DOI 10.29042/2018-4094-4099

Helix Vol. 8(5): 4094- 4099

Reliability Index for Twitter – Twitter Handles' Credibility Assessment

***[1]Kartik Sharma, [2]Sanjeev Dhawan, [3]Kulvinder Singh**
[1]Research Scholar of Computer Engineering,
[2,3]Faculty of Computer Science & Engineering
[1, 2,3]University Institute of Engineering & Technology, Kurukshetra University, Kurukshetra, Haryana, India
*Email: *[1]kartik_s@outlook.in, [2]rsdhawan@rediffmail.com, [3]kshanda@rediffmail.com*

Received: 09th July 2018, Accepted: 14th August 2018, Published: 31st August 2018

Abstract
Social networks have attained a vital place in everyone's lives by emerging as the accustomed means for social interaction among people as well as the primary medium for dissemination of information. Accordingly, these platforms also bear the responsibility of such information being reliable and authentic. In order to endorse such information, the social networks need to vet the sources of such information which in turn means to remove or flag potentially fake accounts as well as automated bot accounts. A social network's trustworthiness is directly proportional to its efficiency in spotting malicious accounts.
In this paper, we propose a fresh approach to assign a reliability measure to every single account on Twitter. In order to compute an account's reliability measure, we factor in certain parameters which are paramount in ascertaining any twitter account's trustworthiness along with their weightage which varies for every such parameter.

Keywords: Social Networks, Twitter, Fake Accounts, Sybil, Reliability Index, Account Credibility

Introduction

Social Networks being a cardinal part of our lives, bring as many threats with them as the convenience they offer. Since quite some time, social networks have befallen as the primary source of information, news, communication etc for a majority of our population. People have started counting on social media for a range of things starting from basic social interaction to their day to day news. In this era, where social networks have emerged as the primary bridge between people, we need to consider the perils of this as well as ways to encounter those. Nowadays, anything can be found on social networks, ranging from personal posts to worldwide news, be it factual or false. There is hardly any way to ascertain the genuineness of such information except for its source profile. Unfortunately, as of now, there is no unswerving way to establish the credibility of an account which offers any such information. In this paper, we propose an approach to do just that. We aim to provide a way to determine a social account's trustworthiness by scrutinizing an assortment of the subject's profile particulars. In our approach, we employ a point system for measuring any account's reliability in quantitative terms. Based on the resulting numerical value, we can easily determine if an account is worth our trust or not along with the degree to which we can rely on the account in question. In the subsequent sections, we will elucidate on our approach along with the careful functioning of the same [1, 2, 3, 4, 5].

Materials and Methods

Proposed Methodology

In this paper, we propose a novel approach named 'Reliability Index for Twitter', which effectively assigns every twitter handle with a numeric value which can be used to judge a profile's authenticity. The reliability measure is a numeric value set between 0 and 1, where 0 signifies an extremely unreliable and possibly fake account while 1 denotes a reliable and possibly genuine, human-operated account. While both of those values are virtually hard to come by, the reliability index would most certainly be a rational number ranging from 0.0 to 1.0. The value is attained by setting up parameters to be taken into consideration for computation of the reliability measure and assigning numeric weightage to the parameter's possible instances, be them constructive or detrimental.

Finding and choosing those parameters is the primary step in this approach while assigning the right weightage to each such parameter is the most crucial one as the efficiency of the system depends on this. If we omit the critical parameters or assign less value to the critical parameters or high value to the trivial ones, then we most certainly botch up the efficiency of the system. In other words, the effectiveness of the system especially depends on the choice of parameters along with their respective weightage. The aforementioned approach is a conceptual one to be employed by Twitter without disclosing the itemized view of the determination process of the same. This approach can be applied to any social network but for every social network, new parameters need to be set along with their respective weightage.

In our approach, we take 20 parameters into consideration in order to calculate a reliability measure for Twitter accounts. Those 20 parameters have been carefully selected and act as partial deciding factors in the calculation of the reliability factor. Those parameters have been coupled with their respective desired values as well as a weightage which is added to the raw reliability value if the desired condition is met. In our model, the weightage ranges from a minimum value of 4 to a maximum value of 33. The parameter associated with the value '4' is account age which holds the number of months an account has been active for. We have set the desired condition to be more than two months. If an account's age is more than two months, a value of 4 is added to the raw reliability measure value which is then divided by 300 to attain the final reliability measure or a scaled value of 0.01 is added to the final reliability measure. Either way, the final value ranges between 0 and 1.

Parameters used in our approach are listed herewith with a brief explanation:

1. *Verified Status*

This parameter deals with the fact, if the twitter account is verified or not. The desired condition for this parameter would be if the account is verified by Twitter and hence bearing a blue tick on the profile page. Earlier research work corroborates this parameter being a full proof one since Twitter itself reckons upon this parameter with a proper modus operandi.

2. *Followers*

The number of followers an account has is a very important factor when making a distinction between genuine and bogus accounts. In our approach, if an account has more than 20 followers, it works in favor of the account in terms of the reliability index [6, 7, 8, 9].

3. *Following*

In this parameter, the number of accounts which are being followed by the subject account is taken into consideration. If an account follows more than 20 other accounts, then the condition for this parameter is satisfied. This parameter weighs less than the previous one in the determination of the reliability measure, as most fake accounts run lower in the number of followers than the number of followings [6, 7, 8].

4. *Tweets*

The quantity of tweets published by an account has considerable substance when assessing its reliability. In our model, we consider the number of tweets to be 30 for being a positive influence on the reliability measure of the account in question [10].

5. *Media Tweets*

Apart from the number of tweets, the content of tweets matters as well. As most phony accounts either have only a few tweets with hardly any with media attached to them, media tweets become pertinent. A Twitter account with minimum five media tweets qualifies with respect to this parameter [10, 11, 12].

6. *Likes*

Twitter allows users to like other user's tweets which then are shown on the user's profile page separately from other tweets. This feature is different from retweets and weighs highly on the reliability factor as bogus accounts are not accustomed to like tweets. The borderline for this parameter has been set to 15 in our proposal [9].

7. *Liked Tweets*

Like our last parameter, Twitter also has a record of tweets of an account which have been liked by other Twitter accounts. That quantity also comes in handy whilst calculating the reliability measure.

8. *Retweets*

Retweets are used to forward or repost a tweet from another user. These belong to the same cluster of parameters as likes and liked tweets and bear an analogous importance as a factor in our calculation. We set this threshold to 25 retweets in order to satisfy this parameter condition [11].

9. *Retweeted Tweets*

Similar to the citation in the preceding parameter, this one deals with the record of tweets which have been retweeted by other Twitter handles. The qualifying figure w.r.t. this parameter has been set to a minimum of 10 in favor of the approach's success.

10. *Account Age*

A twitter account's creation date is commonly a matter of public record and the account's age can be construed from that information. The unit used for this parameter is months in our model. We consider the account's age a minor but imperative factor in determining the account reliability, the threshold value has been set to two months for this clause [13].

11. *Verified Mobile Number*

Twitter allows us to link our mobile numbers with our account using a verification code in order to enhance the accessibility as well as security. A verified mobile number certainly works in the favor of a twitter handle's reliability with a high weightage.

12. *Verified Email Address*

Similar to the above factor, this one is fulfilled if a verified email address is linked to the twitter account at hand. It has comparatively less weightage with respect to a verified email address.

13. *The ratio of Followers and Following*

This is the most vital and decisive parameter in affinity to the rest of them. It refers to the mathematical ratio of the number of followers and the number of following accounts. In our model, we've set the deciding ratio eligibility to be 0.8.

14. *Hashtags Used*

This factor checks if there has been any usage of hashtags in the tweets till date. Presence of hashtags is regarded as a positive influence on the reliability score. [11]

15. Bio Added

If an account has been personalized with a bio then it surely works in the favor of the reliability index since most bogus accounts do not care enough to add that kind of detail in their profile.

16. Mobile Application Attachment

The presence of a linked mobile application on a phone, be it Android or iOS, has a positive impact on the reliability measure of an account. This is a bit similar to the presence of a verified mobile number, but it weighs more standing since this action is a little more personalized than just adding a mobile number and usually, only the accounts used by actual humans feel the need to use an application for better accessibility.

17. Two Factor Authentication

Two-factor authentication adds an extra layer of security into the access of accounts and is hardly ever activated in case of bots or phony accounts. Due to the aforementioned reason, it is a high-value parameter. However, we cannot solely rely on this parameter as even a lot of genuine accounts do not have this option enabled, but it definitely makes a decent addition to an account's reliability measure.

18. Location Tweets

Tweets may contain a lot of data in themselves apart from the text it has. It may contain media as well as the location data of the place the tweet was authored at. In our experience, bots and malicious accounts usually do not care enough to turn this feature on, which is why it can be used to make a distinction between the two.

19. Contacts Uploaded

Twitter provides its users with the facility to upload their contact lists in order to find their friends and acquaintances in a more efficient way by using this information. Even though it has no direct relationship with the accounts being reliable or not, we can still use this as a parameter as this feature is mostly used by genuine accounts only.

20. Reported Fake

There are often instances when an account is reported as fake or malicious by other users on Twitter, often upon an encounter with other accounts. This parameter depends solely on the reporting account's judgment. In our model, if an account has never been reported, then it tops off the reliability score to a considerable extent [14].

The above-mentioned parameters are the very basis of our model and help us achieve numeric reliability measure for every account. However important the above parameters may be, they are of little or no use when used alone and cannot be relied upon for judgment with respect to an account's reliability status. However, when clubbed together, they can act as a full proof way for establishing the same.

As stated earlier, we assign every parameter with a desired value as well as a numeric value to be added to the reliability score after the condition is met. We have taken a different value for every parameter according to their importance and relevance. The same can be ascertained in the below-mentioned table.

Attributes	Desired Condition	Weight	Scaled Weight
Verified Status	Yes	10	0.03
Followers	>20	9	0.03
Following	>20	5	0.02
Tweets	>30	10	0.03
Media Tweets	>5	16	0.05
Likes	>15	17	0.06
Liked Tweets	>5	17	0.06
Retweets	>25	6	0.02
Retweeted	>10	18	0.06
Account age	>2	7	0.01
Mobile Number Verified	Yes	21	0.07
Email Address Verified	Yes	12	0.03
The Ratio of Followers and Followed	>0.8	35	0.12
Hashtags Used	Yes	6	0.02
Bio Added	Yes	12	0.04
Mobile Application Attached	Yes	16	0.05
Two Factor Authentication added	Yes	21	0.07
Location Tweets	Yes	16	0.05
Contacts Uploaded	Yes	13	0.04
Reported Fake	No	33	0.11
		300	1

Figure 1: Twitter Reliability Index Formula with Parameters

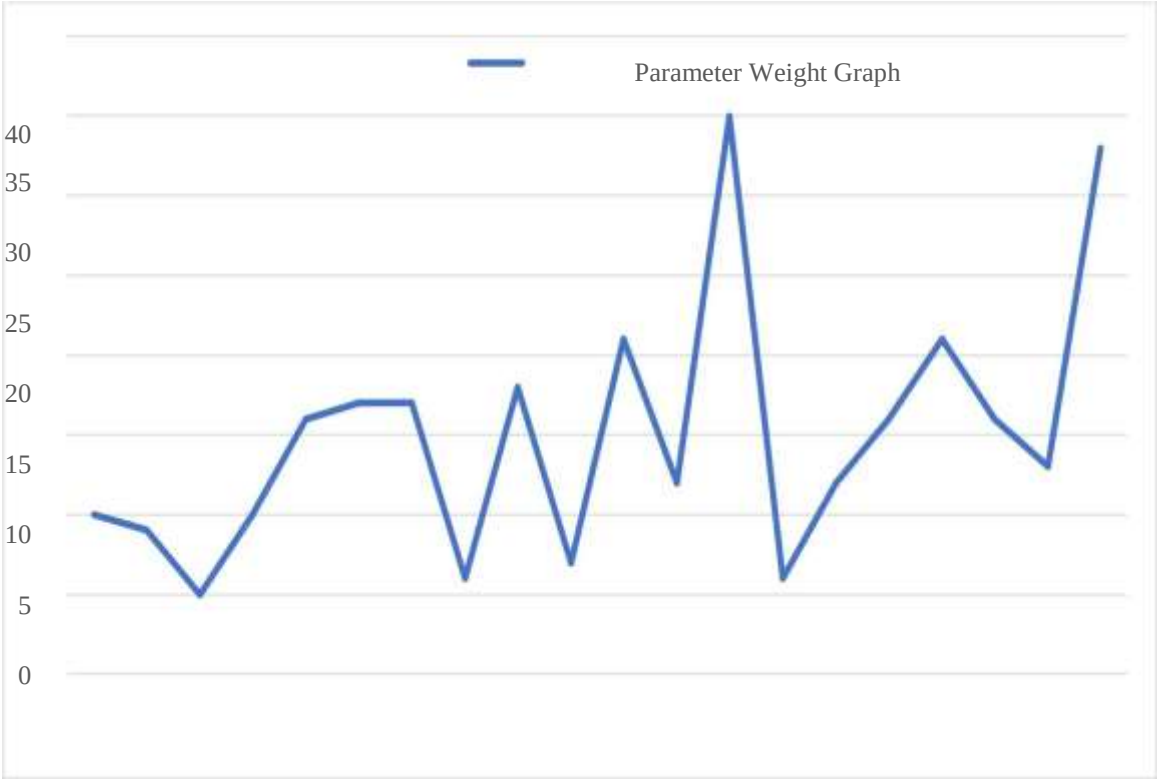

Figure 2: Parameter Weight Graph

Implementation

The approach we proposed for finding out the reliability index of accounts has been tailored for Twitter. It is supposed to be implemented on Twitter's end rather than the user's end as a few of the parameters we used are not a matter of public record. We propose that this model be implemented by Twitter on its end with the score of each profile being visible publicly whilst keeping the process and specific parameter specifics behind the curtains in order to preserve privacy. We implemented the same on real life twitter accounts with the help of a data set and Python Data Analysis Library.

Data Set

No dataset was available on the internet for twitter or any other social network due to the latest changes in their terms and conditions. We found an old dataset from the Stanford University Large Dataset collection (SNAP) but like any other such options available, that was obsolete and gigantic as well, with hardly any relevance. We ended up extracting data directly from Twitter using its API coupled with Python's sentiment analysis. Our dataset contained actual twitter accounts present on Twitter on 3rd May 2018. For the purpose of our concept, we had to append a few details to the extracted data set which weren't available publicly.

Tools and Technologies Used

For the practical realization of our concept, we had to hinge on the Python's Twitter Sentiment Analysis, Natural Language Toolkit (NLTK) as well as its Data Analysis Library – Python Pandas (v.0.23.0) [15, 16]. We used Python 3.6.5, being the latest release. Python is a very strong choice when it comes to text processing and the availability of numerous libraries and supporting tools for specialized purposes also comes in handy.

Python Data Analysis Library provides us with efficient and easy to use data structures as well as analyses tools [17]. These technologies deliver effectual as well as accurate tools for extracting and processing semantic information with minimum hassle and small footprint.

Results and Discussion

The proposed approach, when directed exerted upon two diverse data sets, accomplished convincing results. The subject twitter handles were successfully discerned as reliable ones and questionable ones with optimum accuracy. In this line of research, there has been minimal to no work done with a similar approach. The parameters and weightage considered in our proposal is an outcome of rigorous research in the area with respect to the literature available on the same. Latterly, the parameters being considered and their weightage are vastly accountable for the results achieved.

Conclusion and Future Scope

Social Networks are still way behind in vanquishing the threat of fake accounts, bots (spambots etc), pseudonym accounts, impersonators etc. Our approach is merely a footstep in the course of thwarting such adversaries. It is tailor-made for Twitter and efficaciously delivers a quantifiable reliability factor of twitter handles after a careful analysis of various parameters. The competence of our model depends on the choice of parameters and their weightage, which can be altered to formulate a more or less efficient system. In near future, our approach may be successfully applied to any social network subject to amendments in parameters and their weightage. Our research is an outcome of the contemporary need for dealing with the menaces of fraudulent account and information. There have been numerous efforts in this area of research but none of them have been full proof, which calls for more focus in this prominent area of research.

References

[1] J. Castellini, V. Poggioni and G. Sorbi, "Fake Twitter followers detection by denoising autoencoder," in *Proceedings of the International Conference on Web Intelligence*, 2017.

[2] Z. Chu, S. Gianvecchio, H. Wang and S.Jajodia, "Who is tweeting on Twitter: human, bot, or cyborg?," in *Proceedings of the 26th annual computer security applications conference*, 2010.

[3] F. Benevenuto, G. Magno, T. Rodrigues and V. Almeida, "Detecting spammers on twitter," in *Collaboration, electronic messaging, anti-abuse and spam conference (CEAS)*, 2010.

[4] A. Pathak, "An analysis of various tools,methods and systems to generate fake accounts for social media," *North eastern University Boston, Massachusetts December*, 2014.

[5] A. Mehrotra, M. Sarreddy and S. Singh, "Detection of fake Twitter followers using graph centrality measures," in *Contemporary Computing and Informatics (IC3I), 2016 2nd International Conference on*, 2016.

[6] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi and M. Tesconi, "A Fake Follower Story: improving fake accounts detection on Twitter," *IIT-CNR, Tech. Rep. TR-03,* 2014.

[7] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi and M. Tesconi, "Fame for sale: efficient detection of fake Twitter followers,"*Decision Support Systems,* vol. 80, pp. 56-71, 2015.

[8] M. Egele, G. Stringhini, C. Kruegel and G. Vigna, "Compa: Detecting compromised accounts on social networks.," in *NDSS*, 2013.

[9] Y. Zhang and J. Lu, "Discover millions of fake followers in Weibo," *Social Network Analysis and Mining,* vol. 6, p. 16, 2016.

[10] C. Xiao, D. M. Freeman and T. Hwa, "Detecting clusters of fake accounts in online social networks," in *Proceedings of the 8th ACM Workshop on Artificial Intelligence and Security*, 2015.

[11] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen and B. Y. Zhao, "Detecting and characterizing social spam campaigns," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, 2010.

[12] J. P. Dickerson, V. Kagan and V. S. Subrahmanian, "Using sentiment to detect bots on twitter: Are humans more opinionated than bots?," in *Advances in Social Networks Analysis and Mining (ASONAM), 2014 IEEE/ACM International Conference on*, 2014.

[13] C. A. Davis, O. Varol, E. Ferrara, A. Flammini and F. Menczer, "Botornot: A system to evaluate social bots," in

Proceedings of the 25th International Conference Companion on World Wide Web, 2016.

[14] Y. Shen, J. Yu, K. Dong and K. Nan, "Automatic fake followers detection in chinese micro-blogging system," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2014.

[16] W. McKinney, Python for data analysis: Data wrangling with Pandas, NumPy, and IPython, " O'Reilly Media, Inc.", 2012.

[17] W. McKinney, "pandas: a foundational Python library for data analysis and statistics," *Python for High Performance and Scientific Computing*, pp. 1-9, 2011.

[18] W. McKinney and P. D. Team, "Pandas—Powerful Python Data Analysis Toolkit," *Pandas—Powerful Python Data Analysis Toolkit*, p. 1625, 2015.

[19] J. A. Teixeira da Silva, "Fake peer reviews, fake identities, fake accounts, fake data: beware!," *AME Medical Journal*, vol. 2, 2017.

[20] L. Sloan, J. Morgan, W. Housley, M. Williams, A. Edwards, P. Burnap and O. Rana, "Knowing the tweeters: Deriving sociologically relevant demographics from Twitter," *Sociological research online*, vol. 18, pp. 1-11, 2013.

[21] P. W. L. Fong, "Preventing Sybil attacks by privilege attenuation: A design principle for social network systems," in *Security and privacy (SP), 2011 IEEE symposium on*, 2011.

[22] E. Ferrara, O. Varol, C. Davis, F. Menczer and A. Flammini, "The rise of social bots," *Communications of the ACM*, vol. 59, pp. 96-104, 2016.

  ISSN 2319 – 5592 (Online)