A Survey on Sockpuppet Detection over Collaborative Projects

Roshan Kumari
M.Tech. Scholar
Department Of Computer Science & Engineering
ABES Engineering College, Ghaziabad

Saurabh Kr. Srivastava
Sr. Asst. Professor
Department Of Computer Science & Engineering
ABES Engineering College, Ghaziabad

ABSTRACT
Sockpuppet are the fake identity used by some malicious users. In Wikipedia, there are number of sockpuppets present. Two different online accounts, belongs to the same person are called sockpuppet. Sockpuppets has become significant issues, in which one can have fake identity for some specific purpose or malicious use. Sockpuppet detection are based on binary classification in which given accounts are classified either sockpuppet or non-sockpuppet category. This research synthesizes sockpuppets detection in which various approaches are discussed to identify the fake account and legitimate account.

Keywords
Social media, Sockpuppets, Non-Sockpuppets, Binary classification, Wikipedia and authorship attribution

1. INTRODUCTION
In recent years social media has become most important part of human being as increasing use of internet. Collaborative projects are a platform where two or more people come together and share their knowledge and information related to some specific domain. Collaborative projects like Wikipedia has not mandatory to have one person one account. So there are some malicious user who creates some fake account and use these accounts for personal benefits or malicious use. Sockpuppet has become one of the most important issues over collaborative projects. One can have new account with the help of less information. So it's necessary to have some specific method to find out these sockpuppet cases or suspicious cases because its violates privacy or personal information. As we know, Wikipedia does not provide any specific facility to detect such malicious accounts, so the current process is done manually which is time consuming and cost effective. So identify such accounts as sockpuppet in Wikipedia is an important issue. Sockpuppets detection is based on binary classification in which two classes are predefined and classification process assigns instances either as sockpuppets or non-sockpuppets. To identify such sockpuppet account is a challenging task. Multiple Identity deception (Sockpuppet) has become an increasingly important issue in the social media environment. 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. Fig 1 shows the entities involved in sockpuppet detection.

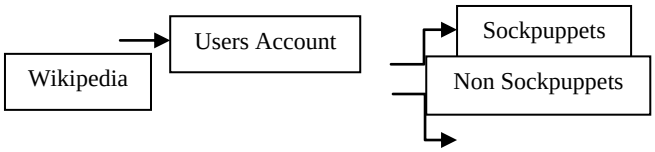

Fig 1: Entity involved in sockpuppet detection

2. RELATED WORK
In the context of sockpuppet detection, Thamar Solorio et.al.[1] has contributed their work and supported experimentally with small case study of sockpuppets detection using authorship attribution. In this approach, unique writing style of a user is identified. On the basis of user's writing style, comparison between suspicious accounts and actual user accounts are identified. It's important to have focus on the talked pages since the articles edited. In Wikipedia, as one person can create multiple accounts with providing less information. So whenever a user in bad faith, vandalizes some existing articles or create some fake articles, the user is banned for editing new account. If a user is banned then next time he/she will again try to create new fake account and try to edit articles. These accounts are called sockpuppets. The primary account is called Sockpuppeteer. Authors detected typical features of authorship attribution. After analyzing authors comment some new features that includes stylist grammatical and formatting preferences of the authors. On the basis of these features such suspicious accounts are discriminated. Some features are total number of characters, total number of sentences, total number of tokens, words without vowels, total alphabet count, total punctuation count, total contraction count, parenthesis count etc. they described top 30 features on the basis of their rank using Information gain. The results are calculated with the help of machine learning algorithm (support vector machine in Weka) sockpuppet cases are identified. In experimental result authors have described that the accuracy of this approach is 68%. After that, authors have again conducted experiment to show the performance of the same method with and without timing features for the problem of sockpuppet cases. So average confidence (84%) of classification and F-Measure (72%) increases with timing features. Authors have taken data for experiment, are described in table 1.

Table 1: Data Taken For Experiment

Training Data	Testing Data
Sockpuppeteer (Blocked Accounts)	Sockpuppet Accounts

Thamar Solorio et.al.[2] has described a corpus of sockpuppet cases from Wikipedia. A Corpus provides a real world dataset of short messages from malicious users. This corpus can be

used by researchers working on authorship attributes problem. Sockpuppet investigations in Wikipedia (SPI) are done using support vector machine to identify confirmed SPI cases and Denied SPI cases. In SPI cases, investigation are requested by real author to detect real account or suspected account. When an SPI concludes with a confirmed sockpuppet cases then sockpuppet accounts are banned. The author has used list of features[13]. Authors has taken 623 cases in a ten cross validation for experiment. So the best results achieves F-Measure (73%) using all features.

Xueling Zheng et.al.[3] proposed two approaches for identification of sockpuppet cases. First approach is based on the relative number of replies between suspected sockpuppets. While second approach is based on to retrieve some keywords from the posts to identify sockpuppet cases. Two methods are described, one algorithm for sockpuppet detection in one discussion forum and second algorithm for sockpuppet detection in two different forums. In first algorithm, authors has defined detection score to find out the total number of topics posted by one account and the relative number of replies from the other account with respect to all his replies. As the detection score grows larger, the more likely the accounts belong to the same person. For the threshold of 1.5, the accuracy is smaller than 0.34 without using the active account time but the accuracy increases to more than 0.9 with active time. In second approach, author created a keyword-based profile of the accounts based on the corresponding posts. It has been found that the two accounts keyword similarity is 0.96. Zaher Yamak et.al.[4] has proposed a method for automatic detection of sockpuppets on Wikipedia using non-verbal behavior. Multiple account are created by manipulator for several manipulations (Sybil, information manipulation, social spams). In collaborative projects it's very important to detect a manipulator who manipulates things for their personal benefits with the help of multiple accounts. Four steps are defined to detect sockpuppet cases: Data extraction from Wikipedia, account selection for the positive and negative sockpuppet accounts, feature extraction and algorithm training and testing. Result obtained the best accuracy using Random Forest(RF) 99.8% and Bayesian Network (BN) 99.6%. Michail Tsikerdekis et.al.[5] has proposed a novel approach for multiple identity deception using non-verbal behavior. Nonverbal communication (user activity or movement) are more powerful than Verbal communication(speech). Identity deception focuses on manipulating the senders information and can be divided into 3 categories-identity concealment, identity theft and identity forgery. Identity Deception focuses on manipulating the sender's information. A major issue with identity deception in social media is the presence of multiple identities by one user. Wikipedia is a free online encyclopedia in which everyone can contribute with an account or without an account. So it is easy to make multiple id and use for some malicious use. Logs of blocked users on Wikipedia during the period since February 2004 until October 2013 are taken for experiment. Two categories of non-verbal behavior are considered for experiment- One time dependent and time independent. Authors have used non-verbal behavior for detection of multiple identity (sockpuppet).

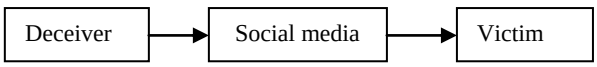

Fig 2: Entities Involved In Online Deception

The best result obtained by Adaptive Boosting approach which provides balance between recall and precision. Overall accuracy achieved by this method is 71.3%. Dhanyasree P. et. al.[6] has proposed a method for sockpuppet detection which use the non-verbal behavior of the user with less detection time and less time complexity. Social media gives its user a great freedom to create more than one account with the help of less information. But it becomes serious issues when someone uses these accounts for some intentional purpose or misuses it. Fig 2 shows the actual entities involved in online deception. It's necessary to have some detection method to detect that one person has only one account and he or she does not involve in forgery cases. In this paper non-verbal behavior are used to detect which is more effective than verbal behavior. The author uses Random Forest and PSO for result evaluation. Sheetal Antony et.al.[7] has proposed a system architecture which combined both verbal and non-verbal behavior. Deception means when people transfers wrong information purposefully to cheat a person. The author has proposed a system that can use verbal and nonverbal behavioral patterns to detect identity deception. There is an admin who manages each account for users. The details and activities of the user are analyzed and detect if there is some deception. The details are verified in database. If it detects that there is some deception then there are some security questions that are asked to users. M Balaanand et.al.[8] has proposed a approach that makes use of non-verbal behavior data in social networks in order to detect multiple account and fake identity. Authors have used user behavior under two major categories: time dependent and time-independent. Obtained result shows the use of non-verbal user activity is an efficient approach for detection of sockpuppetry. Table 2 describes the confusion matrix. with the help if this matrix, precision, recall, accuracy, f-measure, false positive rate and Matthews correlation coefficient(MCC) are derived to test the proposed model. Antu Mary Kuruvilla et.al.[9] proposed method for detecting identity deception by a single user is based on using nonverbal behavior. Some Wikipedia users create multiple accounts and use them for various malicious purposes such as creating fraudulent articles, damaging existing article text etc. So these deceptions cannot easily detected by any authority. Numerous methods have been proposed that can help in detecting multiple accounts owned by the same persons. Non-verbal explains about activities done by each user separately are such as: Number of articles generates, Number of searches done for same articles, Number of bytes added and also removed, Number of times same spelling mistakes carryout constantly and Time taken between each revision.

Table 2: Table representing Confusion matrix for sockpuppet and non-sockpuppet

	Sockpuppet	**Non-Sockpuppet**
Predicted identity deception(Sockpuppet)	True positive	False positive
Predicted legitimate user(Non-Sockpuppet)	False negative	True negative

Michail Tsikerdekis et.al. [10] has contributed their work on online identity deception over social media. Deception means someone transferring wrong or fake messages where recipient is not aware that he/she is getting wrong messages .On social media it's easy for someone to do deception using fake identity for some specific purpose. Deception can be defined in terms of content deception, sender's deception, communication channel deception and hybrid deception. Authors have used algorithms such as Information Manipulation Theory (IMT), Interpersonal Deception Theory

(IDT) and Leakage Theory (LT). Author proposes that there should be some standard rule for users registration and verifying users credentials. A summary of related work are discussed in table 3.

Dataset: All authors have taken dataset from WIKIPEDIA. Taken dataset consist some blocked account (sockpuppets) and some active account. On the basis of these accounts experiment has done and results are presented in summarized format in mentioned table 3.

Performance Measures: There are some measures, on the basis of these measures results are discussed

- Accuracy
- Precision
- Recall
- F-Measures
- True Positive Rate(TPR)
- False Positive Rate(FPR)
- Matthews Correlation Coefficient(MCC)

Table 3: A summary on related work on Sockpuppet detection

Publications	Algorithms	Done Work
Thamar Solorio et. at. [2013, 2014]	SVM	On the basis of authorship attributes sockpuppet cases are detected. Authors have tried to identify sockpuppet cases by comparing writing style of actual author and fake author. Each comment made by a user is considered a "document". There are two steps, in first step, predictions from the classifier on each comment has done. Then in second step, predictions for each comments and combine them in majority voting schema to assign final decisions to each account. The paper presented 68% accuracy of this approach. The author extended his work and on the basis of corpus available on real world deceptive writing, sockpuppets are detected. SPI (Sockpuppet Investigations) are considered for experiment. The reported best result gives 73% F-Measure.
Xueling Zheng et. al. [2011]	Detection Score & Keyword Similarity Method	Sockpuppet pairs are detected on the basis of detection score .The first one is designed for detecting those sock puppets pairs in the same discussion forum while the second one is for detecting sockpuppets pairs that appear in two different forums. In the reported results Sockpuppet pairs are detected.
ZaherYamak et. al. [2016]	SVM, RF, NB, KNN, Bayesian Network & Adaptive Boosting	Authors have tried to select feature for sockpuppet detection on the basis of three sets: contribution behavior, other users behavior towards these contribution and account behavior. Best accuracy given by Random Forest (99.8%) and Bayesian Network (99.6%).
Michail Tsikerdkis et. al. [2014]	SVM, RF & Adaptive Boosting (ADA)	Authors tried to detect multiple account identity deception (sockpuppet) on the basis of non-verbal behavior. It has been observed that non-verbal behavior (user activity and user movement) is more effective than verbal behavior (speech and text). Higher accuracy rate achieved by Adaptive Boosting. Overall accuracy achieved by this method is 71.3%.Author extended his work and addressed the challenges for identity deception and safe social media environment for interaction. He supported that some standard rule for user's registration and verification credentials needs to be investigated. That can minimize the sockpuppetry cases.
M. Balaanad et. al. [2015]	SVM, RF & ADA	Authors have proposed method for detection of multiple account on social media like Facebook, Twitter, LinkedIn etc. Used nonverbal behavior under time -dependent category and time-independent category. Adaptive Boosting gives the best balance between recall and precision with high accuracy.
Dhanyasree P. et. al. [2016]	RF & PSO	Authors have proposed a method using non-verbal behavior of the user on social media with less detection time. Proposed method achieves less time complexity. Both the verbal and nonverbal behavior has been combined and used for sockpuppets detection.
Sheetal Antony et.al. [2016]	Effectiveness of verbal	The author proposed a safe social media environment for users so that they can share their information safely. It has been observed that combination of

	and non-verbal behavior	verbal and non-verbal behavior gives better result in comparison to using individually. Author explains the method to detect deception and provide security in social media using the verbal and non-verbal behavior of a user.
Antu Mary et. al. [2015]	Effectiveness of non-verbal behavior	Using verbal and nonverbal behavior of user we can easily detect the sockpuppet with limited amount of time. Following attributes are considered by the author for non-verbal behavior: number of articles generated, number of searches done for same articles, number of bytes added and removed, number of times same spelling mistakes carryout constantly and time taken between each revision.

3. CONCLUSION AND FUTURE WORK

This paper presented an overview for sockpuppet detection over collaborative projects in which sockpuppet issues are significant issues. On the basis of verbal and non-verbal behavior, it can be easy to detect sockpuppet accounts made by fake users. In this survey, it can be observed that non-verbal behavior is more effective than verbal behavior over collaborative projects (Wikipedia etc.). There are not strict rule to have one person one account. So easily one can make fake account for malicious use. Non- verbal behavior is an effective method which is supported by many authors for sockpuppet detection. In future direction, time complexity can be detected on the basis of verbal and non-verbal behavior. Comparison among all approaches can be discussed in more detailed manner.

4. REFERENCES

[1] Thamar Solorio, Ragib Hasan and Mainul Mizan, "A Case Study of Sockpuppet Detection in Wikipedia", Proceedings of the Workshop on Language in Social Media(LASM 2013),Pages 59-68,Atlanta,Georgia,June 13 2013.@2013 Association for Computational Linguistics**.**

[2] Thamar Solorio, Ragib Hasan and Mainul Mizan, "Sockpuppet Detection in Wikipedia: A Corpus of Real-World Deceptive Writing For Linking Writing", arXiv:1310.6772v1[cs.CL] 24 Oct 2013.

[3] Xueling Zheng, Yiu Ming Lai, K.P. Chow, Lucas C.K. Hui and S.M. Yiu, "Detection of Sockpuppets in Online Discussion Forums", HKU CS Tech Report TR-2011-03.

[4]ZaherYamak, Julien Saunier and Laurent Vercouter, " Detection of Multiple Identity Manipulation in Collaborative Projects", IW3C2, WWW'16 Companion, April 11-15, 2016, Montreal, Quebec, Canada. ACM 978-1-4503-4144-8/04.

[5] Michail Tsikerdekis et. al., "Multiple Account Identity Deception Detection in Social Media Using Non Verbal Behavior", IEEE Transactions on Information Forensics and Security, Vol 9, No 8, August 2014.

[6] Dhanyasree P*, Sajitha Krishnan and Ambika devi Amma T, "Deception Detection in Social Media through Combined Verbal and Non-Verbal Behavior ", International Journal of Advanced Research in Computer Science and Software Engineering , Volume 5, Issue 4, 2015.

[7] Sheetal Antony, Prof. B. S. Umashankar, "Identity Deception Detection and Security in Social Medium, IJCSMC, Vol. 5, Issue 4, April 2016, pg.499-502.

[8] M Balaanand, RSoumipriya, S Sivaranjani and S Sankari, "Identifying Fake Users in Social Networks Using Non-Verbal Behavior". International Journal of Technology and Engineering System (IJTES)Vol 7. No.2 2015 Pp. 157-161©gopalax Journals, Singapore.

[9] Antu Mary Kuruvilla1 and Saira Varghese2, "A Survey on detecting Identity Deception in Social Media Applications", International Journal of Modern Trends in Engineering and Research (IJMTER) Volume 02, Issue 04, [April – 2015] ISSN (Online):2349–9745 ; ISSN (Print):2393-8161.

[10] Michail Tsikerdekiset. al., "Online Deception in Social Media", Information Science Faculty Publications Paper 12.2014.

[11] E.Elangovan1, Dr. D. Chandrakala," IDENTIFICATION AND PREVENTION OF MULTIPLE ACCOUNT IN SOCIAL MEDIA", International Journal of Advanced Technology in Engineering and Science, Volume No 03, Special Issue No. 01, March 2015.

[12] Ms. M. Preensta Ebenazer, Dr. P. Sumathi," An Overview of Identity Deception Approaches and Its Deception Approaches and Its Effects", International Journal of Computer Trends and Technology (IJCTT) – volume 25 Number 3 – July 2015.

[13]http://docsig.cis.uab.edu/media/2014/03/list-of-features.pdf.