Are Attributes on Social Media Platforms Usable for Assisting in the Automatic Detection of Identity Deception?

E. van der Walt and J.H.P. Eloff

Department of Computer Science, University of Pretoria, South Africa
e-mail: estee.vanderwalt@gmail.com, eloff@cs.up.ac.za

Abstract

Social Media Platforms (SMPs) allow any person to easily communicate with their friends or the general public at large. People can now be targeted at great scale, most often for malicious purposes. The mere fact that more people are using SMPs exposes more people to various forms of cyber threats such as cyber-bullying. The problem is that many of these cyber-attacks involve some form of identity deception, where the attackers lie about who they are. The solution proposed in this paper is to work towards developing a model for Identity Deception Detection (IDD) on SMPs by identifying and using metadata that is freely available on SMPs. This metadata includes attributes that describes a user account on an SMP. The aim is to use only these attributes, as opposed to the contents of a communication, for determining if people are lying about their identities. By discarding contents, an identity deception detection model can be developed with lower overhead. A prototype is discussed that runs an experiment using the metadata (attributes) that defines the identity of a user on an SMP. The results show promise for further research in developing solutions for assisting with the automatic detection of identity deception.

Keywords

Cyber-security, identity deception, fake identities, social media, big data, Twitter

1 Introduction

Today, it is said that over 4 billion people, or more than half of the world's population have access to the Internet and that more than 3.1 billion interact with other people on social media platforms (SMPs) (Chaffey, 2018). This phenomenal growth in the number of online identities resulted in new social interaction abilities on SMPs, being added on a daily basis, with the intention to benefit society at large. Some examples of these benefits include the tracking of natural disasters (Chun et al., 2014) and the prediction of public crowd gatherings (ben Khalifa et al., 2016). However, this growth in the number of online identities not only brought social benefits but also facilitates activities that are deceitful and potentially harmful to societies. Consider for example the online activities in February 2018 where 13 Russians were charged by the United States Justice Department for subverting the 2016 political campaign (Apuzzo and LaFraniere, 2018). They created social media accounts as if they were American citizens with the assumed intention to create discord in the democracy system through the content they posted. In these types of deceitful activities attackers lie about their online identities by providing false information for account attributes on SMPs.

Within the cyber-security world, these types of activities are commonly known as impersonation or identity deception (Donath, 1999). The growth in the number of online identities on SMPs and the resulting voluminous of the data has made it very difficult, if not impossible, to know who to trust on SMPs (Ribeiro et al., 2016). Furthermore, humans are gullible and do not, for example, have the ability to discern the truth from lies (Sandy et al., 2017). SMPs on the other hand allow malicious humans, to deceive (Cook et al., 2014).

The cyber-threat of malicious individuals together with the intrinsic vulnerability of SMPs increase the risk for humans to be exposed to identity deception. Most of the countermeasures available today for minimizing this risk can either be classified as of legal or technological nature. The Children's Internet Protection Act (CIPA) (Kierkegaard, 2008) is a good example of a legal instrument but unfortunately does not address the illegal or harmful use of fake identities. Besides legal instruments, various technologies have been proposed to assist in the protection of humans against identity deception on SMPs, for example, plugins (Rashidi and Fung, 2016), APIs (Muller and Thiesing, 2011), and software systems (Egele et al., 2013). These technologies differ in who they protect from, what deception they can detect, and the various methods used to detect identity deception.

This paper focuses on the technological aspect of countering the act of identity deception. For this paper in particular, an attempt is made to determine attributes of accounts on SMPs that have the potential to assist in the automatic detection of identity deception. The contributions of the research results reported on in this paper are summarized as follows:

- To identify the attributes freely available on SMPs that can play a role in detecting identity deception through a literature review.
- To implement and execute an experiment based on supervised machine learning for assisting the automatic detection of identity deception.

Section 2 of this paper identifies existing identity related attributes found on SMPs. The section furthermore discusses how these attributes have been applied in related work on identity deception detection. This discussion leads into a definition for the requirements, such as to use content from humans only, expected of a prototype aiming to assist in the automatic detection of identity deception on SMPs by humans in section 3. Section 4 proposes a high-level design for the prototype. Sections 5 and 6 explicate and discuss the experimental results following the prototype's implementation.

2 Background and related work

Many examples of cyber threats that have materialised in real-life incidents can be found on SMPs, such as a woman who was falsely lured through Facebook to be killed (de Villiers, 2017). In these cases, the attackers lied by changing various of their social media account attributes that defines their identities to hide who they are. SMP data are mostly known for the content added by its users. Past research used the content itself to detect non-humans accounts, also known as bots (Dickerson et al., 2014)

(Rashidi and Fung, 2016). A challenge was held by DARPA in 2016 to detect bots on Twitter specifically (Subrahmanian et al., 2016). The overall conclusion was that a large set of initial bots can be detected through rules based on heuristics, behaviours, linguistics, and inconsistencies. Noteworthy from the DARPA challenge was that not only the content was used to detect these bots. Besides posting content, SMP users are required to open an account with the SMP before they can start posting content (Facebook, 2017). During this registration process they are requested to give information like their name (Facebook, 2017), location (Twitter, 2018), and even birth date in some cases (LinkedIn, 2017). This data is also generally referred to as metadata or attributes (Sloan et al., 2015). These attributes not only identify the user but also serves to distinguish them from another user. Take Twitter for example. In Twitter, the name of the user and the location are examples of attributes describing the user. It is noticeable that the same attributes are found across the different SMPs. This indicates that a proposal towards detecting identity deception could somehow also apply to other SMPs.

Past related work proposed various identity attributes, and also combined some identity attributes to engineer new features, to detect identity deception. Feature engineering is the process of using domain knowledge to construct new pieces of information (Domingos, 2012). In this case, the attributes available in SMPs are used to create new information about the identity of a user. Lee et al. (Lee et al., 2010), Ribeiro et al. (Ribeiro et al., 2018), and Thomas et al. (Thomas et al., 2013) used linguistic features extracted from various SMPs to detect identity deception. Examples of such linguistic features are: the collection of specific words (Ribeiro et al., 2018), repetitions of content (Lee et al., 2010), and sharing the same naming structure, for example "JohnSmith" being very similar to "JohnSmit2" (Thomas et al., 2013). Non-verbal attributes like the date the account was opened (Tsikerdekis, 2017), the type of SMP (Thomas et al., 2013), and profile update time (Gurajala et al., 2016) were useful where the information provided for an account is scarce. Network features, like accounts in the same domain (Thomas et al., 2013), friends (Gurajala et al., 2016), and followers (Gurajala et al., 2016) were used to detect deception. Lastly, identity attributes like gender (Hancock and Toma, 2009), location (Alowibdi et al., 2015), profile image (Hancock and Toma, 2009), age (Tuna et al., 2016), profession (Tuna et al., 2016), name (Peddinti et al., 2017), and email (Xiao et al., 2015) were proposed indicators towards detecting identity deception. Many of the attributes used to detect identity deception, required additional processing to extract knowledge about the identity of a person. For example, the content had to be parsed for specific words to determine sentiment (Ribeiro et al., 2018) and each profile image was manually labelled to determine of that person was an adult or not (Tuna et al., 2016). This additional work required, adds overhead to a model proposing to assist in the automated detection of human identity deception on SMPs.

Cresci et al. (Cresci et al., 2015) and Varol et al. (Varol et al., 2017) used a combination of attributes and features in their research with the aim of reducing the overhead required to develop an identity deception detection model. They showed that the identity and non-verbal attributes were not only easy to mine, but also just as accurate at detecting identity deception for bots, compared to using network, linguistics, or other content related features. Even though Cresci et al. (Cresci et al., 2015) and Varol

et. al (Varol et al., 2017) focussed on detecting deceptive non-human accounts on SMPs, these same SMP attributes apply to humans. For this reason, the authors propose to use the identity and non-verbal attributes on SMPs in an experiment to not only assist in the automated detection of human identity deception, but also to understand which attributes are more indicative of such deceptiveness. The following common attributes, amongst others, were identified in SMPs, like Facebook (Facebook, 2017) and Twitter (Twitter, 2018), through each platform's API reference: the name of the user, their profile image, their location, status description, and the date they created their SMP account.

Cresci et al. (Cresci et al., 2015) furthermore proposed machine learning algorithms like decision tree, random forest, support vector machines (SVMs), adaptive boosting, k-nearest neighbours and logistic regression for their research experiments. Gupta et al. (Gupta et al., 2013) in turn suggested Naïve Bayes and decision trees to detect bots successfully. Xiao et al. (Xiao et al., 2015) proposed logistic regression, random forests, and SVMs to detect deceptive accounts. The related work, although focussed on detecting bots, not only had success in detecting deceptive identities on SMPs, but also used the attributes freely available on SMPs. For these reasons, this paper will use supervised machine learning as a method to develop a model that can assist to detect identity deception by humans on SMPs.

3 Establishing the requirements

The following requirements for the research presented in this paper have been accumulated through related work:

- Use a dataset that consists of a large volume of heterogeneous data, created at high velocity. SMPs, being a big data platform, are a good source of such data (Van der Walt and Eloff, 2015).
- Use only attributes freely available on an SMP (Twitter, 2018) (Facebook, 2017) (LinkedIn, 2017).
- Ignore non-human accounts in the SMP data (Cresci et al., 2015).
- Ignore content posted by users on an SMP (Varol et al., 2017).
- The attributes used for the model, should describe the identity of the person (Meligy et al., 2017).
- Develop a supervised machine learning model (Cresci et al., 2015)
- The data should contain both examples of deceptive and trustworthy people. Supervised machine learning requires a labelled dataset (Kuhn et al., 2016).
- Compare the results from various machine learning models (Varol et al., 2017).
- Automate the detection due to SMPs' big data nature (Chaffey, 2018).

The next section provides a high-level design for the prototype.

4 High-level design of a prototype for the automated assistance of identity deception detection on SMPs

To describe the components of the prototype, the Unified Modelling Language (UML) is employed. UML is a visual modelling language for systems (O'Regan, 2017). It helps to define a prototype during the design phase instead of during development. This approach not only describes the prototype at the beginning of development, but also minimizes the risk of the prototype not complying with the requirements and only finding this out at the end of the development. For this prototype, the authors propose three components:

- Prepare – For the prototype, freely available SMP attributes are available. The attributes should describe the identity of the user and not include any content they posted. The data should also contain examples of both deceptive and trustworthy accounts. To adhere to these requirements, this component retrieves the data from Twitter as an example of an SMP, cleans the data from any non-human accounts, labels the data for supervised machine learning, and finally prepares the data for supervised machine learning.
- Discover – Supervised machine learning is required to build and evaluate models assisting in the detection of human identity deception on SMPs. This component allows for experimentation by using the prepared data to train various supervised machine learning algorithms using different parameters, such as resampling (Domingos, 2012), and hyperparameters (Dickerson et al., 2014).
- Detect – Due to the nature of the data, more specifically its volume and heterogeneity, the process if identity deception detection should be automated. This component allows for unassisted identity deception detection and uses the most accurate machine learning model discovered during experimentation.

For this research, the proposed prototype was built using infrastructure provided by the Future SOC Lab in Potsdam, Germany (FSOC, 2018). The Twitter data was mined using Apache Flume (Apache, 2018) and stored in a SAP HANA (SAP, 2017) in-memory database consisting of 2TB of RAM and 8TB of storage. Machine learning models were built using the Caret package in R (Kuhn et al., 2016). The prototype components, their functions, and how each component addresses the requirements expected of a prototype assisting in the automated detection of human identity deception on SMPs are illustrated in Figure 1. The next section shows some results delivered by the running prototype.

5 Experimental results

Identity attributes from Twitter accounts were mined, using a Java API (Yamamoto, 2018) together with Apache Flume (Apache, 2018) during 2016. Apache Flume was able to import volumes of data whilst ignoring non-English-speaking accounts, before sending the data to SAP HANA for further processing. 224 796 Twitter accounts and 606 million tweets were finally stored in SAP HANA. 53 091 of the Twitter accounts were discarded at this point, using rules from the research of Cresci et al. (Cresci et al., 2015). These discarded accounts were found to belong to non-human or bot

accounts. The tweets were ignored as the prototype requires only those attributes describing the identity of the person. An additional 15 000 deceptive human accounts were manually fabricated and injected into the corpus to create a labelled corpus of deceptive and non-deceptive accounts. These injected examples of deceptive accounts each had one or more identity attributes not representative of the truth. For example, the profile image would be of someone else or the location would be a place different from their indicated GPS location. The attributes in the final prepared dataset only contained those attributes describing the identities of both trustworthy and deceptive humans.

The experiment executed with the prototype, used the aforementioned prepared identity data. The results from this experiment, using supervised machine learning and 10-fold cross validation (Peddinti et al., 2017) (Fire et al., 2014), is shown in Table 1. Given PR-AUC, which measures the Precision-Recall performance of a machine learning model (Davis and Goadrich, 2006), it is shown that at best, the Adaboost and nnet (neural net) algorithms detected identity deception by humans with a score of 0.317 and 0.306 (1 being the best, 0 being the worst) respectively.

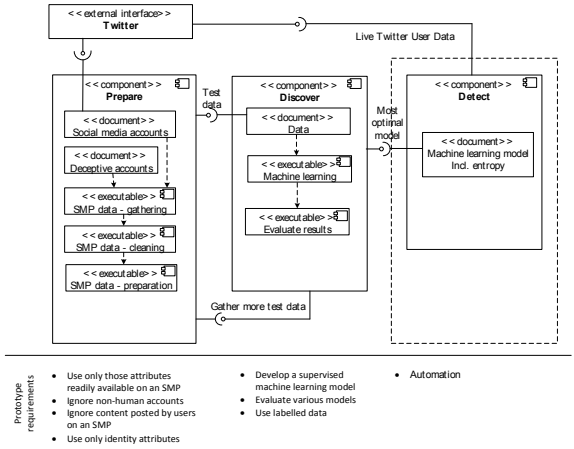

Figure 1: UML component diagram of the proposed prototype

Algorithm	Accuracy	Kappa	F1-score	ROC-AUC	PR-AUC
random forest	0.800	0.236	0.331	0.755	0.280
Adaboost	0.788	0.186	0.287	0.702	0.317
nnet	0.730	0.170	0.282	0.734	0.306

Table 1: Results from the experiment

6 Discussion of results

For this paper, the Accuracy, Kappa, F1 score, and AUC (Area Under the Curve) was considered to assist in the evaluation of the models. The F1 score and AUC metrics are often used in research detecting spam and bot accounts to determine the effectiveness of the machine learning models (Ferrara et al., 2016) (Fire et al., 2014) (Xiao et al., 2015). Although 10-fold cross validation was used to train the models, it is known that the F1 score and ROC-AUC (Receiver Operator Characteristics Curve) suffers (Menardi and Torelli, 2014) (Jeni et al., 2013) in skewed distributions. More recently the PR-AUC (Precision-Recall Curve) has been recommended as an alternative to ROC-AUC (Saito and Rehmsmeier, 2017) (Davis and Goadrich, 2006). Based on all the information provided, the PR-AUC was regarded as the final metric to evaluate the machine learning models with. The results were still suboptimal as the AUC for a random predictor equals 0.50 (Powers, 2011).

The following recommendations are proposed to address the weak model performance results and to improve the prototype:

- Experiment with additional features to increase the accuracy of the prototype. For example, by combining SMP attributes like whether the gender on the profile image matches the gender of the SMP user, further lies can potentially be identified. These attributes should still be freely available on SMPs.
- Improve the completeness of attributes on SMPs as many identity attributes were found to be incomplete i.e. not completed by the users at the time of creating the user account. If some of these attributes, like location and profile image were made compulsory by the SMP provider, identity deception detection accuracy could potentially increase.
- Additional validation could be performed by SMPs upon user registration to ensure the veracity of SMP attributes. By, for example, getting someone else to validate that the profile image is representative of that user, could prevent potential identity deception.

7 Conclusion and future work

This paper showed how, besides the content, many attributes exist on SMPs that could be indicative of human identity deception. A prototype was proposed, showing how metadata (attributes) freely available on SMPs can be used to train supervised machine learning models. It is also shown by means of the experimental results provided in this paper how these supervised machine learning models can be a step in the right direction for assisting with the automatic detection of humans lying about their identities on SMPs. Furthermore, the results also uncovered the influence of a skewed labelled dataset and the difficulty in using only the meta-data (attributes) describing the identity of a human on SMPs. It was found that many of the identity attributes were incomplete and it was therefore difficult to create an accurate model to assist in the automated detection of identity deception by humans on SMPs. Future work will focus on increasing the accuracy of the machine learning models. One way of achieving this will be to introduce engineered features, additional to the attributes, such as “age-determined-from-profile-image”.

8 References

Alowibdi, J. S., Buy, U. A., Philip, S. Y., Ghani, S. & Mokbel, M. 2015. Deception detection in Twitter. *Social Network Analysis and Mining,* 5**,** 1-16.

Apache. 2018. Flume. Available: https://flume.apache.org/ [Accessed 16 May 2018].

Apuzzo, M. & Lafraniere, S. 2018. 13 Russians Indicted as Mueller Reveals Effort to Aid Trump Campaign. *The New York Times*, 16 Feb 2018.

Ben Khalifa, M., Redondo, R. P. D., Vilas, A. F. & Rodríguez, S. S. 2016. Identifying urban crowds using geo-located Social media data: a Twitter experiment in New York City. *Journal of Intelligent Information Systems***,** 1-22.

Chaffey, D. 2018. *Global social media research summary* [Online]. Smart Insights. Available: https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/ [Accessed 23 Jun 2018].

Chun, Y., Hwang, H. S. & Kim, C. S. 2014. Development of a Disaster Information Extraction System based on Social Network Services. *nternational Journal of Multimedia and Ubiquitous Engineering.*

Cook, D. M., Waugh, B., Abdipanah, M., Hashemi, O. & Abdul Rahman, S. 2014. Twitter Deception and Influence: Issues of Identity, Slacktivism, and Puppetry. *Journal of Information Warfare,* 13**,** 58-71.

Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A. & Tesconi, M. 2015. Fame for sale: efficient detection of fake Twitter followers. *Decision Support Systems,* 80**,** 56-71.

Davis, J. & Goadrich, M. The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd international conference on Machine learning, 2006. ACM, 233-240.

De Villiers, J. 2017. Suspects use fake Facebook profile to lure women, rape and kill them. *News24* [Online]. Available: https://www.news24.com/SouthAfrica/News/suspects-use-fake-facebook-profile-to-lure-women-rape-and-kill-them-20171104 [Accessed 4 Nov 2017].

Dickerson, J. P., Kagan, V. & Subrahmanian, V. Using sentiment to detect bots on Twitter: Are humans more opinionated than bots? Advances in Social Networks Analysis and Mining (ASONAM), 2014 IEEE/ACM International Conference on, 2014. IEEE, 620-627.

Domingos, P. 2012. A few useful things to know about machine learning. *Communications of the ACM,* 55**,** 78-87.

Donath, J. S. 1999. Identity and deception in the virtual community. *Communities in cyberspace,* 1996**,** 29-59.

Egele, M., Stringhini, G., Kruegel, C. & Vigna, G. Compa: Detecting compromised accounts on social networks. NDSS, 2013.

Facebook. 2017. The Facebook Graph API. Available: https://developers.facebook.com/docs/graph-api/overview [Accessed 8 Jan 2018].

Ferrara, E., Wang, W.-Q., Varol, O., Flammini, A. & Galstyan, A. Predicting online extremism, content adopters, and interaction reciprocity. International Conference on Social Informatics, 2016. Springer, 22-39.

Fire, M., Kagan, D., Elyashar, A. & Elovici, Y. 2014. Friend or foe? Fake profile identification in online social networks. *Social Network Analysis and Mining,* 4**,** 1-23.

Fsoc. 2018. The HPI Future SOC lab. Available: https://hpi.de/en/research/future-soc-lab.html [Accessed 8 June 2018].

Gupta, A., Lamba, H., Kumaraguru, P. & Joshi, A. Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. Proceedings of the 22nd international conference on World Wide Web, 2013. ACM, 729-736.

Gurajala, S., White, J. S., Hudson, B., Voter, B. R. & Matthews, J. N. 2016. Profile characteristics of fake Twitter accounts. *Big Data & Society,* 3**,** 2053951716674236.

Hancock, J. T. & Toma, C. L. 2009. Putting your best face forward: The accuracy of online dating photographs. *Journal of Communication,* 59**,** 367-386.

Jeni, L. A., Cohn, J. F. & De La Torre, F. Facing imbalanced data--Recommendations for the use of performance metrics. Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on, 2013. IEEE, 245-251.

Kierkegaard, S. 2008. Cybering, online grooming and ageplay. *Computer Law & Security Review,* 24**,** 41-55.

Kuhn , M., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z. & Al., E. 2016. caret: Classification and regression training. R package version 6.0-73.

Lee, K., Caverlee, J. & Webb, S. The social honeypot project: protecting online communities from spammers. Proceedings of the 19th international conference on World wide web, 2010. ACM, 1139-1140.

Linkedin. 2017. LinkedIn Developers. Available: https://developer.linkedin.com/ [8 Jan 2018].

Meligy, A. M., Ibrahim, H. M. & Torky, M. F. 2017. Identity Verification Mechanism for Detecting Fake Profiles in Online Social Networks.

Menardi, G. & Torelli, N. 2014. Training and assessing classification rules with imbalanced data. *Data Mining and Knowledge Discovery***,** 1-31.

Muller, F. & Thiesing, F. Social networking APIs for companies—An example of using the Facebook API for companies. Computational Aspects of Social Networks (CASoN), 2011 International Conference on, 2011. IEEE, 120-123.

O'regan, G. 2017. Unified Modelling Language. *Concise Guide to Software Engineering.* Springer.

Peddinti, S. T., Ross, K. W. & Cappos, J. 2017. Mining Anonymity: Identifying Sensitive Accounts on Twitter. *International AAAI Conference on Web and Social Media.* Montreal, Canada.

Powers, D. M. 2011. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.

Rashidi, B. & Fung, C. BotTracer: Bot user detection using clustering method in RecDroid. Network Operations and Management Symposium (NOMS), 2016 IEEE/IFIP, 2016. IEEE, 1239-1244.

Ribeiro, M. H., Calais, P. H., Santos, Y. A., Almeida, V. A. & Meira Jr, W. Characterizing and Detecting Hateful Users on Twitter. Twelfth International AAAI Conference on Web and Social Media, 2018 Palo Alto, California, USA. AAAI Press, 676-679.

Ribeiro, M. T., Singh, S. & Guestrin, C. Why should i trust you?: Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016. ACM, 1135-1144.

Saito, T. & Rehmsmeier, M. 2017. Precrec: fast and accurate precision–recall and ROC curve calculations in R. *Bioinformatics,* 33**,** 145-147.

Sandy, C., Rusconi, P. & Li, S. 2017. Can Humans Detect the Authenticity of Social Media Accounts? *3rd IEEE International Conference on Cybernetics (CYBCONF).* Exeter, UK.

Sap. 2017. SAP HANA. Enterprise Edition. Available: https://www.sap.com/uk/developer/topics/sap-hana.html [Accessed 10 June 2018].

Sloan, L., Morgan, J., Burnap, P. & Williams, M. 2015. Who tweets? Deriving the demographic characteristics of age, occupation and social class from Twitter user meta-data. *PloS one,* 10**,** e0115545.

Subrahmanian, V., Azaria, A., Durst, S., Kagan, V., Galstyan, A., Lerman, K., Zhu, L., Ferrara, E., Flammini, A. & Menczer, F. 2016. The darpa twitter bot challenge. *IEEE Computer Magazine,* 49**,** 38-46.

Thomas, K., Mccoy, D., Grier, C., Kolcz, A. & Paxson, V. Trafficking Fraudulent Accounts: The Role of the Underground Market in Twitter Spam and Abuse. USENIX Security, 2013. Citeseer, 195-210.

Tsikerdekis, M. 2017. Identity Deception Prevention Using Common Contribution Network Data. *IEEE Transactions on Information Forensics and Security,* 12**,** 188-199.

Tuna, T., Akbas, E., Aksoy, A., Canbaz, M. A., Karabiyik, U., Gonen, B. & Aygun, R. 2016. User characterization for online social networks. *Social Network Analysis and Mining,* 6**,** 104.

Twitter. 2018. Twitter API. Available: https://dev.twitter.com/overview/api [Accessed 8 Jan 2018].

Van Der Walt, E. & Eloff, J. H. P. 2015. Protecting minors on social media platforms - A Big Data Science experiment *HPI Cloud Symposium "Operating the Cloud".*

Varol, O., Ferrara, E., Davis, C. A., Menczer, F. & Flammini, A. Online human-bot interactions: Detection, estimation, and characterization. Eleventh International AAAI Conference on Web and Social Media, 2017 Montreal, Canada. 280-289.

Xiao, C., Freeman, D. M. & Hwa, T. Detecting clusters of fake accounts in online social networks. Proceedings of the 8th ACM Workshop on Artificial Intelligence and Security, 2015. ACM, 91-101.

Yamamoto, Y. 2018. Twitter4J. Available: http://twitter4j.org/en/ [Accessed 16 May 2018].