éventuelle que ce message a engendrée chez l'individu cible, contenue dans tout message en réponse à un message *stimulus*. Le poids d'un message *stimulus* peut prendre trois valeurs : 0,25 en l'absence de réaction, 1 (la valeur maximale) si le message en réaction inclut l'expression d'un changement d'avis, et 0,5 pour toute autre réaction

- 3. La centralité des individus représentés dans le graphe d'influence individuelle :** la centralité doit indiquer dans quelle mesure chaque individu propage de l'influence, la mesure de centralité doit donc se focaliser sur les liens sortants du nœud analysé

4.2 Cadre expérimental

Pour le développement et l'évaluation de notre système hybride, nous avons utilisé 50 fils de discussion qui proviennent du corpus *Change My View* en partie utilisé pour la campagne d'annotation, avec les critères de filtrage sur le minimum de participation et de changement d'avis. Ces 50 fils de discussion contiennent 4 664 messages, 1 435 participants dont 5% d'influenceurs. Les données utilisées pour l'apprentissage des classificateurs sont distinctes de celles utilisées pour l'hybridation : elles correspondent à l'ensemble des messages annotés durant la campagne.

Pour une évaluation comparative, en plus de la construction du graphe correspondant à notre modèle hybride, nous avons aussi construit un graphe *baseline* qui représente la transmission des messages, quels qu'ils soient, entre les individus : il y a un arc orienté depuis un nœud représentant un individu *Orig* vers un nœud représentant un individu *Dest* si *Orig* a envoyé un message à *Dest*, ou si *Dest* a répondu à un message de *Orig*. Nous avons appliqué deux mesures de centralité *sortante* sur ces graphes : le *degré sortant* (Shaw, 1954), basé sur le nombre d'arcs sortants, et la mesure *hits relais* (Kleinberg, 1999), qui ajoute une pondération des arcs sortants d'autant plus grande que leurs nœuds cibles ont des liens entrants, constituant ainsi des pôles dans le graphe.

Nous avons expérimenté trois configurations du graphe de *stimuli* et du graphe *baseline* : (1) aucune pondération dans les deux graphes, (2) une pondération plus importante des arcs lorsque les messages correspondants ont engendré des messages en réponse (0,5 contre 0,25), et (3) ajoute à (2) une pondération maximale (un poids de 1) des arcs *stimuli* si les messages correspondants ont engendré un changement d'avis. Nous avons aussi comparé l'utilisation de la centralité des participants mesurée par discussion, sur des *graphes locaux*, chacun étant construit sur une seule discussion du jeu de données, avec la mesure de centralité des participants mesurée sur un *graphe global*, construit à partir de l'ensemble des discussions du jeu de données.

Pour l'évaluation, nous avons établi les classements des participants par discussion selon leurs scores de centralité dans les différentes configurations présentées auparavant, puis appliqué la mesure *Mean Average Precision* (MAP) pour comparer ces classements à une catégorisation binaire des participants, considérés comme des influenceurs s'ils ont reçu une récompense « *delta* » validée par

la modération du forum *Change My View*. Le score, de 0 à 1, est d'autant plus grand que les influenceurs sont situés en haut du classement évalué.

4.3 Résultats

Évaluation	MAP avec l'hybridation				MAP avec la <i>baseline</i>			
	Graphe global		Graphes locaux		Graphe global		Graphes locaux	
Mesure de centralité	Degré	Hits	Degré	Hits	Degré	Hits	Degré	Hits
Sans pondération	0,26	0,24	0,32	0,33	0,205	0,204	0,22	0,25
Réactions diverses	0,27	0,25	0,349	0,348	0,21	0,20	0,30	0,22
Ch. d'avis distingué	0,31	0,28	0,48	0,55	Seulement par hybridation			

TABLEAU 1 : Résultats des expériences sur l'hybridation

L'approche par hybridation supplante globalement la *baseline* (jusqu'à 10 points en plus). La mesure Hits ne supplante notablement la mesure par degré sortant, pourtant plus simple, que sur les graphes locaux, avec la distinction du changement d'avis (7 points en plus), ce qui semble indiquer l'importance de ce trait détecter les influenceurs.

5 Conclusion et perspectives

Nous avons proposé un modèle d'influence à partir de messages d'un forum d'argumentation, réalisé une caractérisation linguistique de ce modèle par des traits énonciatifs originaux et mené une campagne d'annotation de ces traits, qui nous a permis de développer et d'évaluer des systèmes à base de TAL, qui ont montré la possibilité de les détecter automatiquement. Nos expériences par hybridation ont prouvé qu'une approche avec du TAL peut apporter un gain substantiel pour détecter les influenceurs par rapport à la seule approche par centralité, plus classique.

Il serait intéressant d'utiliser d'autres sources de données pour la robustesse de notre système, d'ajouter la détection automatique des traits de *stimulation* et d'améliorer la détection des traits déjà intégrés. Nous pourrions aussi affiner la sortie du système avec, par exemple, un classement des influenceurs et des *stimuli* d'après l'étendue de leur impact, ainsi qu'une analyse des relations particulières qui existent peut-être entre des *stimuli* et des *stimulations*.

Références

- AJZEN, I. (1996). The social psychology of decision making. *Social Psychology: Handbook of Basic Principles*, p. 297-325.
- BAVELAS A. (1948). A mathematical model for group structures. *Human organization*, p. 16–30. DOI : [10.1121/1.1906679](https://doi.org/10.1121/1.1906679).
- BAVELAS A. (1950). Communication patterns in task-oriented groups. *The journal of the acoustical society of America*, p. 725–730. DOI : [10.1121/1.1906679](https://doi.org/10.1121/1.1906679).
- DETURCK K. (2021). *Détection des influenceurs dans des médias sociaux*, Doctoral dissertation, Institut National des Langues et Civilisations Orientales-INALCO PARIS-LANGUES O'.
- DETURCK, K. (2018). Détection d'influenceurs dans des médias sociaux. In *Actes de la conférence TALN, Volume 2-Démonstrations, articles des Rencontres Jeunes Chercheurs, ateliers DeFT*, Rennes, France, p. 117-130 : ATALA.
- FENG Y., CHEN, H., & KONG, Q. (2021). An expert with whom I can identify: The role of narratives in influencer marketing. *International Journal of Advertising*, 40(7), p. 972-993. DOI : [10.1080/02650487.2020.1824751](https://doi.org/10.1080/02650487.2020.1824751).
- FREEMAN, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, p. 35-41.
- FREEMAN, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, 1(3), 215-239.
- KATSIMPRAS, G., VOGIATZIS, D., & PALIOURAS, G. (2015). Determining influential users with supervised random walks.
- KLEINBERG, J. M. (1999). Authoritative sources in a hyperlinked environment. *The Structure and Dynamics of Networks*, 9781400841(5), p. 514–542. DOI : [10.1145/324133.324140](https://doi.org/10.1145/324133.324140).
- LI, Y., MA, S., ZHANG, Y., HUANG, R., & KINSHUK. (2013). An improved mix framework for opinion leader identification in online learning communities. *Knowledge-Based Systems*, 43, p. 43–51. DOI : [10.1016/j.knosys.2013.01.005](https://doi.org/10.1016/j.knosys.2013.01.005).
- MATHET, Y. (2017). The Agreement Measure γ cat a Complement to γ Focused on Categorization of a Continuum. *Computational Linguistics*, 43(3), p. 661-681. DOI : [10.1162/COLI_a_00296](https://doi.org/10.1162/COLI_a_00296).
- PAGE, L., BRIN, S., MOTWANI, R., & WINOGRAD, T. (1998). The PageRank Citation Ranking: Bringing Order to the Web. *World Wide Web Internet And Web Information Systems*. DOI : [10.1.1.31.1768](https://doi.org/10.1.1.31.1768).
- ROSA, H., CARVALHO, J. P., ASTUDILLO, R., & BATISTA, F. (2018). Page rank versus katz: Is the centrality algorithm choice relevant to measure user influence in twitter? *Studies in Computational Intelligence*, 1-9 : Springer. DOI : [10.1007/978-3-319-74681-4_1](https://doi.org/10.1007/978-3-319-74681-4_1).
- ROSENTHAL, S., & MCKEOWN, K. (2017). Detecting influencers in multiple online genres. *ACM Transactions on Internet Technology*, 17(2), 1–22. DOI : [10.1145/3014164](https://doi.org/10.1145/3014164).
- SHAW, M. E. (1954). Group Structure and the Behavior of Individuals in Small Groups. *Journal of Psychology: Interdisciplinary and Applied*, 139-149. DOI : [10.1080/00223980.1954.9712925](https://doi.org/10.1080/00223980.1954.9712925).
- SINGH R. R. (2022). Centrality measures: a tool to identify key actors in social networks. *Principles of Social Networking*, p. 1-27. Springer, Singapore. DOI : [10.1007/978-981-16-3398-0_1](https://doi.org/10.1007/978-981-16-3398-0_1).

SONG, X., CHI, Y., HINO, K., & TSENG, B. L. (2007). Identifying opinion leaders in the blogosphere. *International Conference on Information and Knowledge Management, Proceedings*, 07(January 2007), p. 971–974. DOI: [10.1145/1321440.1321588](https://doi.org/10.1145/1321440.1321588).

TAN, C., NICULAE, V., DANESCU-NICULESCU-MIZIL, C., & LEE, L. (2016). Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions. *25th International World Wide Web Conference, WWW 2016*, p. 613–624. DOI: [10.1145/2872427.2883081](https://doi.org/10.1145/2872427.2883081).

TURNER, J. C., & OAKES, P. J. (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology*, 25(3), 237–252. DOI: [10.1111/j.2044-8309.1986.tb00732.x](https://doi.org/10.1111/j.2044-8309.1986.tb00732.x).

WIBISONO, A. I., & RULDEVIYANI, Y. (2021). Detecting Social Media Influencers of Airline Services through Social Network Analysis on Twitter: A Case Study of the Indonesian Airline Industry. *3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)*, p. 314-319. IEEE. DOI: [10.1109/EIConCIT50028.2021.9431876](https://doi.org/10.1109/EIConCIT50028.2021.9431876).