International Journal of Scientific Research & Engineering Trends
Volume 7, Issue 1, Jan-Feb-2021, ISSN (Online): 2395-566X

Segregating Spammers from Social Digital Platform by Genetic Algorithm

Rahul Kumar, Dr. Avinash Sharma
Millennium Institue Of Science And Technology
Bhopal, Madhya pradesh, India
rahulmits57@gmail.com

***Abstract*- Presence of online Social digital platform in human life from past few decade influence advertiser to place their product in different style. Some of digital marketing companies plays a giant role for promoting a product, service, thought, etc. to social digital users. Digital platform do not want to lose their control so spammer who work professionally to promote a brand need to identify ad remove from the platform. This paper has work in this field of online social spammer detection by developing a unsupervised algorithm which identify the user class (Spammer or real) as per sequential behavior done. Proposed work utilized a intelligent water drop genetic algorithm for segregating the users into desired class. Some of digital platform inherent features of the user profile were used for finding the fitness class of the user to the corresponding class of user. Experimental work was done on real twitter dataset and comparison of proposed model was done with existing method of spammer detection. Results shows that proposed model has increased the accuracy of spammer detection.**

***Keywords*:- Digital social media, Genetic algorithm, Spammer, Malicious Post.**

I.INTRODUCTION

By using the Internet, it has become quite unpretentious to receive any type of information from any source around the world. The increased demand from social sites allows users to gather an abundance of user information and data. Enormous amounts of data on these pages often draw the attention of fake users [1]. Twitter has rapidly become an online source for acquiring real-time information about users. Twitter is an Online Social Network (OSN) where users can share anything and everything, such as news, opinions, and even their moods. Several arguments can be held over different topics, such as politics, current affairs, and important events.

When a user tweets something, it is instantly conveyed to his/her followers, allowing them to outspread the received information at a much broader level [2]. With the evolution of OSNs, the need to study and analyze users' behaviors in online social platforms has intensity. Many people who do not have much information regarding the OSNs can easily be tricked by the fraudsters.

There is also a demand to combat and place a control on the people who use OSNs only for advertisements and thus spam other people's accounts. Recently, the detection of spam in social networking sites attracted the attention of researchers. Spam detection is a difficult task in maintaining the security of social networks. It is essential to recognize spams in the OSN sites to save users from various kinds of malicious attacks and to preserve their security and privacy. These hazardous maneuvers adopted by spammers cause massive destruction of the community in the real world. Twitter spammers have various objectives, such as spreading invalid information, fake news, rumors, and spontaneous messages. Spammers achieve their malicious objectives through advertisements and several other means where they support different mailing lists and subsequently dispatch spam messages randomly to broadcast their interests. These activities cause disturbance to the. original users who are known as non-spammers. In addition, it also decreases the repute of the OSN platforms.

Several research works have been carried out in the domain of Twitter spam detection. To encompass the existing state-of the- art, a few surveys have also been carried out on fake user identification from Twitter. Tingmin et al. [4] provide a survey of new methods and techniques to identify Twitter spam detection. The above survey presents a comparative study of the current approaches.

On the other hand, the authors in [5] conducted a survey on different behaviors exhibited by spammers on Twitter social network. The study also provides a literature review that recognizes the existence of spammers on Twitter social network. Despite all the existing studies, there is still a gap in the existing literature.

Therefore, to bridge the gap, we review state-of-the-art in the spammer detection and fake user identi_cation on Twitter. Moreover, this survey presents a taxonomy of the Twitter spam detection approaches and attempts to offer a detailed description of recent developments in the domain.

© 2021 IJSRET

II. RELATED WORK

Ahmed & Abulaish in 2012 [6] work for determining the spam profiles over OSNs. A dataset comprising of Facebook profiles (both real and spam profiles) was utilized for the experiment. A weighed graph was applied which helped in modeling the social networks in which the profiles were characterized as nodes and their intersections as edges. Also, in case of a cluster comprising both real and spam profiles, the idea of majority voting was established. The outcomes have revealed that the introduction of majority voting not only reduces the sum of clusters but also increases the performance values of FP and FB measures.

Al-garadi, et.al, 2016 [7] A set of characteristic attributes from Twitter, concerning network, activity, user, as well as tweet content were utilized in developing a supervised machine learning solution for recognizing cyberbullying on Twitter. Resulted in an fmeasure of 0.936 with an area under the receiver-operating characteristic curve of 0.943. These outcomes suggested that the model based on these features offered a reasonable solution for identifying Cyberbullying in online communication backgrounds. Lastly, the outcome acquired with the proposed features was proven the best by comparing with results from two baseline features for their performance.

Liu, et.al, 2017 [8] suggested a fuzzy-based oversampling method that offered artificial data samples. These samples were obtained from limited observed samples depending on their idea of classification using fuzzy based information. Moreover, an ensemble technique was developed which examined certain classifiers based on their imbalanced data in three phases. It could be concluded from the study that by means of imbalanced class distribution, the rate of spam detection can greatly be influenced.

Muhammad U.S. Khan et. al. 2017 [9] proposed a framework that separates the spammers and unsolicited bloggers from the genuine experts of a specific domain. The proposed approach employs modified Hyperlink Induced Topic Search (HITS) to separate the unsolicited bloggers from the experts on Twitter on the basis of tweets. The approach considers domain specific keywords in the tweets and several tweet characteristics to identify the unsolicited bloggers.

Rathore, et.al, in 2018 works accordance with recent activities of Facebook, the outline presented the urgency of setting out a new feature set for detecting spams. A baseline dataset was used from Facebook which involved 300 spammers in addition to 700 legitimate user profiles. The baseline dataset encompassed feature set for an individual profile, which were extracted through a novel dataset construction mechanism. Furthermore, an intelligent decision support scheme was employed to differentiate spammers from genuine users. The evaluation outcomes proved that the suggested outline was precise and effective to deliver excellent performance with a maximum accuracy of 0.984 and Mathew correlation coefficient of 0.977.

III. PROPOSED METHODOLOGY

Explanation of proposed model SDIWDA (Spammer Detection by Intelligent Water Drop Algorithm) was done in this section by flow chart of figure 1.

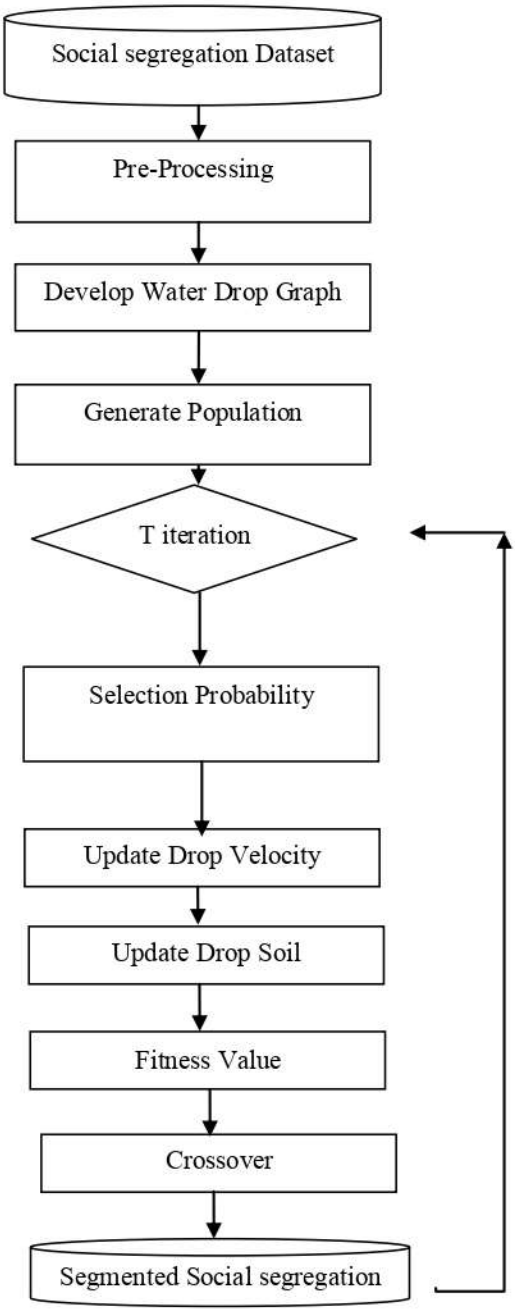

Fig.1 Block diagram of proposed IWD based Spammer detection.

© 2021 IJSRET

1. Pre-Processing:
As the dataset is a collection of data which is unorganized and need to retrieve important information which is fruitful for the work in this work dataset contain time, date, protocol, session, etc. Here data is clean and transform this as per working environment. Preprocessing is a procedure utilized for transformation of content into feature vector. As tweet content on webpage have words which need pre-processing by removing stopwords [14]. So set of stopwords are removed and filtered words were further process to collect keywords.

Hence each tweet has its own set of keywords depend on type of content. Although common keywords may exist between users of same domain. So let tweet Tm have content {w1, w2, s1, w3, s2, s1…….wn} where n is total number of words in Tm page. After stopword {s1, s2,…..} removal important words will be {w1, w2, w3,…….wn} [12].

2. Feature Collection:
In this work twitter dataset was consider as the input where nine features of each user were extract. These features F represent the user behavior on the social network. Some of features extract from twitter dataset are used in the work: Has tag, URLS used by user, Average URLs in a tweet, Average URLs in a tweet, Average hashtags in a tweet, Words in Tweet, Final tweets, Single user Tweets/Total, Number of Tweet, Average words in tweet of single user, User's tweets that contain the URLs.

3. IWD (Intelligent Water Drop):
In this IWDGA each user act as a water drop where one drop move towards other drop where soli between the drop position is less. So soil act as distance between the user value. This movement of user or water drop creates a segment in the social social segregation. Output of this module was segment representative user values.

4. Develop Water Graph:
First phase of this intelligent water drop genetic algorithm was to develop a graph where each node is user and other distance between them act as weight of graph. Soil of the graph act as graph weight which was estimate by Eq. 4.

$Soil(i,j) = \text{Euclidian}(i,j)$ --Eq. 1
Where i, j are user value in a social segregation and S is soil matrix of users.

5. 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 β_L and β_G are initialize by 0.9. Values of constants may be vary as per algorithm requirement.

6. Generate Population:
Possible set of solutions which are terms as chromosome in this algorithm are generate in this step [14]. Each water drop is set of possible cluster center set. Hence water drop is combination of H={U_1, U_m}, where population have total n number of water drops. Hence population generation function in this algorithm is shown by Eq. 2.

H←Generate_population(m, n)----Eq. 2

7. Drop Movement Probability:
As social digital users was segment into s class, so each representative user move toward other user of social segregation as per soil value. So movement probability of water drop help in this selection by Eq. 3.

$$DMP(i,j) = \frac{FS(i,j)}{\sum_{k=1}^{N} FS(i,k)} \text{ -----Eq. 3}$$

$$FS(i,j) = \frac{1}{\delta + GS(i,k)} \text{-----Eq. 4}$$

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

Where WS is Weighted Soil, FS is Feasible Solution

8. Modify Drop Velocity and Soil Values:
D^{th} Drop velocity was update by Eq. 5 when it move towards D'^{th} drop.

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

D^{th} Drop soil was update by Eq.6 when it move towards D'^{th} drop.

$$\Delta S(D,D') = \frac{S_1}{S_2 + S_3 * T(t+1)^2} \text{ --------Eq. 6}$$

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

HD is heuristic durability a constant value range in 0-1.

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

Where v_1, v_2, s_1, s_2, s_3, β_L are constant range between 0 to 1.

9. Fitness Function:
Fitness value of chromosome depends on the segment user representative set. So an summation of minimum distance between the segment representative users to the segmenting user of social segregation act as fitness value. Chromosome having minimum summation value is consider as best solution in the current population.

10. Crossover:
Genetic algorithm works on random sampling of solution. To increase the sampling reach chromosome need to be regenerate or mdify the existing population. Hence crossover operation was adopt by the work. One of best chromosome in the population modify other chromosome cluster center value.

© 2021 IJSRET

11. Spammer and Real User cluster:
So water drop which have best fitness value after sufficient number of iteration is consider as resultant cluster. Hence as per cluster center user were identified as the true user or spammer of the social media. As each spammer set of instance sequence were totally different from real user set of instances, so distance from other existing nodes were high.

IV. EVALUATION PARAMETERS

1. Precision:

$$\text{Precision} = \left(\frac{\text{True}_{\text{positive}}}{(\text{False}_{\text{positive}} + \text{True}_{\text{positive}})}\right)$$

2. Recall:

$$\text{Recall} = \left(\frac{\text{True}_{\text{positive}}}{\text{False}_{\text{negative}} + \text{True}_{\text{positive}}}\right)$$

3. F-Measure:

$$\text{F} - \text{Measure} = \left(\frac{2\text{xPrecisionxRecall}}{(\text{Recall} + \text{Precision})}\right)$$

4. Accuracy:

$$\text{Accuracy} = \left(\frac{\text{Correct_class}}{(\text{Correct_class} + \text{In Correct_class})}\right)$$

5. Dataset:
Experiment was perform on twitter dataset having 30 users with 1055 tweets. Work has segregate users into two class first was spammer and other was real users. Comparison of proposed model was done with Hyperlink Induced Topic Search algorithm proposed in [13].

V. RESULTS

Table 1 Precision value based comparison of users segregating models.

User Set	SDIWDA	Hyperlink Induced Topic Search
Set1	0.85	0.461538
Set2	0.88	0.5
Set3	0.84	0.555556
Set4	0.85	0.5385
Set5	0.86	0.6667

Precision values obtained from various spammer detection techniques, shown in table 1. This improvement was achieved by involving pre-processing filter. Use of genetic algorithm increase the feature selection for blogger and spammer classification work.

Table 2. Recall value based comparison of users segregating models.

User Set	SDIWDA	Hyperlink Induced Topic Search
Set1	1	0.857143
Set2	0.933333	0.727273
Set3	0.9375	0.833333
Set4	0.944444	0.6364
Set5	0.9	0.7143

Recall values observed in table 2 shows that proposed SDIWDA algorithm was better as compared to other spammer detection algorithm Hyperlink Induced Topic Search. As per selected features graph weight values of edges has increase the recall value of SDIWDA. As SDIWDA has utilized content similarity between social users spammer identification accuracy was increase.

Table 3. F-measure value based comparison of users segregating models.

User Set	SDIWDA	Hyperlink Induced Topic Search
Set1	0.917	0.6
Set2	0.903	0.592593
Set3	0.88	0.666667
Set4	0.895	0.5833
Set5	0.88	0.6897

F-measure values obtained from various spammer detection techniques, shown in table 1. This improvement was achieved by involving pre-processing filter. Use of genetic algorithm increase the feature selection for blogger and spammer classification work.

Accuracy values observed in table 4 shows that proposed SDIWDA algorithm was better as compared to other spammer detection algorithm Hyperlink Induced Topic Search. As per selected features graph weight values of edges has increase the accuracy value of SDIWDA. As SDIWDA has utilized content similarity between social users spammer identification accuracy was increase.

Table 4. Accuracy value based comparison of users segregating models.

User Set	SDIWDA	Hyperlink Induced Topic Search
Set1	0.875	0.5
Set2	0.85	0.45
Set3	0.8183	0.5456
Set4	0.831	0.5833
Set5	0.81	0.6667

VI. CONCLUSIONS

Social media sites engage people by various means of social activities and public interest advisement. This gathering of people attract advertiser to increase their product reach with filtered audience. As social media want

© 2021 IJSRET

to maintain real users on platform so detection of spammers are proposed by this paper. This work has utilize soft computing technique, Intelligent water drop genetic algorithm for feature based user segregation. Users were segment into two classes first was spammer and other was real user. Propsoed model do not need prior training, leaning, etc. as genetic algorithm IWD, dynamically segregate users. Expeirment was done on twitter dataset and result shows that proposed has increase the precision value by 34.38% while accuracy of correct class identification was improved by 36.39%. Infuture researcher can adopt other language user to segregate.

REFERENCES

[1] Van Dijck, José. The culture of connectivity: A critical history of social media. Oxford University Press, 2013.
[2] Boyd, D., & Ellison, N. (2008). Social network sites: Definition, history, and scholarship. Journal of Computer Mediated Communication, 13(1), 210—23.
[3] Lenhart, Amanda. "Teens, social media & technology overview 2015." Pew Research Center 9 (2015).
[4] Gauri Jain* , 2Manisha, 3Basant Agarwal. "An Overview of RNN and CNN Techniques for Spam Detection in Social Media". Volume 6, Issue 10, October 2016 ISSN: 2277 128X.
[5] I. Drost and T. Scheˌer. Thwarting the nigritude ultramarine: Learning to identify link spam. In Proceedig of the 16th European Conference on Machine Learning, ECML'05, 2005.
[6] Ahmed, F., & Abulaish, M. (2012, June), "An mcl-based approach for spam profile detection in online social networks". In Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on (pp. 602-608), IEEE.Al-garadi, M.
[7] A. Varathan, K. D., & Ravana, S. D.(2016), "Cybercrime detection in online communications: The experimental case of cyber bullying detection in the Twitter network", Computers in Human Behavior, 63, 433-443.
[8] Liu, S., Wang, Y., Zhang, J., Chen, C., & Xiang, Y. (2017), "Addressing the class imbalance problem in Twitter spam detection using ensemble learning", Computers & Security, 69, 35-49.
[9] Muhammad U. S. Khan, Mazhar Ali, Assad Abbas, Samee U. Khan, *and* Albert Y. Zomaya. "Segregating Spammers and Unsolicited Bloggers from Genuine Experts on Twitter". IEEE Computer Society, 2017.
[10] Rathore, S., Loia, V., & Park, J. H. (2018), "SpamSpotter: an efficient spammer detection framework based on intelligent decision support system on facebook", Applied Soft Computing, 67, 920-932.
[11] F. Masood et al., "Spammer Detection and Fake User Identification on Social Networks," in IEEE Access, vol. 7, pp. 68140-68152, 2019.
[12] R. Abinaya, B. Niveda E. and P. Naveen, "Spam Detection On Social Media Platforms," 2020 7th International Conference on Smart Structures and Systems (ICSSS), Chennai, India, 2020.https://humanities data.com/resources/157
[13] Atta, Rand & Hassan, Soukaena & Hashem, & Ekhlas, Khalaf & Gbashi, Ekhlas. (2019). Spam Classification Using Genetic Algorithm. Iraqi Journal of Information Technology. V.9 N.2. 2018.

© 2021 IJSRET