Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 13, Issue 1, January 2022: 1306-1314

Identification of Harmful Social Bots using Learning Automata in Social Networks

Dr. D. Bujji Babu, Professor , Dept Of Master Of Computer Applications in QIS College Of Engineering & Technology from JNTUK.,(A.P),India .
Mr. Sk. Anjaneyulu Babu, Assistant Professor, Dept Of Master Of Computer Applications in QIS College Of Engineering & Technology from JNTUK,(A.P),India .
Mr. D. Ananda Rao, PG Scholar, Dept Of Master Of Computer Applications In QIS College Of Engineering & Technology from JNTUK,(A.P),India .
Mr. P. Suman, PG Scholar, Dept Of Master Of Computer Applications In QIS College Of Engineering & Technology from JNTUK,(A.P),India .
Mr. M. Lakshmi Babu, PG Scholar, Dept Of Master Of Computer Applications In QIS College Of Engineering & Technology from JNTUK,(A.P),India .
Mr. T. Kishore Kumar, PG Scholar ,Dept Of Master Of Computer Applications In QIS College Of Engineering & Technology from JNTUK,(A.P),India .

ABSTRACT :
New approaches to collecting and analysing such large data have emerged as the amount, speed, and variety of user data on online social networks (e.g., user-generated data) has increased dramatically. Social bots, for example, have been used to provide automated analytics services and to provide better customer service. Malicious social media bots, on the other hand, have been used to propagate false information (e.g., fake news) that has real-world implications. As a result, identifying and removing hazardous social bots from online social networks is critical. Examining the quantitative aspects of their activity is one of the most common strategies for identifying malicious social bots. Social bots can readily replicate these characteristics, resulting in lower analytical accuracy. This study presents a novel method for detecting hostile social bots that includes feature selection based on clickstream sequence transition probability as well as semi-supervised clustering. This strategy takes use of both the ephemeral nature of user behaviour and the possibility for clickstreams to emerge. The results of our experiments on real onlinesocial networking platforms show that the

identification method based on the probability of transformation of user behaviour clickstreams is more accurate than the method based on quantitative analysis of user behaviour in identifying different types of malicious social bots. It will rise by an average of 12.8 percent as a result of this.

Keywords*:* Social bots,Social network

I INTRODUCTION

On social media networks, a malicious Socialbot is a computer programme that impersonates a human user (OSNs). Furthermore, hostile social bots engage in a number of criminal actions, including social spam distribution, the creation of phoney identities, scamming online reviews, and phishing. When a Twitter user wishes to share a URL (s) with a nearby participant (i.e. followers or followers), the participant utilises the URL abbreviation (i.e. bit.ly) to abbreviate the URL (because a tweet is limited to 140 characters). Fraudulent social bots can also tweet phishing URLs in shortened form. Participants' requests are redirected to intermediate URLs connected to malicious servers when they click on the abbreviated phishing URL, as shown in Fig.

1. These intermediate URLs may redirect malicious URLs in tweets, detecting all dangerous social bots is challenging. As a result, it's critical to spot dangerous URLs (i.e. malicious URLs) shared on Twitter by bad social bots. Most present solutions depend on monitored learning algorithms, where the model is trained on labelled data, to identify hostile bots in OSNs. These methodologies, on the other hand, focus on statistical traits rather than analysing customer social behaviour. Furthermore, because hostile bots modify their behaviour over time to avoid being caught, these technologies are less efficient at identifying transient data patterns with noisy data (i.e., when the data is tainted by untrustworthy or phoney information). As a result, we proposed that ad hoc data models (such as the Learning Automaton (LA) model) be maintained using reinforcement learning methodologies.

II. LITERATURE SURVEY

Using Machine Learning to Detect Fake Accounts in Media Applications

The social network, which is such a vital aspect of our lives, is affected by online fraud and false accounts. Intruders frequently create phoney profiles on social media sites to carry out unwanted acts such as hurting people, stealing identities, and violating privacy (OSN). As a result, confirming whether the account is legitimate is the most important problem in OSN. In this study, we used the support vector machine approach and deep neural networks to introduce numerous classification methods. Taxonomic algorithms are also compared. The SpamUser dataset is utilised to choose the best one.

Using Variational AutoEncoder and k- Nearest Neighbor to Detect Social Media Bots

Harmful social media bots propagate malicious content on social media and have a substantial influence on network security. An accurate and reliable categorization of social media bots is critical for detecting information manipulation in social networks. To correct expensive labelling and unbalanced positive and negative sampling errors in existing social media bot detection methods and reduce training patterns to minimise abnormal patterns in the model, we

propose an extraordinary identification framework based on a combination of variant auto-encoder and anomaly detection algorithm. The purpose is to automatically encode and decode sample characteristics using a variation auto-encoder. The overall design elements are similar to the early aspects after decoding; however, the odd patterns and initial features are distinct. The original characteristics and decoding representation are combined, and abnormalities are detected using an unique detection approach. Studies on public datasets demonstrate that the suggested model's area below the curve for identifying social media bots has reached 98 percent, successfully differentiating bots from the general population and verifying the proposed model's usefulness.

Machine Learning Algorithms for Bot Detection on Twitter

Twitter is a prominent social media platform that allows users to express their opinions on a wide range of issues, including politics, sports, the stock market, and entertainment. This is one of the most efficient methods of data transport. It has the potential to significantly alter people's viewpoints. As a result, legal entities are compelled to use tweets rather than Twitter bots to communicate messages. The Twitter bot is used to send spam messages. As a result, detecting bots can aid in the detection of spam communications. This work uses machine learning to develop a method for detecting Twitter bots. Compare and contrast the Decision Tree, Multinomial Nave Base, Random Forest, and Bag of Words algorithms.

III SYSTEM ANALYSISEXISTING SYSTEM

Examining the quantitative features of their behaviour is one of the most frequent ways for spotting harmful social bots. Social bots can readily replicate these characteristics, resulting in lower analytical accuracy. To identify harmful social bots, a new approach has been presented, which involves both clickstream sequence transition probability and feature selection based on semi- supervised clustering.

PROPOSED SYSTEM

It is suggested that visual analysis be combined with software service requirements analysis, making it easier to analyse changes in client demands. Understanding the changing demands of the software services ecosystem need research like this.

We present a heuristic monitored booster model that uses the ratio of tweets delivered to tweets published on Twitter, average tweet length, URL, and forwarding time toraise the recall rate and identify dangerous bots.

A supervised machine learning method has been proposed to identify social bots based on the age, position, and other static characteristics of active, passive, and inactive Twitter users, as well as Twitter users based on interactions, interaction content, interactions, and certain dynamic characteristics. Concerning interpersonalsocial relationships.

Using our proposed semi-supervised clustering detection technique, we then evaluate and

categorise contextually conscious user behaviour on social networks. We can rapidly detect dangerous social bots by tagging a few users.

IV IMPLEMENTATION

Architecture:

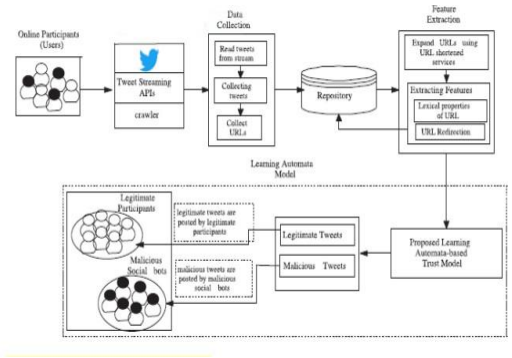

Fig-1. Architectures of the system model

MODULES:

Tweet Admin :

This module requires the administrator to log in with a valid username and password. He is capable of performing specific tasks, such as B. After the person has successfully signed in, view and authorise (provide link to the user to view the profile). This is where you'll find all of the uses for buddy requests and responses. Add a spam url name, then go through and block all spam accounts with profile information. View all information about people in order to unlock them using the decision tree or by clicking on their username. All of a user's tweet items, including interactions and ratings, may be found in one spot. Show all spam URL accounts (virus and malware oriented) as well as normal Forest Tree Reasons accounts. Show all spam and frequent URLs based on interactions by filter name, with a link to a graph that counts both users. Show all spam and generic URLs based on tweet information for each filter name, as well as a link to a chart with two user numbers. The amount of spam accounts and ordinary accounts may be seen in the graph.

Friend Request & Response :

This module allows the administrator to review all friend requests and answers. Here you'll see all of your requests and answers, as well as tags like ID, requested user picture, requested username, username request, status, and time and date. The status changes to authorised if the user approves the request; otherwise, the status remains pending.

User :

There are an unknown number of users in this module. Before engaging in any activity, the user must first register. A user's information is saved in a database when they register. He must log in with his approved username and password after finishing the registration procedure. After successfully checking in, the user may read their profile in the community, look for friends in the community, view friend requests and answers, and view my friends in the community, as well as explore TUses, tcontent desc, metadata desk, TweetURL, TDate and time, toner, and more. TImage Create a tweet item with the following fields: Tweet Post name, TAbout, TUses, and tcontent desc. Search for the tweet item using a term, then rank your interactions (the higher your score, the better) and look at the URL to view the webpage. All of your tweets, as well as other people's interactions and scores, are available in one spot. All of your friends' tweets, as well as other interactions and scores, may be viewed and rated. View all comparable friend tweets and show friend items with profiles of allspamming URLs.

Searching customers to make friends

In this module, the user searches for users on the same network or networks as them and sends friend invites to them. People can only create friends with users on other networks if they are approved.

V RESULT AND DISCUSSION

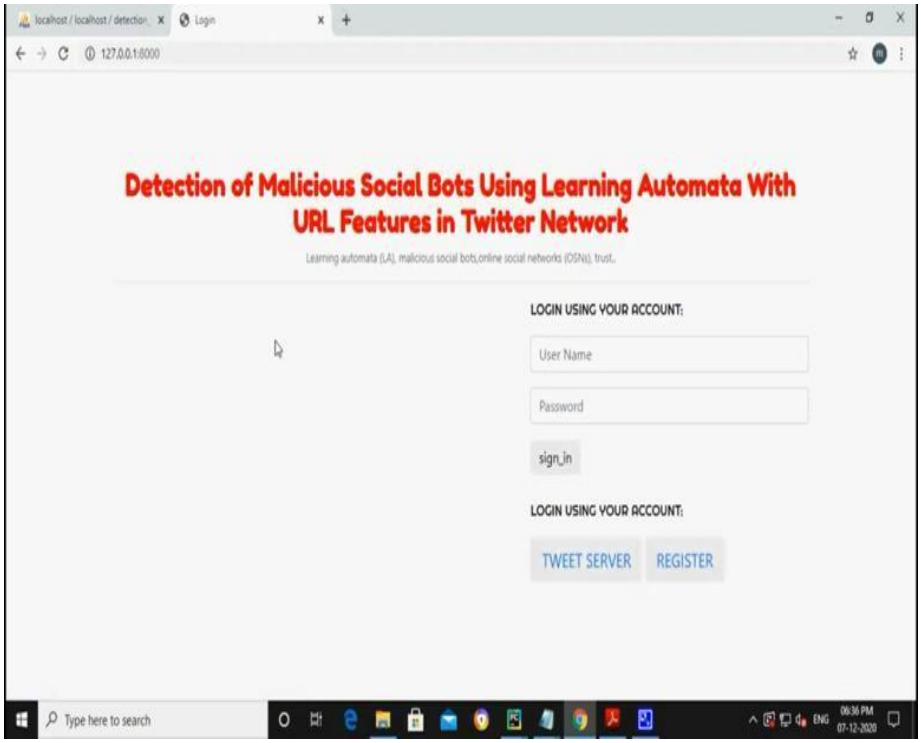

Fig 2 : User Registration

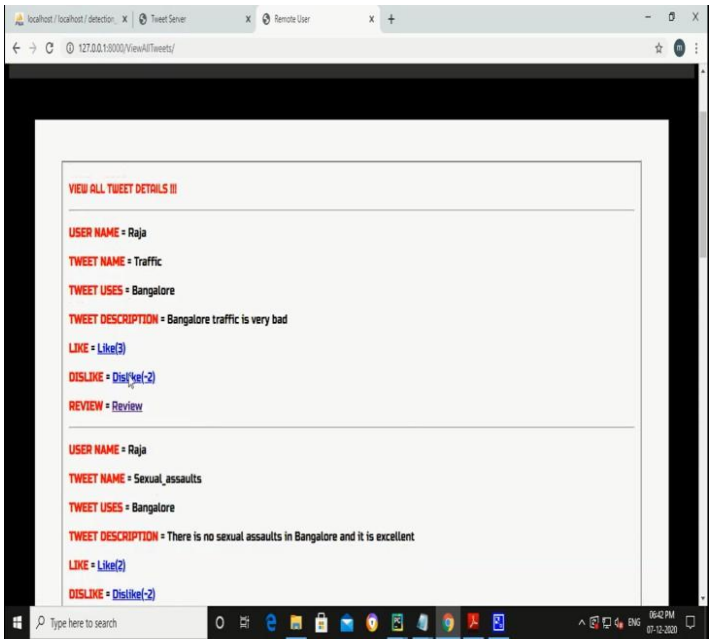

Fig 3: Tweet Post In Line Chart

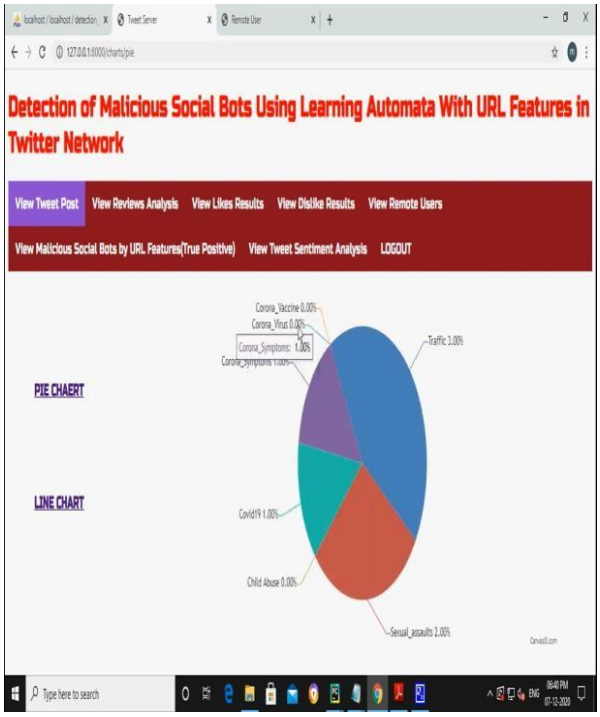

Fig 4 : Tweet Post In Pie Chart

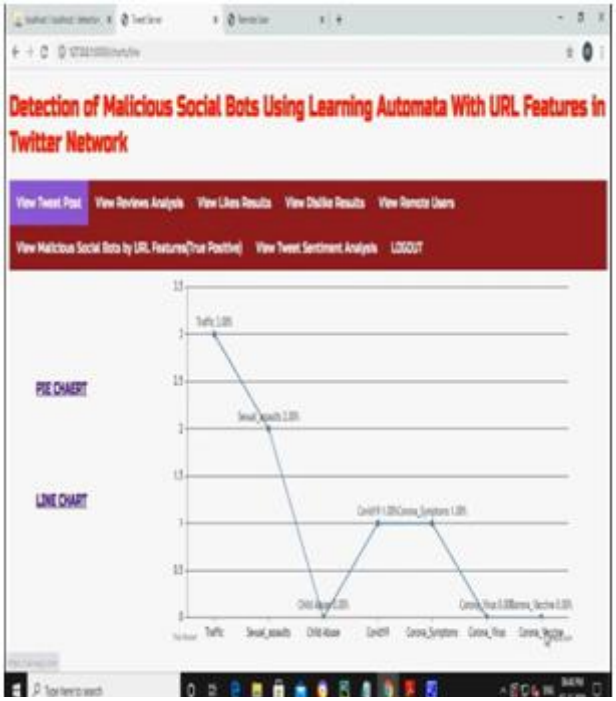

Fig 5: All Tweet Details

VI CONCLUSION

The LA-MSBD technique is presented in this study, which combines a dependable computational model with a set of URL- based MSBD characteristics. To check the legitimacy of tweets, we also apply Bayesian learning and DST (posted by each participant). In addition, the suggested LA- MSBD algorithm updates the action probability value with a minimal number of learning actions (i.e., the probability that participants will post malicious URLs in tweets). The suggested LA-MSBD method

improves learning efficiency. Two Twitter datasets will be used to test the performance of our proposed LA-MSBD algorithm. In comparison to other current algorithms, experimental findings show that the suggested LA-MSBD technique improves accuracy by up to 7%. The suggested LA- MSBD approach has an MSBD accuracy of
95.37 percent and a The Fake Project and Social Honeypot dataset accuracy of 91.77 percent. Furthermore, we wish to explore the interdependence of symptoms and their influence on MSBD as a future research issue.

VII REFERENCES

[1] P. Shi, Z. Zhang, and K.-K.-R. Choo, "Detecting harmful social bots based on clickstream sequences," IEEE Access, vol. 7, pp. 28855–28862, 2019. 2] P. Shi, Z. Zhang, and K.-K.-

R. Choo, "Detecting malicious social bots based on clickstream sequences," IEEE Access, vol. 7, pp. 28855–28862, 2019.

[2] G. Lingam, R. R. Rout, and D. V. L. N. Somayajulu, "Adaptive deep Q-learning model for identifying social bots and important people in online social networks," Appl. Intell., vol. 49, no. 11 (November 2019), pp. 3947–3964.

[3] T. T. Kwon, D. Choi, J. Han, S. Chun, E. Rappos, S. Robert, and D. Choi, "Bit.ly/practice: Uncovering content publishing and sharing using URL shortening services," Telematics Inform., vol. 35, no. 5, pp. 1310–1323, 2018.

[4] "Fluxing botnet command and control channels via URL shortening services," Comput. Commun., vol. 36, no. 3, pp. 320– 332, February 2013.

[5] S. Madisetty and M. S. Desarkar, "A neural network-based ensemble approach for spam detection in Twitter," IEEE Trans. Comput. Social Syst., vol. 5, no. 4, December 2018, pp. 973–984.

[6] H. B. Kazemian and S. Ahmed, "Comparisons of machine learning techniques for detecting malicious webpages," Expert Syst. Appl., vol. 42, no. 3, pp. 1166–1177, Feb. 2015. [7] H. B. Kazemian and S. Ahmed, "Comparisons of machine learning techniques for detecting malicious webpages," Expert Syst. Appl., vol. 42, no. 3, pp. 1166–11

[7] H. Gupta, M. S. Jamal, S. Madisetty, and M. S. Desarkar, "A framework for real-time spam identification in Twitter," in Proc. 10thInt. Conf. Commun. Syst. Netw. (COMSNETS), Jan. 2018, pp. 380–383, inProc. 10th Int. Conf. Commun. Syst. Netw. (COMSNETS), Jan.

[8] T. Wu, S. Liu, J. Zhang, and Y. Xiang, "Deep learning for Twitter spam detection," in Proc. Australas. Comput. Sci. WeekMulticonf. (ACSW), 2017, p. 3.

[9] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, "Key issues in guarding against harmful socialbots," in Y. Boshmaf, I. Muslukhov, K. Beznosov, and

M. Ripeanu, Y. Boshmaf, I. Muslukhov, K. Bez Large-Scale Exploits Emergent Threats, the 5th USENIX Workshop, 2012, pp. 1–4.

[10] G. Yan, "Peri-watchdog: Hunting for hidden botnets in the perimeter of online social networks," Computer Networks, vol. 57, no. 2, Feb. 2013, pp. 540–555.

[11] D. Canali, M. Cova, G. Vigna, and C. Kruegel, "Prophiler: A fast filter for the large-scale detection of malicious Web pages," in Proc. 20th Int. Conf. World Wide Web (WWWW), 2011, pp. 197–206; D. Canali, M. Cova, G. Vigna, and C. Kruegel, "Prophiler: A fast filter for the large-scale detection of malicious Web pages

[12] A. K. Jain and B. B. Gupta, "A machine learning-based approach for phishing detection using hyperlinks information," J. Ambient Intell. Hum. Comput., vol. 10, no. 5, May 2019, pp. 2015–2028.

[13] C. Chen, J. Zhang, X. Chen, Y. Xiang, and W. Zhou, "6 million spam tweets: A huge ground truth for timely Twitter spam detection," IEEE International Conference on Communications (ICC), Jun. 2015, pp. 7065–7070.

[14] "Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?" by Z. Chu,

S. Gianvecchio, H. Wang, and S. Jajodia. IEEE Transactions on Dependable Secure Computing, vol. 9, no. 6, November 2012, pp. 811–824.

[15] C. Chen, Y. Wang, J. Zhang, Y. Xiang,W. Zhou, and G. Min, "Statistical features- based real-time detection of drifting Twitter spam," IEEE Trans. Inf. Forensics Security, vol. 12, no. 4 (April 2017), pp. 914–925.

[16] N. Rndic and P. Laskov, "Practical evasion of a learning-based classifier: A case study," IEEE Symp. Secur. Privacy, pp.197–211, May 2014.