International Journal of Current Trends in Engineering & Technology
www.ijctet.org, ISSN: 2395-3152
Volume: 07, Issue: 04 (July-August, 2021)

# Social BOT Detection for Twitter Dataset by User Action and IWD Algorithm

**Kamlesh Kori [1], Rajesh Ku. Nigam [2], Dr. Bhupendra Verma [3]**

[1] Research Scholar, CSE Department., Technocrats Institute of Technology & Science, Bhopal (India) (kamleshkori45@gmail.com)
2.Associate Professor, CSE Department., Technocrats Institute of Technology, Bhopal (India) (rajeshrewa37@gmail.com)
3. Professor, CSE Department., Technocrats Institute of Technology, Bhopal (India) (bkverma3@gmail.com)

***Abstract*—** Digital platform dependency of today era attract promoters to brand product services. So unwanted posting was done by some programs known as a bot. Some researchers have proposed different techniques to identify these bots, which was a post by both programs. This paper has developed a model to identify bots from a real user. User action was analysed as features for the classification of bots and real user. The whole process adopts graph-based clustering and genetic algorithm based cluster representative feature. Graph-based clustering classifies the user into two classes, and the Intelligent water drop genetic algorithm finds the class representative action sequence in features. An experiment was done on a real Twitter dataset, and the result shows that the proposed model has increased the detection accuracy of work.

***Keywords****: Clustering, Data mining, Genetic Algorithm, Social Network.*

## I. INTRODUCTION

Social networks are very well-known networks through which data or thoughts of individual or community are exchanged across the globe. A social organisation is formed of nodes that are normally entities or associations. Individuals are communicating in Social Networks and developing
relationships with one another. In Social Networks, websites like Facebook, twitters, my webspace and LinkedIn are highly liked websites. Millions of clients are fascinated with these websites, and many have taken these websites as part of their living. For the past few years, the Social Networking websites like Facebook, Twitter, LinkedIn. have achieved so much recognition as it becomes the everyday routine of approximately every individual to check their profile every day as recognised by Michael Fire et al. [1]. Although it comprises many clients and is a hub of data, this has become a feasible path for attackers to use or assault. Many websites offer diverse things to prevent these sorts of assaults, but it is complicated to end them because they have various fresh methods each day to perform assault. Due to the user-friendly environment of Facebook, users are expected to reveal much private information about themselves and their links, as offered by Abu-Nimeh et al. [2]. The information may contain a date of birth, private pictures, service, email address, high school name, relationship status, and even mobile number. If a hateful user takes this private data, it is necessary to carry out malicious actions on their timeline or even in their private life [3]. For example, a hateful user can utilise the private data taken from the Facebook website to send customised spam posts. In Facebook, there are various third party requests accessed by the web user. When a user wants to drive any third-party request, the user must permit the authorisation to access some profiles information by the application. When the user permits the authorisation, the application can see the user's private information like name, email, and friends list. Occasionally hackers generate these applications and influence the user to utilise these hateful Apps. Customer accesses malicious Apps and has to share its private details with App. Hacker takes benefit of user's private details and posts hateful stuff on user's wall.

Among the diverse examination relating to Twitter, spam accounts recognition is mainly considered and applicable. In universal terms, spammers are beings, real users, or mechanical bots whose intent is to frequently distribute messages that include useless content for profitable functions [13], links to hateful websites to extend malware, phishing attacks, and other damaging action [5].

## II. Related Work

Daya L. Mevada [6] prescribed strategies to find opinion spam from the large measure of unstructured records has become critical exploration trouble. This examination prompts a sentiment spam analyser which over and again sorts input text data into either spam or non-spam gathering. The arranged framework applies artificial intelligence directed strategy.

M.N. Istiaq et al. [7] authors have presented an article that finds the chance of starting dynamic learning to recognise Review spams performed on genuine records that exhibit promising results. All through the methodology, they qualified model using dynamic taking in procedure which gains from the most phenomenal models in various emphases.

Rashmi Gomatesh et al. [9] anticipated their perception in the article "Recognition of Fake Review and Brand Spam Using Data Mining Technique". This strategy proposed a conducted way to deal with spot audit spammers who

attempt to control the evaluations on hardly anything. The author determines a joined activity method for grade analysts dependent fair and square to check spamming practices. They affirmed anticipated methods by performing client estimation on an Amazon dataset that holds surveys of various organisation's things.

S P. Rajamohana et al. [10] focused on effectively misdirecting surveys open on the web, which slowly influences organisations and customers. Along these lines, it is critical to notice and expel such bogus feeds from online locales. This record uncovers a few methodologies used for audit spam acknowledgement, and execution measures were perceived.

Mubarak et al. [11] introduced an effortless method for understanding the hypothesis. People may like to channel information for various reasons, such as ordering information, killing obscene substance from the media stream, or preventing kids from seeing unambiguous posted messages. Every one of these targets manual for component learning interchanges with the Twitter API and different limits. A further inside and out an assessment of spamming hurts reveals building calculations; for example, NB IBK (which may allude to Ibk calculation, applies the k-NN calculation) to find answers for the trouble.

Ameen and Kaya [12] proposed a related work and found that easygoing backwoods had the greatest accomplishment at 92.95%. An examiner must research to discover the best calculation to use before going with further examination. There is no fastidious calculation that goes past all others under all conditions; this explains the need to explore various methodologies. Before moving towards higher classifier techniques, it is important to value why most analysts have released SVM classifiers, such as a sack of words and a pack of implies.

Alshehri et al. [13] use hashtags and N-grams to show out grown-up Arabic substance. The pack of words procedure uses twofold qualities to guarantee positive words in a posted substance, while a sack of implies includes discovering a normal of word vectors. The result of their inspection was a 79% precision of preparation.

## III. Proposed Methodology

This section proposed bot detection model was explained where working steps were explained by the block diagram shown in figure 1. Each block detailed below subsection.

### Pre-processing

Preprocessing is a procedure utilised for the transformation of content into a feature vector. Much the same as content orders, the preprocessing additionally have a debate about its division. This work uses tweets preprocessing, which comprise words in charge of bringing down the execution of learning models. The dataset contains some of the columns that are not fruitful for this work; hence, information is either deleted or transformed into different states. Information preprocessing diminishes the extent of the info content records fundamentally. It includes exercises like sentence limit assurance, common language, particular stop word disposal and stemming.

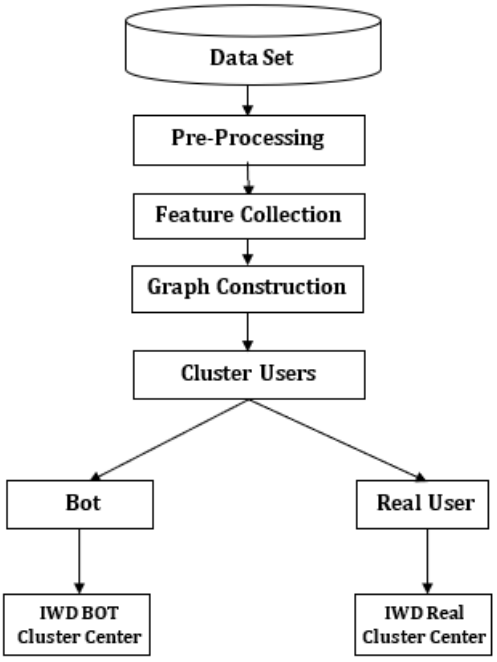


Figure 1 IWDBD Block diagram.

### Featured collection:

This work twitter dataset was considered as the input where nine features of each user were extracted. These features F represent the user behaviour on the social network. Table 1 shows the feature set utilise in this work. Feature

1. Sequence of Shares
2. The sequence of Likes instance
3. The sequence of Tweets perform by the user
4. The sequence of re-tweet perform by the user
5. Time instance of the event

Feature from 1 to 5 can be easily extracted from the dataset where the user follows the behaviour. Transition probability between action events steams: P(i, j) represents the probability that the click action is j at timestamp t followed by click action at timestamp t-1. So, probability finds the relation between the i and j features of the use in the form of transition done in a specific set of durations.

$$P(i,j) = \frac{\sum_1^t X_i \rightarrow X_j}{\sum_1^t X_i}$$

In above equation $X_i \rightarrow X_j$act as a transition from i to j feature instance, while t as the total time instance when i feature were applied. So, if work uses n number of

features, each user has a feature vector of nxn where cell contain a probability transition value.

$$F_x = \begin{vmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{vmatrix}$$

**Graph Construction**

In this step, develop a completely connected graph where each node is connected with another node, and the distance between them acts as the edge's weight. Estimation of the distance was done by using the X and Y-axis of the system. Here this can be understanding as let nodes are N = {$n_1$, $n_2$, $n_3$…….$n_m$} and distance between them are evaluated by Euclidian distance formula. Now sort graph edges with Minimum Weight in decreasing order. This can be understood as matrix S[] of three columns and rows depending on the number of edges present in the graph.

**Resultant cluster**

So, nodes present with less distance edge weights are considered the true user or real social media user. At the same time, nodes whose distance values are larger in the outer cluster were considered the bots. As each bot set of instance sequence was different from the real user set, distance from other existing nodes was high. Hence, cluster selection of real or bot depends on the partial tree's weight value present in the cluster.

**IWD Algorithm**

The IWD drop moves towards other drops and forms a group as per low soil between the two merging drops [14]. In this paper, the drop is user, and soil is graph weight values. So, the soil values between each drop were summarised in a graph. As a graph has a node, so unique user counts are present as graph nodes. While the distance between each node from other node act as weight or soil in the graph.

$$Soil(x, y) = \text{Euclidian}(x, y) \text{ --Eq. 1}$$

Where x, y is the nodes in the graph and Soil(x, y) is the distance between users.

**Static and Dynamic Parameter**

In this step some of constant were initialize before the start of algorithm such as soil updating parameters $S_1$ = 1, $S_2$ = .01, and $S_3$ = 1, velocity updating parameters $V_1$ = 1, $V_2$ = .01, and $V_3$ = 1. Finally, global and local soil constants $\beta_L$ and $\beta_G$ are initialised by 0.9. Values of constants may be varying as per algorithm requirement.

**Population Generation**

A random set of users were collected as a cluster centre candidate. The collection of these candidates were termed as population. Each candidate has two user value set, which is also known as chromosome [15]. So, if population PW has n number of candidates, then Eq. 2 gives the population set for processed image PI.

PW←Population_Generation(PI, n)----Eq.2

**Drop Movement Probability**

The Association of a drop towards another drop depends on movement probability. So as per soil weight value obtained from

$$WS(i,j) = \begin{cases} Soil(i,j) & \text{if min (Soil(i, all nodes)} > 0 \\ & \text{Otherwise} \\ Soil(i,j) - \min(Soil(i, \text{all nodes}) & \end{cases}$$

$$FS(i,j) = \frac{1}{\delta + GS(i,k)} \text{-----Eq. 3}$$

$$DMP(i,j) = \frac{FS(i,j)}{\sum_{k=1}^{N} FS(i,k)} \text{-----Eq. 4}$$

Movement Probability MP was evaluated by eq. 3 and 4. FS is a feasible value solution as per soil.

**Update Drop and Soil Values**

Drop movement change velocity of drop as per soil and n merging drop velocity. So, eq. 5 gives velocity update value for $t^{th}$ iteration.

$$DV(t+1) = V(t) + \frac{V_1}{V_2 + V_3 * Soil(D,D'')^2} \text{-----Eq. 5}$$

Similarly, soil value was updated by Eq. 6.

$$\Delta S(D, D'') = \frac{s_1}{S_2 + S_3 * T(t+1)^2} \text{--------Eq. 6}$$

$$T(t+1) = \frac{h}{V(t+1)}$$

$$Soil(i,j) = (1 - \beta_L) * Soil(i,j) - \beta_L * \Delta S(i,j)$$

Where $v_1$, $v_2$, h, $s_1$, $s_2$, $s_3$, $\beta_L$ are constant range between 0 to 1.

**Fitness Function**

Finding the difference between user's Fitness function similarity was used to evaluate the similarity between the profile, content, and self-mention. Sort the similarity matrix in descending order to assign the user to the centroid. Each feature gives its separate index to the genetic algorithm population, so different weight assign for each feature.

**IWD Crossover**

Some groups of chromosomes were prepared, and a good solution was considered for the local parent, which crossover with other chromosomes in the group [16]. One random position pixel value was copied from the local parent chromosome in a crossover and replaced the same

position other chromosome pixel value of the same group. The obtained new chromosome was further evaluated to get its fitness value; if the new chromosome fitness value was better than the group chromosome, it replaces one of the poor parents in the population; otherwise, the parent will continue.

**Final Solution**

After a sufficient number of iteration, cluster centres are obtained and assign users to those clusters. Here each cluster is representing by its cluster centre. So as per the different number of user type available in the dataset number of clusters are generated.

## IV. Experimental Setup

The whole work was implemented on MATLAB software. It is utilised on account of its rich library, which has numerous inbuilt storage that can be specifically used in this work for various reason. Out of various storage, few are crossing point, contrasting of the string, and so forth. One more essential factor is its GUI by which one who doesn't know about the code can straightforwardly run the storage without having earlier information.

**Dataset**

This work experiment is done on social dataset content obtained from https://botometer.iuni.iu.edu/bot-repository/datasets.html, whereas per the user-related Twitter comments of respected user with different action and timestamp were available. The description of the dataset is available in table 1.

Table 1 Experimental dataset detailed description.

| Features | Values |
|---|---|
| Instances | 50000 |
| Users | 35 |
| Data Year | 1 |
| User Actions | Tweet, Share, Re-tweet, Like |

**Results**

The proposed model results were compared with previous work, Click Stream Sequence BOT Detection (CSSBD), proposed in [17].

Table 2 Precision Based Comparison

| Data Size | IWDBD | CSSBD |
|---|---|---|
| 15000 | 1 | 0.25 |
| 20000 | 0.7778 | 0.5556 |
| 25000 | 0.7857 | 0.7857 |
| 30000 | 0.9444 | 0.7778 |
| 35000 | 0.7826 | 0.7696 |

Table 2 and figure 2 have shown that the proposed model has increased the classification accuracy of work. Values show that the use of graph clustering algorithm has reduced the confusion of user identity. The further paper has involved the IWD algorithm, which gives user features as per class.

Table 3 Recall Based Comparison

| Data Size | IWDBD | CSSBD |
|---|---|---|
| 15000 | 1 | 1 |
| 20000 | 1 | 0.7143 |
| 25000 | 1 | 0.6471 |
| 30000 | 0.7391 | 0.7 |
| 35000 | 0.9474 | 0.7692 |

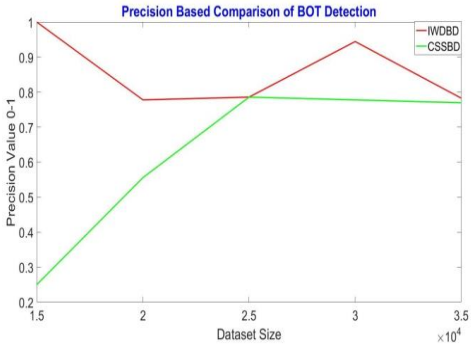


Figure 2 precision value of Bot detection at different dataset size.

Above table 3 has shown that the proposed model has increased the classification accuracy of work. Values show that the use of graph clustering algorithm has reduced the confusion of user identity. The further paper has involved the IWD algorithm, which gives user features as per class.

Table 4 F-measure Based Comparison.

| Data Size | IWDBD | CSSBD |
|---|---|---|
| 15000 | 1 | 0.4 |
| 20000 | 0.875 | 0.6250 |
| 25000 | 0.88 | 0.7097 |
| 30000 | 0.8293 | 0.7368 |
| 35000 | 0.8571 | 0.8163 |

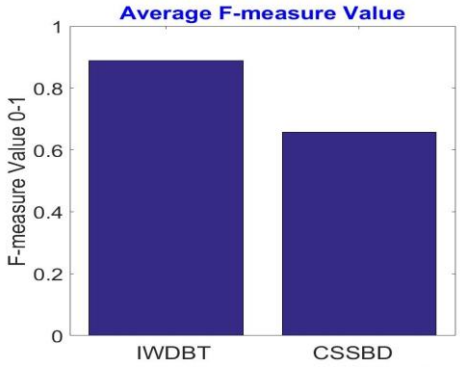


Figure 3 Average f-measure value of Bot detection techniques.

Table 4 and figure 3 have shown that the proposed model has increased the bot detection accuracy of work. Feature selection has increased the graph clustering algorithm accuracy. The further paper has involved the IWD algorithm, which gives user features as per class.

Table 5 Execution time Based Comparison.

| Data Size | IWDBD | CSSBD |
|---|---|---|
| 15000 | 1.52 | 1.92 |
| 20000 | 1.54 | 2.23 |
| 25000 | 2.23 | 2.844 |
| 30000 | 2.46 | 2.945 |
| 35000 | 3.37 | 4.011 |

In table 5 shows that the proposed model has required less time to find the real user class.IWD algorithm has increased the detection ratio.

Table 6 Accuracy Based Comparison.

| Data Size | IWDBD | CSSBD |
|---|---|---|
| 15000 | 1 | 0.25 |
| 20000 | 0.8182 | 0.4545 |
| 25000 | 0.85 | 0.55 |
| 30000 | 0.7083 | 0.5833 |
| 35000 | 0.7931 | 0.6897 |

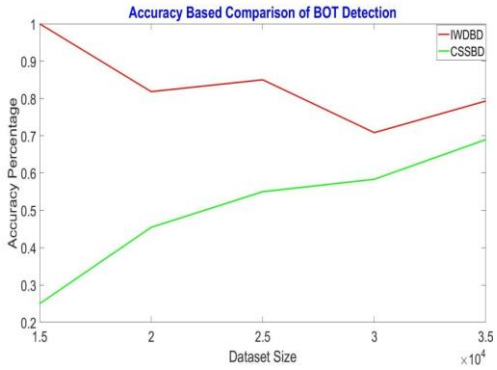


Figure 4 Accuracy value of Bot detection at different dataset size.

Table 6 and figure 4 have shown that the proposed model has increased the bot detection accuracy of work. Feature selection has increased the graph clustering algorithm accuracy. The further paper has involved the IWD algorithm, which gives user features as per class.

## V. Conclusions

The life of social media depends on real user action, but digital user performs unfair action and reduces overall trust value. Many social sites execute bot detection algorithm. This paper has proposed a graph-based IWD algorithm for bot detection. The input of this algorithm was a set of user action and based on these action transition probability features, and the user was cluster into two class. The output of the graph-based clustering algorithm was passed into the genetic algorithm. IWD algorithm finds the bot user feature set and real user feature set. The experiment was performed on a Twitter real dataset, and results show that the proposed model has increased the bot detection accuracy by 39.38%. In future, the researcher can perform bot detection by using the artificial intelligence method.

## References

1. Morris, M. and Ogan, C. (1996). The internet a mass medium. Journal of communication, 46(1):39-50.
2. Lee, K., Eo_, B. D., and Caverlee, J. (2011). Seven Months with the Devils: A Long-Term Study of Content Polluters on Twitter. In Proc. AAAI Intl. Conf. on Web and Social Media (ICWSM).
3. Ferrara, E., Varol, O., Davis, C., Menczer, F., and Flammini, A. (2016a). The rise of social bots. Communications of the ACM, 59(7):96{104.
4. Jun, Y., Meng, R., and Johar, G. V. (2017). Perceived social presence reduces fact-checking—proceedings of the National Academy of Sciences,114(23):5976{5981.
5. Jagatic, T. N., Johnson, N. A., Jakobsson, M., and Menczer, F. (2007). Social phishing. Communications of the ACM, 50(10):94{100.
6. Mevada D. L., Daxini V., "An opinion spam analyser for product Reviews using the supervised machine learning method." pp.03, (2015).
7. M. N. Istiaq Ahsan , Tamzid Nahian , Abdullah All Kafi , Md. Ismail Hossain, Faisal Muhammad Shah "Review Spam Detection using Active Learning." 978-1-5090-0996-1, pp.16, (2016).
8. Michael C., et al. "Survey of review spam detection using machine learning techniques." Journal of Big Data 2.1, pp.9, (2015).
9. Adike R. G., Reddy V, "Detection of Fake Review and Brand Spam Using Data Mining Technique.", pp.02,(2016).
10. Rajamohana S. P, Umamaheswari K., Dharani M., Vedackshya R., "Survey of review spam detection using machine learning techniques.",978-1-50905778-8, pp.17 (2017).
11. Mubarak, H.; Darwish, K.; Magdy, W. Abusive language detection on Arabic social media. In Proceedings of the First Workshop on Abusive Language Online, Vancouver, BC, Canada, 4–7 August 2017; pp. 52–56.
12. Ameen, A.K.; Kaya, B. Detecting spammers in a Twitter network. Int. J. Appl. Math. Electron. Comput. 2017, 5, 71–75.
13. Alshehri, A.; Nagoudi, A.; Hassan, A.; Abdul-Mageed, M. Think before your click: Data and models for adult content in Arabic Twitter. In Proceedings of the 2nd Text Analytics for Cybersecurity and Online Safety (TA-COS-2018),

2018.

14. J. Benesty, J. Chen and Y. Huang, "Study of the widely linear Wiener filter for noise reduction," *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, Dallas, TX, 20104.
15. B. Z. Dadaneh, H. Y. Markid and A. Zakerolhosseini, "Graph colouring using Intelligent Water Drop algorithm," *2015 23rd Iranian Conference on Electrical Engineering*, Tehran, 20155.
16. Basem O. Alijla, Chee Peng Lim, Li-Pei Wong, Ahamad Tajudin Khader, Mohammed Azmi Al-Betar. "An ensemble of intelligent water drop algorithm for feature selection optimisation problem". Applied Soft Computing, Volume 65, 2018.
17. Peining Shi, Zhiyong Zhang And Kim-Kwang Raymond Choo. “Detecting Malicious Social Bots Based on clickstream Sequences”. IEEE Access March 18, 2019.