Biogeography Genetic Algorithm Based Social Platform Spammer Identification Using Content Feature

Rupali Vishwakarma[1], Dr. Pratima Gautam[2]
[1]Phd Scholar Department of CSE, [2]Department of CSE
AISECT University, Bhopal, India
rupalidohare25@gmail.com, pratima_shkl@yahoo.com

***Abstract* —** *The most popular and leading social network service have the probability of threats and unwanted posts. So to identify and block such Spams, there are a few techniques were developed. Number of researchers have proposed different techniques to identify malicious accounts and spammers over last two decades. This work has also proposed an un-supervised technique for identifying the real users from the social network spammers. Here clustering of social users were done by using tweleve features on the basis of words, hastags, number of blogs (Tweet), URL, etc. Here for unsupervised spammer identification genetic algorithm biogeographic genetic algorithm was used. As this algorithm perform chromosome modification on the basis of immigration and emigration rate, so reaching a good solution is easily achieved. Proposed model cluster the user on the basis of its social activities in articular duration of time. Experiment was done on real dataset from twitter social network. Proposed algorithm BGOA (Biogeographic Optimization Algorithm) for spammer detection in social network was compared with other existing algorithm on different evaluation parameters and results shows that proposed model was better than other.*

Index Terms — *Online Social Networks (OSNs), Twitter, Spammers, Legitimate users*

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 organization 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 web space and LinkedIn are highly liked websites. Millions of clients, are fascinated with these websites and many of them have taken these websites as part of their living. From the past some years, the Social Networking websites like Facebook, twitter. Linkedin etc. have achieved so much recognition as it becomes the everyday routine of approximately every individual to check their profile every day as recognized by Michael Fire et al. [1]. Although it comprises a vast number of clients and it a hub of data, this has become a feasible path for attackers to use or assault. Many websites offers diverse things to prevent these sorts of assaults but it is complicated to end them because they have a variety of fresh methods each day to performing assault. Due to the user friendly environment of Facebook, users are expected to reveal many private information about themselves and their links as offered by Abu-Nimeh et al. [2]. The information may contain date of birth, private pictures, place of service, email address, high school name, relationship status, and even mobile number. If this private data is taken by hateful user then it is to them to carry out malicious actions on their timeline or even in their private life [3]. For example, a hateful user can utilize the private data taken from the Facebook website to send customized spam posts to user. In Facebook there are various third party requests accessed by the web user. When user wants to drive any third party request then user must permit the authorization to access the some profiles information by the application. When user permits the authorization then application can see the user's private information like name, email id and friends list etc. Occasionally hackers generate these applications and influence the user to utilize 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.

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

II. Related Work

In 2015, Daya L. Mevada, [6] recommended techniques to locate opinion spam from enormous amount of unstructured records has become an significant research difficulty. This research advises an opinion spam

analyzer which repeatedly categorize input text information into either spam or non-spam group. The planned system will apply machine learning supervised method.

In 2016, M.N. Istiaq Ahsan et. al. [7] This article discovers the chances of initiating active learning for identifying Review spams performed on real life records which demonstrates promising outcomes. Throughout the procedure, they qualified model utilizing active learning technique which learns from the most excellent examples in numerous iterations.

In 2016, Miss. Rashmi Gomatesh Adike et. Al. [9] projected their observation in the article "Detection of Fake Review and Brand Spam Using Data Mining Technique". This method suggested a behavioral approach to spot review spammers those who are trying to control the ratings on few items. Writer derives a combined action techniques for grade reviewers based on the level that they have verified the spamming behaviors. They confirmed projected techniques by performing user estimation on an Amazon dataset which holds reviews of diverse company's items.

In 2017, SP.Rajamohana, et. al. [10] Focused light on misleading reviews that are easily accessible in the internet which gradually more affects businesses and clients. Therefore it is significant to notice and remove such false feeds from online sites. This document discloses some approaches utilized for review spam recognition and performance measures were recognized.

Mubarak et al. [11] presented a graceful means of understanding the theory. Individuals may prefer to filter data for numerous reasons, such as the want of classifying data, eliminate pornographic substance from the media stream, or stop kids from seeing unambiguous posted messages. All these objectives guide to mechanism learning communications with the Twitter API and other boundaries. A further in-depth examination of spamming harms discloses engineering algorithms such as NB IBK (which is may refer to Ibk algorithm, applies the k-NN algorithm) as means of discovering solutions to the difficulty.

Ameen and Kaya [12] proposed out a related work and found that casual forest had the maximum success at 92.95%. An investigator must research to find out the greatest algorithm to use before going with further analysis. There is no meticulous algorithm that goes beyond all others under all conditions; this clarifies the want of research with different approaches. Before moving towards higher classifier methods, it is necessary to appreciate the cause that the majority of researchers have discharge SVM classifiers such as bag-of-words and bag-of-means.

Alshehri et al. [13] utilize hashtags and N-grams to display out adult Arabic substance. The bag-of-words technique utilizes binary values to ensure for definite words in a posted content, while bag-of-means involve finding out an average of word vectors. The outcome of their examine was a 79% accurateness of processing.

III. Proposed Methodology

Explanation of proposed model SDBOA (Spammer Detection by Biogeography Optimization Algorithm) was done in this section by flow chart of figure 1. Whole work was broadly perform in two section first was developing a ontology by using web content from social network set of actions taken by user. Second was Testing where input are user feature set. So output of second model is predicted class (Spammer user or Real user) of digital social network user.

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].

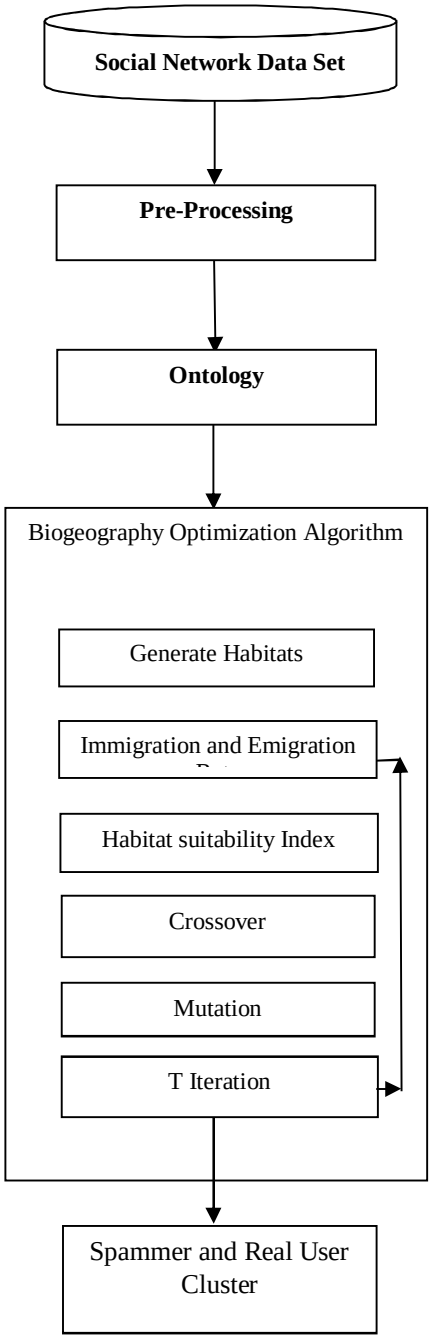

Fig.1 Block diagram of proposed BGOA based Spammer detection.

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. Table 1 shows feature set utilize in this work.

Table 1 Feature used for spammer identification.

1	Total number of Has tag
2	Total number of URLS used by user
3	Average number of URLs in a tweet
4	Average number of URLs/words in a tweet
5	Average number of hashtags/words in a tweet
6	Total number of words
7	Total tweets
8	Single user Tweets/Total Number of Tweet
9	Average number of words in tweet of single user
10	User's tweets that contain the URLs/Tweet
11	Average number of HAstags in a tweet
12	User Inter content relation

Total number of Has tag: Its an summation of HasTag used by the single user in all number of tweets for particular time duration.

Total number of URLS used by user: Its an summation of URL used by the single user in all number of tweets for particular time duration.

Average number of URLs in a tweet: Its an ratio of total URL used by the single user in all tweet to the total number of tweets done in particular time duration.

Average number of URLs/words in a tweet: Its an ratio of total URL used by the single user in all tweet to the total number of words used by single user tweets done for particular time duration.

Average number of hashtags/words in a tweet: Its an ratio of total HasTags used by the single user in all tweet to the total number of words used by single user tweets done for particular time duration.

Total number of words: Its an summation of words used by the single user in all number of tweets for particular time duration.

Total tweets: Its an total number of tweets done by the single user in particular time duration.

Average number of HAstags in a tweet: Its an ratio of total hastags used by the single user in all tweet to the total number of tweets done in particular time duration.

Single user Tweet/Total Number of Tweet: Its an ratio of total tweets done by the single user to the total number of tweets done in particular time duration by all user under observation.

User Inter content relation: Its an similarity between user on the basis of common number of words for all number of tweets [15].

Average number of words in tweet of single user: Its an ratio of total words used by the single user in all tweet to the total number of words used in tweets done for particular time duration.

User's tweets that contain the URLs/Tweet: Its an count of total number of tweet having URL to the total number of tweets done by single user in particular time duration.

BOA (Biogeography Optimization Algorithm)

Species in nature adopt changes as per suitable environmental conditions. So change of habit is one of type of change adopt by species time to time. Based on this MacArthur and Wilson [16] proposed an mathematical algorithm in early 1960, where main concern of this model to understand the migration of species from one habitat to other, Biogeography was a trending research area at that time. So in 2008 Simon [17] proposed an generalized genetic algorithm to resolve similar type of issues. Some of basic terms related to this work are:

Habitat Suitability Index (HSI): This is term as fitness value of the habitat, means higher value shows that poor place to live while low value means good place to live in terms of resources, life, etc.

Immigration and Emigration Rate Some of basic terms of immigration λ and Emigration α was done by Eq. 1, 2 [16, 20, 21]:

$$\lambda_R = (1 - \alpha_R)\text{----------Eq. 1}$$

$$\alpha_R = \frac{R}{h}\text{----------Eq. 2}$$

Where R is rank of habitat in terms of HSI value, while h is total number of habitats.

Generate Habitats: Possible set of solutions which are terms as habitat in this algorithm are generate in this step. Each habitat is set of possible cluster center set. Hence habitat is combination of H={U_1, U_m}, where population have total h number of habitats. Hence population generation function in this algorithm is shown by Eq. 3.

H←Habitat(m, h)----Eq. 3

Fitness Function (HSI): Habitat suitability Index of any habitat depends on the distance. Estimation of the distance was done by using F_X and F_Y user features. Here this can be evaluate by Euclidian distance formula.

$$DD^{f}_{x,y} = \sqrt{(XX_{f1} - YX_{f1})^2 + (XX_{f2} - YX_{f2})^2 \dots \dots + (XX_{fn} - YX_{fn})^2}\text{ ----Eq. 4}$$

Where X and Y are feature values of ending nodes, while XX_{f1} is the X user feature value 1. Hence rank i of the habitat depend on the summation of distance $F_u$value of each habitat from other user.

HSI = Rank(F_u, H)----Eq. 5

Crossover

Emigration of user in form of species from one habitat to other is depend on emigration rate. While permitting species to enter in a habitat is depend on immigration rate. Hence for crossover from one habitat to other both type of rate need to find. So crossover depends on following condition.

```
Loop x=1:h
    If Cross_Over_Limit > λR
        Loop y=1:h
            If Cross_Over_Limit > αR
                M←Rand()
                H[x, m]← H[y, m]
            EndIf
        EndLoop
    EndIf
EndLoop
```

Where Cross_Over_Limit is random number range between 0-1, x and y is habitat position specify immigration, emigration operation.

Mutation

In this work after crossover mutation was also perform so chance of new solution get increases. For this paper has involved mutation probability where as per HSI value mutation was performs in selected habitats.

$$M_R = \frac{HSI_R}{sum(h)}\text{----------Eq. 6}$$

$$M_p = \frac{M_R}{Max(M_R)}\text{ ----------Eq. 7}$$

Hence habitat which cross a constant mutation_cross_limit range in 0-1, M_R gives an mutation rank for the habitat as per HSI value. So higher value have higher mutation rank. Hence those habitat which have higher mutation rank have higher mutation probability. So habitat which have lower

Mutation Probability as compared to Mutation-Cross_Threshold undergoes to mutation.

Spammer and Real User cluster

So habitat 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.

Proposed Algorithm
Input: DS // DS: Dataset
Ouput: C_r C_b// C_r: Cluster of real user, C_b : Cluster of social Spammer

1. PD←Pre-Processing(DS)
2. Loop 1:n// n: number of users
3. Loop 1:i // i, j: User Features
4. F(i,j)←Feature(PD,n,i)
5. EndLoop
6. EndLoop
7. H←Habitat(m, h)
8. Loop 1:iteration
9. HSI = Rank(F_u, H) // Fitness Function
10. H←Crossover(λ_R, α_R, HIS, H)
11. H←Mutation(HIS, H)
12. EndLoop
13. HSI = Rank(F_u, H)
14. Loop 1:n
15. If $DD_{1,n}$(HIS) > $DD_{m,n}$(HIS)
16. C_r←n
17. Otherwise
18. C_b←n
19. EndIf
20. EndLoop

Above algorithm takes tweets of users as input and gives an cluster of spammer, real user. Hence tweets post by spammer can be removed easily as spammer is detected. Pre-processing steps has reduce the features extraction time by utilizing most of features in numeric type. This reduces the execution time of work as well, as working with numeric data is quit easy.

IV. Evaluation Parameters

Precision: Precision value is the ratio of predicted positive user to the totalpredicted user.

$$Precision = \left(\frac{True_{positive}}{(False_{positive} + True_{positive})}\right)$$

Recall: The recall is the fraction of relevant users that have been predicted over the total amount of input users. It is also known as Sensitivity or Completeness.

$$Recall = \left(\frac{True_{positive}}{False_{negative} + True_{positive}}\right)$$

F-Measure: Harmonic mean of precision value and recall value is F-measure.

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

Accuracy: This act as the percentage of correct prediction from the total set of prediction.

$$Accuracy = \left(\frac{Correct\_class}{(Correct\_class + InCorrect\_class)}\right)$$

Dataset
In this work twitter dataset named as twitter_cikm_2010 [] was used which have three column first was USER-ID, Twitter-ID and Tweet. For testing 1057 tweets were used for the classification of user into Real or Fake class.

Results
Results of the proposed work **BGOA, was compared with GBSD** (Graph Based Spammer Detection), FCMRF (Fuzzy C Mean Random Forest) [18] and HITS [19].

Table 2 Precision value comparison of **GBSD** and HITS work.

User Set	BGOA	GBSD	FCMRF	HITS
14	0.818182	0.818182	0.545455	0.454545
16	0.846154	0.923077	0.615385	0.461538
20	0.875	0.875	0.5	0.5
22	0.833333	0.833333	0.666667	0.555556
24	0.85	0.75	0.45	0.55
26	0.857143	0.714286	0.619048	0.571429

Above table 2 shows that Precision value of proposed BGOA was high as compared to previous algorithm GBSD, FCMRF and HITS. Here proper leaning feature with pre-processing filter increase the efficiency of the work. It has been observed that proposed work content feature selection plays an important role for unsupervised classification of data into blogger and spammers.

Table 3 Recall value comparison of proposed and previous work.

User Set	BGOA	GBSD	FCMRF	HITS
14	1	0.75	0.75	0.833333
16	1	0.8	0.8	0.857143
20	0.933333	0.823529	0.8	0.727273
22	0.9375	0.882353	0.923077	0.833333
24	0.944444	0.882353	0.818182	0.846154
26	0.9	0.882353	0.866667	0.857143

Above table 3 shows that Recall value of BGOA was high as compared to previous algorithm GBSD, FCMRF and HITS. Here proper weight assignment of edges as per features values gives better result. It was obtained that user content relation building increase the efficiency of work as it directly identify the similarity between user content.

Table 4 F-Measure value comparison of proposed and previous work.

User Set	BGOA	GBSD	FCMRF	HITS
14	0.9	0.782609	0.631579	0.588235
16	0.916667	0.857143	0.695652	0.6
20	0.903226	0.848485	0.615385	0.592593
22	0.882353	0.857143	0.774194	0.666667
24	0.894737	0.810811	0.580645	0.666667
26	0.878049	0.789474	0.722222	0.685714

Above table 4 shows that F-Measure value of proposed BGOA was high as compared to previous algorithm GBSD, FCMRF and HITS.. Here proper leaning feature with pre-processing filter increase the efficiency of the work. It has been observed that proposed work content feature selection plays an important role for unsupervised classification of data into blogger and spammers.

Table 5 Accuracy value comparison of proposed and previous work.

User Set	BGOA	GBSD	FCMRF	HITS
14	0.857143	0.642857	0.5	0.5
16	0.875	0.75	0.5625	0.5
20	0.85	0.75	0.5	0.45
22	0.818182	0.772727	0.681818	0.545455
24	0.833333	0.708333	0.458333	0.541667
26	0.807692	0.692308	0.615385	0.576923

Above table 5 shows that Precision value of proposed BGOA was high as compared to previous algorithm GBSD, FCMRF and HITS.. Here proper weight assignment of edges as per features values gives better result. It was obtained that user content relation building increase the efficiency of work as it directly identify the similarity between user content.

V. Conclusions

Social network is place to connect and share thoughts with each other. But most of people get attract from the social audience gathering for there personal or professional advantages. In the propose work no need of any configuration for the information, for example, speakers recognizable proof image or exceptional character. This work presents a study of methods for detection of user profiles as real or social spammer. Here a biogeographic optimization algorithm was proposed for classifying the social nodes into two cluster, where digital social network features were collect from the social action perform by the user. Use of these features were done by fitness function which have increase the work performance as well. Results shows that proposed BGOA based clustering technique has increase the precision value as compared to previous approach HITS by 39.11%. While recall value was also increase by 13.31%, at the same time accuracy of the social spammer identification was also increase. By 38.22% In future researcher can adopt other genetic algorithm with different feature sets to increase detection percentage.

References

[1] Morris, M. and Ogan, C. (1996). "*The internet as 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 re-duces 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 analyzer for product Reviews using 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 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] Boshmaf, Y., Muslukhov, I., Beznosov, K., and Ripeanu, M. (2012). "*Key chal- lenges in defending against malicious socialbots*". In Proc. 5th USENIX Conference on Large-Scale Exploits and Emergent Threats (LEET).

[15] Dandan Jiang1, Xiangfeng Luo1, Junyu Xuan, And Zheng Xu *."Sentiment Computing for the News Event Based. on the Social Media Big Data". Digital Object Identifier 10.1109/ACCESS.2016.2607218 IEEE Acess 2017.*

[16] MacArthur R., Wilson E. "*The Theory of Biogeography"*. Princeton, NJ, USA: Princeton University Press; 1967.

[17] Simon D. Biogeography-based optimization. IEEE Transactions on Evolutionary Computation. 2008;12(6):702–713. doi: 10.1109/tevc.2008.919004.

[18] Rupali Vishwakarma, Dr. Pratima Gautam . "*Un-Supervised Random Forest Tree and Content Feature-Based Blog Spammer Identification*". International Journal of Computer Sciences and Engineering (ISSN: 2347-2693), Vol.7, Issue.9, September 2019.

[19] Muhammad U. S. Khan, *Member,* Mazhar Ali, *Member,* Assad Abbas, *Student Member,* Samee U. Khan, *Senior Member and* Albert Y. Zomaya. "*Segregating Spammers and Unsolicited Bloggers from Genuine Experts on Twitter*". IEEE Computer Society 2016.

[20] Mohammed Alweshah. "*Construction biogeography-based optimization algorithm for solving classification problems*". Neural Computing and Applications, Springer volume 28 February 2018

[21] Raju Pal, Mukesh Saraswat. "*Enhanced Bag of Features Using AlexNet and Improved Biogeography-Based Optimization for Histopathological Image Analysis*". 2018 Eleventh International Conference on Contemporary Computing (IC3)