Manipulation and fake news detection on social media: a two domain survey, combining social network analysis and knowledge bases exploitation

G. Gadek[1], V. Justine[1], J. Everwyn[1,2]

[1] Airbus Defence and Space
[2] Normandie Université, Université de Caen Normandie, GREYC UMR 6072, Caen, France

Abstract. Social media have to be seen as an adversarial space because of the presence of adversaries, manipulation and disinformation. On classic sources of information, these challenges are usually handled by qualifying content (truth likelihood), and emitters (actor credibility). To adapt this approach to social media, we use influence models, behaviour analysis and community detection for emitters characterisation. This can be combined with the exploitation of knowledge bases for automatic fact checking. This paper proposes a review of this multi-domain challenge.

Keywords:

social media, reasoning, knowledge base entities, graph analytics, credibility, likelihood

1 Introduction

The prevalence of online social networks in the commercial discussion and in the public political debate is not to be proved anymore. Platforms such as Facebook, Weibo or VK function as catalysts to aggregate the opinions of citizens, as well as the feedback of consumers that policy makers and companies take into account.

On these media, adversaries are eager to attack: there is always some uncertainty about the legitimacy of the opinions, as they may constitute a destabilisation intent during a manipulation campaign, launched either by malicious groups, activists, dishonest companies or state-sponsored groups. The countermeasures are monitoring tools, which enable social network analysts to identify the opinions, investigate user importance and relevance, and detect the social patterns of activity that may hide an organised manipulation campaign.

However, these tasks cannot be totally automated due to the creativity of the malicious emitters: there are patterns, yet they do evolve over time. Various

approaches intent to identify manipulations, either through content or structure analysis (i.e., *clickbait* like), while others focus on the social topology of opinion propagation. All these techniques enable the analysts to qualify the emitters and propagators of information and their patterns, which eventually results in a score of **credibility**.

A second approach used to trigger the "manipulation" bell relies in checking the **likelihood** of every piece of information. While fake news look real at first sight, they often do not resist against a well-informed mind.

The automation of this approach can be summed up as an automatic extraction of the informative elements from every publication, and their comparison against a knowledge base. Let us picture it through an example: an entity (say, a boat) is said to have been bombed at some location, while previous information (e.g., maximal boat speed and previous location) physically contradicts this possibility. The likelihood of such message is then automatically degraded.

In this paper, we present the opportunities and highlight the challenges that still lay between today's capabilities, and tomorrow's.

2 Manipulations on social media

The malicious presence on social media takes many forms and is also in itself a source of fake news (e.g., confusion between content moderation, and censorship). This section provides concrete elements to document the current situation of this informational space, beginning with a group of manipulation operations attributed to Russia.

2.1 The social media manipulations attributed to Russia

Very heavy suspicions weigh on the involvement of Russian services in the conduct of the 2016 US presidential campaign, which resulted in the election of Donald Trump.

At least two specific operations have been spotted: the database hack of the mail server of the Democratic Party, disclosing the internal mode of operation of the party in the middle of the campaign, and the massive use of false accounts on the social networks Facebook, Twitter and Reddit. These false accounts propagated content issued by doubtful news outlets: even if some of these sites were motivated solely by money, such as these Macedonian entrepreneurs designing clickbait sites,[1] others seem more closely connected to the Kremlin, which has a long history of cyberspace strategy.[2]

[1] `https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo`

[2] `https://www.researchgate.net/publication/276570725_La_Russie_dans_le_cyberespace_representations_et_enjeux`

The US parliamentary inquiry, which was initially unclear, first highlighted the role played by Russian state media, such as *RussiaToday* and *SputnikNews*, in propagating pro-Trump and anti-Clinton information, rumours and arguments. The Kremlin has denied, saying that these media act similarly to *VoiceOfAmerica*, or *France24*.

Another type of suspicious action goes through the distribution of advertisements on social media, and especially on Facebook via *memes*. A short list of such ads is made public by the Democratic Party,[3] and suggests that the ads were paid for by "nearby companies of the Kremlin". In Figure 1, an example of such an announcement is given: it insists in representing the federation as a way to impose Islam, terror and LGBT while opposing a strong Texan identity. Of course, a real secession is not likely; the goal is not even to get Donald Trump elected, but rather to sow division and discord among the electorate, polarising the public debate.

Fig. 1. A discord-setting ad on Facebook, during the elections meddling

It is only in a second time that less conventional traces appeared, such as this list of Twitter accounts, banned but stamped as fake accounts handled from Russia.[4] This attribution remains denied by the Kremlin, and globally unverifiable. The very existence of a Russian disinformation agency seems to be part of the power strategy: it is about claiming a capacity of confusion, which they do not necessarily have.

[3] `https://democrats-intelligence.house.gov/hpsci-11-1/`
[4] `https://democrats-intelligence.house.gov/uploadedfiles/exhibit_b.pdf`

However, the impact of these "Russian" accounts is real, as they have consistently been able to appear in the mainstream media:[5] often on questionable news sites (Telegraph, Buzzfeed), but also on references (BBC, The Guardian). It is necessary to clarify the argument: until now, no "real" proof of the Russian state involvement could be established.[6] On the other hand, it is now clear that a large number of sources of unreliable content has conquered the media space. Vaguely humorous messages are consumed as information, sometimes pushing to the extreme without any informational basis: the term "fake news" is now rooted in our everyday lives.

2.2 A few astro-turfing examples

Russia is not the only source of social media manipulation: the company *Cambridge Analytica* claimed (without proofs) to have tipped the vote to "Leave" in the referendum for Brexit in June 2016. Among their tools appear the massive distribution of content on social networks, to contain the adversary narrative opposing impose theirs on audiences.

Other countries use such techniques, sometimes continuously. The British daily *The Guardian* exposed in November 2016 that some thirty states around the world relied on "opinions shapers", online opinion formers, to occupy the media space and legitimise the speeches and actions undertaken by their governments [7]. Opinions shapers can be deployed with a low cost and complexity, through crowdsourcing platforms like Amazon Mechanical Turk while being paid only a few cents per post. There may even be no need to create new accounts or to insert them into the global social network as AMT contractors may use and disguise their own accounts. Some countries militarise these operations with an "Internet Water Army" model.[8] Allied services, such as the UK's GCHQ, also have an interesting arsenal,[9] including identity theft as well as psychology based persuasion techniques.

These operations may require sockpuppets, or false identities (sometimes called *sybils*), which reach very unequal levels of sophistication. They simply are fake accounts, manipulated under a false identity [1]. This process enable malicious actors to post attitudes, sometimes extreme, without assuming them directly. They are sometimes used in the field of recommendation systems [13]. Some accounts are quite hollow, created in a hurry without any profile picture. Others seem quite realistic, have credible social links as well as a careful geographical coherence. In 2012, Facebook estimated that there were at least 83

[5] https://www.theguardian.com/media/2017/nov/20/russian-troll-army-tweets-cited-more-than-80-times-in-uk-media

[6] https://www.monde-diplomatique.fr/2017/12/MATE/58207

[7] https://www.theguardian.com/technology/2017/nov/14/social-media-influence-election-countries-armies-of-opinion-shapers-manipulate-democracy-fake-news

[8] https://en.wikipedia.org/wiki/Internet_Water_Army

[9] https://theintercept.com/2014/02/24/jtrig-manipulation/

million *sybils* accounts on their service.[10] These accounts have a value on the black market [32], both for the artificial increase in the number of friends and for the publication of manufactured notices.

2.3 Back to fake news

Various definitions and categorisations have been proposed in the domain of fake news analysis; basically, a distinction is done between "serious fabrications", when news article are forged, mentioning events that never happened, "hoaxes" or rumours that only aim to be spread (and are sometimes referred to as *bullshit*), and "satire": fabrications with an obvious humoristic goal (e.g., *The Onion*[11]) [26]. Overall, the term "fake news" refers both to the globally speaking post-truth informational space, and to pieces of information that are intentionally diffused, while knowing they are false.

Beyond this term, the real problem deals with information manipulation, and the many ways to mix intent, information (e.g. messages) and knowledge (e.g. facts). An exhaustive formalisation of this problem has been recently proposed, with a special focus on the act of *lying* [11]: everything is not only based on content falsehood, but also on content perception based on its source, and its propagation.

Social media are a primary environment for fake news propagation because of their very nature [27]. Social psychology already offers the ground for opinion acceptance by social contact, as in the classic models of opinion propagation [10]; this human behaviour is increased by the recommendation algorithms, creating at the same time echo chambers and viral diffusion. Nowadays, this space seem structured by clusters of like-minded people, exchanges of emotion-loaded content and a clear polarisation of opinions.

3 Holistic social media analysis: towards credibility assessment

We propose to split the huge task of social media analysis into three separate challenges, as illustrated in Figure 2. For various social network platforms, content are collected as *messages*; text mining algorithms computes sentiment and topics. Relations are extracted from the exchanged messages, enabling the construction of a social graph and thus, community detection.

This high level information provides a basis for credibility assessment, through behaviour analysis, at user level, and through a social analysis, resulting in a visualisation of the network through its active communities.

[10] `http://www.bbc.co.uk/news/technology-18813237`
[11] `https://www.theonion.com/`

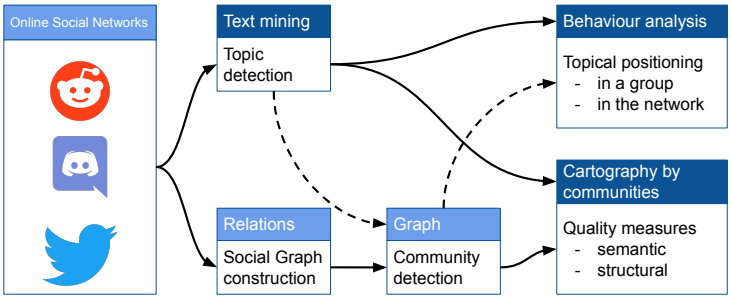

Fig. 2. Social media mining as three separate technological challenges

3.1 Text mining

The first level of online social network analysis is based on its main resource, the textual contents. Often cited as a reference for text clustering, the Latent Semantic Analysis [14] can briefly be resumed as a dimensionality reduction applied on a matrix where each line consists of the $tf.idf$ representation of a text of the training corpus. A similar method, the Latent Dirichlet Allocation (LDA) [3] is said adapted to short text topic classification.

More recent word representation are based on embeddings, such as *Word2Vec* [19], where each word is placed in a high dimensional space, grouping them as they occur in a learning corpus. Word2Vec has since been extended to place sentences and documents in a similar space, thanks to Doc2Vec [15]; it still builds a semantic vectorial space, while also providing a reliable way to handle complete documents instead of isolated words.

3.2 Behaviour analysis

To better explore the identity of the numerous actors, we propose to briefly review the techniques of characterisation of user behaviour. The first approach of *profiling* tries to categorise the accounts; in the second approach named *influence scores*, the goal is to evaluate the impact of each group, based on the impacts of each user.

Through crowdsourcing, the behaviour of Facebook and Twitter users was analysed through AMT[12] and extracted various psychological clusters [24]. Aiming to provide better recommendation to OSN users, TUMS[13] proposes to model accounts through the hashtags, named entities or topics emitted [29].

A temporal analysis can bring valuable information, appearing as the histogram of user activity along the hours of the day, or the days of the week [17, 34]. The strong temporal cyclic patterns of user activity along the day have

[12] Amazon Mechanical Turk: users were paid to be monitored.
[13] TUMS: Twitter-based User Modeling Service

long been underlined [35]; the division of the day in 4 (or more) periods to predict a user activity enables the detection of anomalies [9].

3.3 Cartography by communities

Finally, one efficient way to model influence in networks relies on the social groups of interaction between users. Louvain [4] is the reference algorithm for partition through modularity optimisation. Good modularities (not the global optimum) are reached relatively quickly; however it falls into the problem of "modularity resolution", as an example, by producing too large communities.

Topological graph features such as the modularity value or the density of the detected groups can be considered to evaluate a community detection algorithm. Based on a comprehensive review of scoring functions for community evaluation [33], we adapt their definitions and formalism in the following.

To measure a given community based on the topics expressed by its users, two topical metrics ξ and ρ, inspired from machine learning precision and recall, have been proposed in the literature [8]. They enable a precise characterisation of both the topical cohesion, and the group influence on a given topic.

3.4 Towards automatic risk level prediction

We presented the main features that can be automatically computed to help qualify the importance, seriousness and habit of social media content emitters: this represents a huge amount of evolving data to help the human analysts **evaluate a final credibility** (or risk) score. This evaluation can only be made with respect to a specific usecase, letting the machine learn to evaluate this risks *with* the human.

4 Content analysis and smart likelihood evaluation

While assessing the credibility of a source of information can allow to dismiss facts without analysing them, some situations cannot be solved by this first approach. In the world of media, a well respected journal might be considered a reliable source of information; yet, they occasionally may publish mistakes, or, in the case of voluntary disinformation, be manipulated or hacked into publishing false information. Moreover, an intelligence knowledge base might depend on multiple types of sources, which would most likely be approved of individually. However, these systems can sometimes be fooled by malign intent or plainly hacked into to modify their content.

In either cases, a second approach is necessary, centred on the analysis of the content of the information itself. Multiple solutions have been proposed in the past for this purpose.

4.1 Man-powered, technology-extended detection

The European project InVID[14] aims to detect inconsistencies in videos to prevent the spread of false information based on videos, allowing media to quickly detect if the outsourced videos have been manipulated. The video stream is fragmented into frames for the annotation and detection of the concepts involved in each frame, as well as for the extraction of metadata (author, creator, concepts, etc).

Other recent initiatives have been added to these research projects, such as the CrossCheck[15] project launched in 2017 with Google News Lab in partnership with more than 20 media outlets. In parallel, Facebook associated with eight French media to reduce the amount of false information on its website. A similar project had already been launched in the United States with the support of ABC News, AP, FactCheck.org, Politifact and Snopes. CrossCheck is a collaborative journalism project that brings together editorial teams from all over the world to accurately deal with false, misleading or confusing statements circulating online, studying topics, comments, images and videos.

On the machine learning side, the constitution of datasets and corpora is fed by the various reported hoaxes, as well as the satire websites, and often target celebrities or politics. From these datasets, the classic approach of feature selection (such as n-grams, and syntactic features) and classification gives promising results [25], even though clickbait and deception techniques continue to evolve, in real time.

Here, the machine learning approaches are based on recognising either the semantics for classic rumours (Flat Earth, chemtrails...) or on the syntax for patterns such as clickbait articles. It lacks an automatic knowledge exploitation.

4.2 An ambitious approach: ontologies and knowledge bases

The term ontology[22], borrowed from philosophy, refers to a method of representation describing the types of entities in the world and how they are related. It defines a common vocabulary for all actors in a system to share and process information, including machine-interpretable definitions of basic concepts in the domain and relations among them. An OWL[16] (Web Ontology Language) ontology can contain descriptions of classes, properties, and their instances, and constitutes the data structure of a **knowledge base**. Because of the relational nature of ontologies, knowledge bases are often actually represented as knowledge graphs [7], which have been the focus of research in the past few years. This representation sharpens the importance of relations between things, and add a layer of intrinsic information in the semantic of relations and entities that would

[14] `http://www.invid-project.eu/`

[15] `http://www.lemonde.fr/les-decodeurs/article/2017/02/28/lutte-contre-les-fausses-informations-le-monde-partenaire-du-projet-crosscheck_5086731_4355770.html`

[16] https://www.w3.org/OWL

not appear in a more traditional database. Approaches using knowledge bases intend to capitalise on this to add new and relevant meaning to data.

An example of a knowledge base is shown in Figure 3, displaying Sci-Fi characters. In this view, each node and edge is furthermore described by the underlying ontology: it may be stated that "starredIn" is a relation, and the object "Obi-Wan Kenobi" is a fictional character. This exhustive description of the database schema enables a precise exploitation of the content, including reasoning and systematic consistency checks.

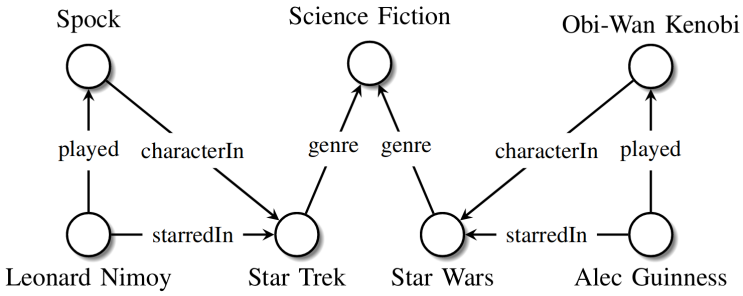

Fig. 3. Sample knowledge graph. Nodes represent entities, edge labels represent types of relations, edges represent existing relationships [20]

Indeed, the detection, identification and qualification of misinformation also takes place at the level of knowledge bases, in addition to the textual analysis that can be conducted. As an example, several pieces of text examined separately will present high likelihoods, while their juxtaposition within a knowledge base triggers the alert of misinformation.

Semantic matching models are similarity-based and compare the latent semantics of entities and relations embeddings. *RESCAL* [21] was the first to do this and has been extended multiple times. Neural network architectures have also been tried with *NTN* [28].

A specific method for the discovery of relations is based on **link prediction** techniques: for instance, translational models evaluate a fact by measuring the distance between two entities, generally using the relation during the translation. *TransE* [5] is its most known representative. More recent techniques also take time or attributes into account, like *Know-Evolve* [31] and *SLIDE* [16].

TransE [5] has been extended by Pan et al. [23] into a Binary-TransE model. They use three knowledge graphs: one based on fake news article base, one on a reliable article base and one on background knowledge from open knowledge databases such as DBpedia, to give additional information if the source is only composed of news articles. They showed that even an incomplete or imprecise knowledge graph can help detect fake news. Zhou et al. [36] demonstrated that fake news detection via NLP is vulnerable to adversarial attacks such as fact-

distortion, subject-object exchange and cause confounding. After experiments on Fakebox,[17] they suggest that the use of a crowd-sourced, dynamically updated (by local and well-informed people) knowledge graphs can improve fact-checking and stop fake news propagation at an early stage.

In short, the addition of semantic and structured data can help detect fake news better by adding prior and contextual information.

4.3 Evolving knowledge

Finally, another type of misinformation consists in achieving a modification of the base itself. The knowledge base actually evolves over time, because of the numerous, similar but different inputs. This problem can be stated as analysing the changes made between two different states of the intelligence base, evaluating the similarity between the entities of this database.

Until recently, the similarity calculation functions used are mostly based on the Jaccard index [12] and thus only compare the sets of instances between two versions of an ontology without questioning the values of the attributes and relationships. More subtle, semantic similarity calculation approaches are grounded in different criteria, such as the similarity between classes, properties, relationships, and attribute values [2, 6, 30]. The main limits of these approaches lie in the definition of similarity distances between all these elements, first between properties - which are of various types and semantic - but also to combine these in a meaningful way as a measure between entities.

Last but not least, the "Semantic concept drift analysis" consists in identifying the evolution of the meaning of the concepts used in different graphs of knowledge. This phenomenon resulted in changes in meaning, in Wikipedia, by taking into account the temporal evolutions of DBPedia, Wikipedia's reference ontology [18]. While this approach might solve some cases of the evolution of entities through time in a knowledge base, many types of families resist predictive models so far (i.e., their evolution seem unpredictable), especially for domains where the huge quantity of data required to apply deep learning techniques is simply not available.

5 Conclusion

Manipulation detection has long been a problem in politics and human sciences; it always had the power to confuse people and organisations, sometimes resulting in irrational decisions. Moreover, this problem has gained in complexity during the current technological revolution. The automation and up-scaling of media production through social media, is at the same time a greater challenge, and an opportunity.

[17] `https://towardsdatascience.com/i-trained-fake-news-detection-ai-with-95-accuracy-and-almost-went-crazy-d10589aa57c`

A greater challenge, because of the breathtaking number of stakeholders and active emitters of content; the adversaries have no difficulty to hide, they can stand among the crowd and act. Even for a well-formed mind, it is no more possible to understand the complex relations between emitters and readers on social media.

An opportunity, because machines and computer scientists provide tools and methods of high efficiency to deal with data, and turn it into consistent, up-to-date knowledge. They however need to work *with* the end-users, to finally enable a true manipulation detection, capitalising expert- and automatic- knowledges.

Bibliography

[1] B. T. Adler, K. Chatterjee, L. De Alfaro, M. Faella, I. Pye, and V. Raman. Assigning trust to wikipedia content. In *Proceedings of the 4th International Symposium on Wikis*, page 26. ACM, 2008.

[2] A. Bellenger, S. Gatepaille, and H. Abdulrab. Method allowing the fusion of semantic beliefs, June 11 2015. US Patent App. 14/391,594.

[3] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

[4] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[5] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. Translating embeddings for modeling multi-relational data. In *Advances in neural information processing systems*, pages 2787–2795, 2013.

[6] C. d'Amato, S. Staab, and N. Fanizzi. On the influence of description logics ontologies on conceptual similarity. In *International Conference on Knowledge Engineering and Knowledge Management*, pages 48–63. Springer, 2008.

[7] L. Ehrlinger and W. Wöß. Towards a definition of knowledge graphs.

[8] G. Gadek, A. Pauchet, N. Malandain, K. Khelif, L. Vercouter, and S. Brunessaux. Measures for topical cohesion of user communities on twitter. In *Proceedings of the International Conference on Web Intelligence*, pages 211–218. ACM, 2017.

[9] M. Gatti, P. Cavalin, S. B. Neto, C. Pinhanez, C. dos Santos, D. Gribel, and A. P. Appel. Large-scale multi-agent-based modeling and simulation of microblogging-based online social network. In *International Workshop on Multi-Agent Systems and Agent-Based Simulation*, pages 17–33. Springer, 2013.

[10] M. Granovetter. Threshold models of collective behavior. *American journal of sociology*, 83(6):1420–1443, 1978.

[11] B. Icard. *Lying, deception and strategic omission: definition et evaluation.* PhD thesis, Paris Sciences et Lettres, 2019.

[12] P. Jaccard. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bulletin de la Société Vaudoise de Sciences Naturelles*, 37:547–579, 1901.

[13] A. Jøsang and J. Golbeck. Challenges for robust trust and reputation systems. In *Proceedings of the 5th International Workshop on Security and Trust Management (SMT 2009), Saint Malo, France*, page 52. Citeseer, 2009.

[14] T. K. Landauer, P. W. Foltz, and D. Laham. An introduction to latent semantic analysis. *Discourse processes*, 25(2-3):259–284, 1998.

[15] Q. Le and T. Mikolov. Distributed representations of sentences and documents. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pages 1188–1196, 2014.

[16] J. Li, K. Cheng, L. Wu, and H. Liu. Streaming Link Prediction on Dynamic Attributed Networks. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining - WSDM '18*, pages 369–377, Marina Del Rey, CA, USA, 2018. ACM Press.

[17] J. Liu, P. Dolan, and E. R. Pedersen. Personalized news recommendation based on click behavior. In *Proceedings of the 15th international conference on Intelligent user interfaces*, pages 31–40. ACM, 2010.

[18] A. Meroño-Peñuela, E. Kontopoulos, S. Darányi, and Y. Kompatsiaris. A study of intensional concept drift in trending dbpedia concepts. In *SEMANTICS Workshops*, 2017.

[19] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119, 2013.

[20] M. Nickel, K. Murphy, V. Tresp, and E. Gabrilovich. A review of relational machine learning for knowledge graphs. *Proceedings of the IEEE*, 104(1):11–33, 2015.

[21] M. Nickel, V. Tresp, and H.-P. Kriegel. A three-way model for collective learning on multi-relational data. In *ICML*, volume 11, pages 809–816, 2011.

[22] N. F. Noy, D. L. McGuinness, et al. Ontology development 101: A guide to creating your first ontology, 2001.

[23] J. Z. Pan, S. Pavlova, C. Li, N. Li, Y. Li, and J. Liu. Content Based Fake News Detection Using Knowledge Graphs. In D. Vrandečić, K. Bontcheva, M. C. Suárez-Figueroa, V. Presutti, I. Celino, M. Sabou, L.-A. Kaffee, and E. Simperl, editors, *The Semantic Web – ISWC 2018*, volume 11136, pages 669–683. Springer International Publishing, Cham, 2018.

[24] E. T. Panek, Y. Nardis, and S. Konrath. Mirror or megaphone?: How relationships between narcissism and social networking site use differ on facebook and twitter. *Computers in Human Behavior*, 29(5):2004–2012, 2013.

[25] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea. Automatic detection of fake news. *COLING*, 2018.

[26] V. L. Rubin, Y. Chen, and N. J. Conroy. Deception detection for news: three types of fakes. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community*, page 83. American Society for Information Science, 2015.

[27] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1):22–36, 2017.

[28] R. Socher, D. Chen, C. D. Manning, and A. Ng. Reasoning with neural tensor networks for knowledge base completion. In *Advances in neural information processing systems*, pages 926–934, 2013.

[29] K. Tao, F. Abel, Q. Gao, and G.-J. Houben. Tums: twitter-based user modeling service. In *Extended Semantic Web Conference*, pages 269–283. Springer, 2011.

[30] M. D. Tran, C. d'Amato, B. T. Nguyen, and A. G. Tettamanzi. Comparing rule evaluation metrics for the evolutionary discovery of multi-relational association rules in the semantic web. In *European conference on genetic programming*, pages 289–305. Springer, 2018.

[31] R. Trivedi, H. Dai, Y. Wang, and L. Song. Know-evolve: Deep temporal reasoning for dynamic knowledge graphs. In *Proceedings of the 34th International Conference on Machine Learning*, 2017.

[32] G. Wang, M. Mohanlal, C. Wilson, X. Wang, M. Metzger, H. Zheng, and B. Y. Zhao. Social turing tests: Crowdsourcing sybil detection. *arXiv preprint arXiv:1205.3856*, 2012.

[33] J. Yang and J. Leskovec. Defining and evaluating network communities based on ground-truth. *Knowledge and Information Systems*, 42(1):181–213, 2015.

[34] H. Yin, B. Cui, X. Zhou, W. Wang, Z. Huang, and S. Sadiq. Joint modeling of user check-in behaviors for real-time point-of-interest recommendation. *ACM Transactions on Information Systems (TOIS)*, 35(2):11, 2016.

[35] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann. Time-aware point-of-interest recommendation. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, pages 363–372. ACM, 2013.

[36] Z. Zhou, H. Guan, M. Bhat, and J. Hsu. Fake News Detection via NLP is Vulnerable to Adversarial Attacks:. In *Proceedings of the 11th International Conference on Agents and Artificial Intelligence*, pages 794–800, Prague, Czech Republic, 2019. SCITEPRESS - Science and Technology Publications.