



IDENTIFYING FAKE USER'S IN SOCIAL NETWORKS USING NON VERBAL BEHAVIOR

M.BALAAANAND¹, R.SOWMIPRIYA², S.SIVARANJANI³, S.SANKARI⁴
V.R.S College of Engineering and Technology, Arasur.
¹balavdy@gmail.com, ²priyasowmi94@gmail.com,
³ranjanisaminathan@gmail.com, ⁴4sankaricse14@gmail.com,

ABSTRACT

Big data is a term with large or complex dataset where traditional data processing applications are inadequate. Challenges include analysis, capture, curation, search, sharing, storage, transfer, visualization, and information privacy. Now a days, Usage of social networks has been increasing which leads the path for users to have multiple accounts for sharing their thoughts. On the other hand, Fakers pretend like an authorized users and cause discomfort to the users during the brainstorming sharing. Verbal behavior encourages to identify the fake and multiple account holders in the social networks only among active users. The optimality of the solution is not guaranteed if algorithm alone applies on the problem, since the resource shared to the deceptive user is also same as the authorized user. In this [project] paper we proposing the nonverbal behavior to identify the deceptive users in the social networks like Facebook, Twitter, LinkedIn etc. Using Deceptive detection algorithm we are identifying the fake users and improving the resource utilization with high reliability and performance in the social networks

Index Terms— Algorithm, illusory, identity, performance, social networks.

I.INTRODUCTION

In past decade we have experienced an incrementing level of interest in online gregarious media, which enables users to not only engender content but withal exchange it utilizing Web 2.0 technologies[2]. Social network usage has increased by 64% since 2005 [3]. The facileness with which we can engender online profiles at a low cost has additionally led to ample opportunities for identity deception, which at times can have fatal consequences. Other gregarious media accommodations such as collaborative projects have to engage in “cat-mouse” games by perpetually having to block utilizer accounts for individuals joining in with different account names not long after a block has been applied. Solutions have been proposed that can assist in detecting multiple accounts owned by the same individual but their effectiveness vary in terms of computational efficiency and complexity of practical implementation depending on the availability of the appropriate data [5], [6]. Moreover, these past methods have mainly focused on detecting deception through

verbal communication (e.g., speech or text) and have ignored the potential of non-verbal (e.g., user activity or movement) deception detection, which has shown high success rates in the offline world [7], considering that non-verbal cues are 4.3 times more powerful than verbal cues in face-to-face communication [8]. This is a promising detection method that we have identified in our previous work and for which we presented experimental results in [9].

In this paper we propose a novel approach that makes use of user non-verbal behavior data in social networks in order to detect multiple account and fake identity deception. The rest of the paper is organized as follows. In Section II, we present an overview on deception and identity deception, and discuss some of the problems with current identity deception detection methods and highlight the research contributions of this paper. In Section III, we describe our proposed method. Section IV presents the performance results obtained

with our proposed method. Finally, Section V discusses the implications of our proposed technique in the growing field of identity deception detection for the social media domain.

II. RELATED WORKS AND CONTRIBUTIONS OF THIS WORK

A. Illusory and Identity Deception

Deception has been defined as the deliberate transfer of false information to a recipient that is not aware that the information received has been falsified [6]. Similarly, human deception is motivated by instrumental (goal-driven), relational (relationship-driven) and identity-driven goals. Online, the success of an attempt to deceive others is dependent upon multiple factors associated with the components involved: deceiver, social medium, potential victim and deceptive action [9]. Factors that affect a deceiver's behavior and effectiveness in achieving deception include a deceiver's expectations, goals, motivations, his/her relation to target and a target's degree of suspicion. A deceiver's goal is to use everything at his/her disposal to keep a low suspicion from his/her target and this applies to both verbal and non-verbal behaviors.

The deceptive action transmitted through cyberspace also has attributes such as the number of targets and the expiry date associated with it that influence its success [9]. An important factor is a victim's Information Communication Technology (ICT) literacy [9]. Deception is achieved by manipulating content, the communication channel, the sender information, or any combinations of these three components [9]. Manipulating content involves tampering with images or even text as can be seen in collaborative projects such as Wikipedia where special user task forces are focused on monitoring for text manipulation with the intention to spread inaccurate information [5]. Identity deception (a subcategory of deception) focuses on manipulating the sender's information and can be divided into three categories: identity concealment (e.g., concealing or altering part of an individual's identity), identity theft (e.g., mimicking another person's real identity) and identity forgery (e.g., forging a fictional identity) [6].

B. Deception Detection

Deception detection theories are divided into those that are based on *leakage cues* (cues sent by the deceiver unwillingly due to factors such as cognitive overload) and *strategic decisions* (cues indicative of deception that are willingly transmitted by a deceiver in order to ensure deception success). To detect deception, both categories pick up cues from verbal and non-verbal communications. Human deception detection is arguably the most widely used method. Individuals can pick up

cues from the environment in which an interaction takes place (e.g., a photo-graph that looks edited) with a deceiver and interpret these cues by understanding a deceiver's goals. The most critical factor in detecting deception is the time, which can vary from days to months, until a truth is uncovered by a previously deceived individual. However, people are bad at detecting deception with detection success bounded between 55 to 60 percent at best while others have measured an even lower success of 34 percent. Even more troublesome is that a study has found that upon training people in detecting verbal and non-verbal cues detection accuracy actually decreased. A more standardized perspective of examining deception detection is necessary to achieve and engineer deception detection solutions with high success rates. Three of the most popular theories used in the deception field are Interpersonal Deception Theory (IDT), Leakage Theory (LT), and Expectancy Violations Theory (EVT).

C. Identity Deception Detection

A particular issue with identity deception in social media is the presence of multiple identities by one user. Both online and offline studies have been conducted in an attempt to solve the problem of detecting duplicate account records. The most direct solution to identify duplicates in a database with the highest accuracy is a cross-comparison for the full length of accounts in a database. A more recent study by Solorio et al. attempted to detect sock puppets (these are new accounts of previously blocked users) on Wikipedia [5]. They used natural language processing techniques to detect users who maintain multiple accounts based on their verbal output. Textual features were used such as punctuation count, quotation count or the variation between using capital or lowercase "I". These features were tested against all revisions made by the users on pages throughout Wikipedia. Due to the volume of users on Wikipedia in conjunction with the number of revisions that each account may have (which can reach thousands), the similarity-based method used to identify a positive match between two accounts needs to receive manual input (an individual needs to set which two accounts need to be compared).

As such, the method can be considered as a human-augmenting deception detection technique since it requires individuals to provide input for two potential accounts that match. A Support Vector Machine (SVM) model has shown 68.83% overall accuracy against an experimental dataset of 77 cases of legitimate users and sock puppets. The limitation of this method is its computational cost involved if one would like to test all accounts against all accounts in a database.

D. Contributions of This Work

The main contributions of this work can be summarized as follows:

- We propose a computationally efficient method (applicable to all social media classifications [2]) for detecting identity deception through the use of non-verbal user activity in the social media environment. This contribution ensures that a relatively high level of overall detection accuracy is obtained that is comparable to similar methods that make use of verbal communication [5], [6] but with lower computational overheads.
- To demonstrate the computational efficiency (to withstand the immense traffic experienced by social media services) of our proposed non-verbal method to deception detection we use publicly available data from Wikipedia and machine learning algorithms.
- Finally, we present design guidelines for designers and developers interested in implementing this method as an added level of security for their social media communities and additional considerations based on various social media classifications in existence today.

III. PROPOSED METHOD FOR DETECTING ONLINE IDENTITY DECEPTION

A. Research Objectives

Our research goal in this work is to develop a method that can automatically detect online identity deception which can be very useful in many online social media scenarios. For instance, one scenario where such detection would be useful is in the case of an open source software development collaborative project website where, for security reasons allows just one account per individual. Since new account registration is available to anyone, a user can therefore register an unlimited number of times every time his/her account gets blocked. Succeeding in identity deception is important for a deceiver who wants to inject malicious code into a project. Once an account is discovered, all changes made to the code by the owner of that account will be investigated and closely examined. We argue that an early detection system can help identify those individuals who experience a disproportionate familiarity with the collaborative software (according to their non-verbal behavior), which may indicate that they are not in fact newcomers or novices. Post-examination and close monitoring of suspect cases will help ensure the security of an open source project. To demonstrate our proposed method's effectiveness we use Wikipedia, which falls under the collaborative projects classification of social media [1] as our experimental case. We used publicly available data for

Wikipedia in order to evaluate our approach. It is worth pointing out that although we have used Wikipedia as an example of a social medium, our method can be applied to virtually any other social medium environment. We briefly describe below some of the non-verbal user activities that can be observed on Wikipedia before describing our proposed method.

B. The Wikipedia Environment

Wikipedia is a free online encyclopedia in which everyone can contribute without an account (anonymously when only IP address is visible) and with an account using a pseudonym or even real name. Wikipedia operates on the concept of namespaces where each namespace is meant to include a specific type of content. Wikipedia's policy pages and discussion on Wikipedia proposals or projects belong to different namespaces. Wikipedia has 28 namespaces. A single user interaction with the Wikipedia's environment and two of its namespaces. The logged data on page revisions provide us with non-verbal user behavior on Wikipedia. For example, the time taken between each revision is a measurable non-verbal behavior.

C. Non-Verbal Behavior Variables

We used simple and more complex variables to represent user behavior. Variables of online non-verbal behavior fall under two major categories: time-independent and time-dependent (henceforth these variables are denoted with index t).

D. Data Retrieval and Model Testing

We collected a list of all publicly available logs of blocked users on Wikipedia. The logs include various reasons for blocking user accounts including account blocks for verified sock puppet cases. Using regular expressions we kept only sock puppet cases with an infinite time of block issued for these accounts. On average it takes approximately 75 days for a sock puppet account to get blocked (median is 3.19 days) as evident in our block log dataset. About 38.96 percent of sock puppets have their accounts blocked during the first day after their first revision on Wikipedia. Ten days after their first revision, the percentage of sock puppets being caught rises to 62.24. By 30 days, the percentage rises to 74.43.

For testing our proposed method we sampled 7,500 cases of sock puppets. In addition, we retrieved a list of all users who made at least one revision through the revision records on all Wikipedia namespaces (these are provided as dump xml files and were parsed). Verified sock puppet cases were removed from this list and an additional sample of non-blocked users was obtained so that our final user list contained 7,500 verified sock

puppet cases and 7,500 legitimate user cases. As such, a fair coin toss for our sample would produce approximately 50% accuracy in detecting sock puppets. Human deception detection is usually placed at much lower rates (as low as 30-50%).

IV. PERFORMANCE EVALUATION

We used a popular set of machine learning algorithms, which includes support vector machine (SVM), Random forest (RF) and Adaptive boosting (ADA), to implement our proposed models.

A. Performance Metrics Used

To evaluate the efficiency of our models for our proposed method we used the following classification matrix. Using this matrix, we derive results to measure the following performance metrics in order to evaluate the performance of our models for our proposed method: *recall* (the fraction of valid sock puppet cases that are returned), *precision* (the fraction of returned cases that are valid sock puppet cases), *F-measure* (the test of a model's accuracy bounded between 0 and 1 that combines recall and precision), *accuracy* (the fraction of true positives and true negatives returned over the total number of cases), *false positive rate* (indicating the rate of falsely identified sock puppets), and *Matthews Correlation Coefficient (MCC)*.

B. Experimental Procedure

To evaluate the performance efficiency of our models for our proposed method we repeated ten times a ten-fold cross-validation procedure to obtain the mean values for all of our performance metrics.

V. DISCUSSION OF RESULTS

Based on the results obtained, we found that Adaptive Boosting appears to provide the best balance between recall and precision whereas maintaining the highest achieved accuracy. Recall levels are relatively high. If the detection method is used to report suspect cases so that administrators can keep a close eye on or restrict certain features for suspect accounts for a time period, then recall is the most important feature and low precision can be tolerated. The results obtained show that the use of non-verbal user activity is a viable and efficient method for detecting identity deception (specifically sock puppetry). Moreover, although we have used Wikipedia as an example of a social medium, this deception detection method can be applied to other social media domains. In fact, the detection method can be used with any social media service that contains user foot prints that are not only verbal (e.g., text, audio, video) but also non-verbal.

A. Limitations of Our Proposed Detection Method

The efficiency and effectiveness of our proposed detection method is influenced by several context specific factors. It can also affect the efficiency if the window is too large given that more data will be needed to be examined by the detection method. Finally, the social medium under examination will also determine the data that can be used. It is worth pointing out that although our method is portable to any social media classification, adaptations may be needed to ensure its proper implementation.

B. Future Work

Future work will need to examine other non-verbal behavior variables in different social media services that can be used as good indicators of deception. Moreover, combining research on verbal detection deception with the non-verbal behavior deception detection method presented in this study may help improve prediction accuracy.

VI. CONCLUSION

Non-verbal behavior monitoring for deception detection is an alternative path that can be used as a leading or complimentary detection solution. A coordinated effort is required to test these solutions on different platforms and advance the field of social media identity deception detection.

REFERENCES

- [1] Michail Tsikerdekis, ansherali Zeadally, Senior member, IEEE.
- [2] A.M Kalpana and Haenein, "Users of the world, unite! The challenges and opportunities of social media," *Business Horizons*, vol.
- [3] C.S. Bhat, "Cyberbullying: Overview and all school personnel" *Australian J.Guid Counselling*, vol.18, no.1 pp. 53-56, Jul 2008.
- [4] T.Solorio, R. Hasan and M.Mizan, "A case study of sockpuppet detection in Wikipedia," in *proc.Workshop Lang.Anal. Social media*, 2013, pp.59-68
- [5] G.A Wang, H. Chen, J. J. Xu, and H.Atabaksh, "Automatically detecting criminal identity deception: An adaptive detection algorithm," *IEEE Trans.Syst., Man C.Cybern. A.Syst., Humans.*, vol.36. no. 5, pp.988-999, sep.2006.
- [6] T. O .Meseryy et al, "Deception detection through automatic, unobtrusive analysis of nonverbal behavior." *IEEE Intell syst.* vol.20, no.5, pp. 36-43, sep/Oct.2005.
- [7] M. Argyle, V.Salter, H. Nicholson, Williams, and P. Burgess, "The Communication of interior and superior

- attitudes by verbal and non-verbal signals, British J. Social Clinical Psychol., vol, no.3 .pp. 222-231, sep.s1970.*
- [8] D.B.Buller and J.K.Burgoon. "Interpersonal deception theory," *Common Theory*, vol.6, no.3, pp. 203-242, Aug.1996.
- [9] J .S. Donath "Identify and deception in the virtual counity," in *Comunities in Cyberspace*, M. A. Smith and P. Kollock, Eds. London, U.k : Routledge, 1999.
- [10] M.Tsikerdekis and S.Zeadally,"Online deception in socialmeadial," *Comomon ACM*, vol.57, no.9, sep.20.