International Journal on Emerging Technologies **11**(5): 205-208(2020)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

Predicting Attackers of Cyberbullying using Ant Colony Optimization in Combating Psychological Effects among the Victims

Rizauddin Saian[1], Noor Shafiqa Fazlien Mohamad Fauzi[2], Zeti Zuryani Mohd Zakuan[3], Haswira Nor Mohamad Hashim[4] and Anida Mahmood[5]

[1]Senior Lecturer, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Cawangan Perlis, Malaysia.
[2]Student, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Cawangan Perlis, Malaysia.
[3]Senior Lecturer, Faculty of Law, Universiti Teknologi MARA, Cawangan Perlis, Malaysia.
[4]Associate Professor, Faculty of Law, Universiti Teknologi MARA, Malaysia.
[5]Senior Lecturer, Faculty of Law, Universiti Teknologi MARA, Malaysia.

(Corresponding author: Rizauddin Saian)
(Received 24 June 2020, Revised 28 July 2020, Accepted 17 August 2020)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Cyberbullying involves the action of attackers, which publicly embarrasses the victim by sharing some personal information of the victim. The data is shared publicly via social media. This is because the social media platform makes it easier to share something that can shame others in the real world. This study aims to develop a classification model to predict attackers of cyberbullying using Ant Colony Optimization. Once the attackers' tendency can be predicted, the Act of attacking can be prevented at an early stage. As a result, the study managed to develop a model to predict cyberbullying which can prevent innocent people from becoming a victim. Indirectly, it will decrease the rate of physiological effect due to cyberbullying. The results have shown that the developed classification model produced better predictive accuracy than J48. The accuracy of the model for J48 is 66.88%, while the accuracy of Ant-Miner is 69.69%. In future, the government authority can develop a mobile app using this model to combat cyberbullying.

Keywords: ant colony optimization, classification, data mining, cyberbullying, psychological effect.

I. INTRODUCTION

Cyberbullying is an act of bullying committed using technological devices [1]. Cyber Security Malaysia revealed that cyberbullying cases involving students happen every day, with 1524 cases reported from 2012 to 2016 [2]. Cyberbullying occurs when the attackers attack the victims by posting their values and believes when any breach of societal expectations and norms are committed.

Cyberbullying is meant to harass the victim by showing something personal publicly. Sometimes it can be retribution for a certain action. Doxing is an example of cyberbullying whereby the attackers publish sensitive information about the targeted individuals with the idea of embarrassing them. The action leads to putting pressure on the victims.

On the other hand, bullying or shaming others via the internet is also considered a social control [3]. It is being used as a sanction to punish the victim for their actions, which is perceived as a violation of a social norm. However, this phenomenon affects the victims emotionally and can sometimes lead to depression. Thus, this kind of action should stop and let the law punish the victim.

Cyberbullying is found to affect the victims psychologically [4]. The typical impact of cyberbullying includes frustration, anger, and sadness [5]. Victims of cyberbullying are largely related to a high level of depressive affect, and the victims also prone to suicidality [6]. Apart from depression, cyberbullying also may lead to anxiety, low self-esteem, insecurity, and loneliness [7].

The problem of cyberbullying in Malaysia can be seen in a case involving a 14-year-old schoolgirl from SMK Methodist Nibong Tebal. She was accused of stealing her teacher's iPhone. As a result, she was cyberbullied by her schoolmates. She denied the accusation. However, being a young schoolgirl, she could not handle the situation. As a result, she committed suicide [8]. Another case of cyberbully was the case of the famous Malaysian skateboarding, Fatin Syahirah Roszizi. She did not perform well during the 2018 Asian Games in Jakarta. As a result, she was roughly criticized and bullied on social media. Due to this, she became depressed, deleted all her photos on her Instagram, and shun herself from society [9].

Cyberbullying in Malaysia is provided under a few laws. Defamation Act 1957 states that if a person by his words spoken or written defame others' reputation, they can be sued under the Act. In order for the victim to sue, he must prove that "the words are defamatory, the words refer to the victim and the words have been published". The victim will succeed if all the three elements are established. The victims also have the right under the Communication and Multimedia Act 1998. Section 233 Communication and Multimedia Act 1998 provides that if the attackers used "any network, service or application" to bully the victim, the attacker shall be guilty of an offence. If found guilty, the attackers may be fine or imprisonment or both. The victim of cyberbullying can also rely on Sedition Act 1948 since the conduct of cyberbullying does fall within the ambit of seditious tendency under section 3 of Sedition Act 1948. Any person who is guilty of the offence shall be, imprisonment or both. The laws are enacted to protect the victims and at the same time act as a deterrence to the attackers. However, in reality, both victims and attackers are not aware of the existing laws. On the other hand, predicting attackers' tendency for cyberbullying can also protect the victims. Successful prediction of attacker's tendency can prevent them from committing the Act of bully. As a result, psychological effects on the victims can definitely be prevented.

The study developed a classification model by using ACO Algorithm. The study enables those with a cyberbully tendency to be predicted and corrected at an early stage. Doing so can reduce the rate of cyberbullying, hence preventing innocent people from becoming victims of a cyberbully. As a result, it will combat the psychological effects among victims of a cyberbully.

II. METHODS

Dorigo, Maniezzo, and Colorni introduced the Ant Colony Optimization (ACO) in 1996 [10]. The behavior of ants inspires ACO. Ants live in colonies. Their job is to find food for their colonies. Ants need to move from their colony to food source as much as possible to collect food. While moving, ants will deposit pheromone on the paths. Other ants can smell the pheromone and will follow the same route to the food source. This is how ants communicate effectively with each other to get to the food source without fail [11].

Table 1: Data description.

Attributes	Values
Gender	Male, Female
Age	17-20, 21-25, 26-30, 31-35, 36-40, 41-45, 46-50,
Marital Status	Single, Married, Divorced
Academic Program	Bachelor's degree, Advance Diploma, master's degree, Doctorate
Student Accommodation	Student hostel, rented house, stay with parents/relatives, Stay in your own house
Permanent address	Traditional Village House, Bungalow unit, Linked house/semi-detached, Apartment/condominium
Social Bond	Close, Loose, Not bonding, No family/friend
Mode of Interaction	Face to face, Online, rarely internet, Living in isolation
Amount of hour using internet in a day	20-24, 15-19, 10-14, less than 10
Have you been shamed online?	Yes, No
Have you ever like/share/comment online posting that shame other people?	Yes, No

In 2002, Ant-Miner was introduced [12]. Ant-Miner is an ACO for rule induction. It requires each ant to begin with an empty rule. Each term is generated one at a time, and it is added to the rule.

Despite the effectiveness of Ant-Miner, it cannot handle continuous data. Hence, discretization can be performed before the rule construction process [13]. For example, at the pre-processing step, use the C4.5-disc method for discretization [14, 15].

This study consists of several processes. The first process is the data pre-processing. The study used data on the respondents' demographic and behavioural characteristics collected from a survey carried out in Universiti Teknologi MARA, Shah Alam, Malaysia. Table 1 shows the description data of the study. The data has been converted into a nominal form.

The second process is model development. The model development process produced the classification model for predicting attackers of cyberbullying using Ant-Miner. The classification model consists of a set of rules. There are three steps required by Ant-Miner to produce a rule. The first step is computing the entropy of each terms Eqn. 1

$$H(W|A_i = V_{ij}) = -\sum_{c=1}^{k} \frac{n_c}{N} \log_2 \frac{n_c}{N} \qquad (1)$$

Where n_c is the number of example in c^{th} decision, and $N = \sum_{c=1}^{k} n^c$ is the number of example in dataset.

The next step is to calculate the heuristic of each terms Eqn. 2

$$n_{ij} = \frac{\log_2 k - H(W|A_i = V_{ij})}{\sum_{i=1}^{a} x_i . \sum_{j=1}^{b_i} (\log_2 k - H(W|A_i = V_{ij}))} \qquad (2)$$

Where *a* is the total number of attributes, b_i is set to 1 if the attribute A_i is not yet selected, else it is set to 0. For this case, value of *k* will be two because there are two decisions. The value of denominator is calculated earlier to make the calculation of heuristic easier.

Finally, compute the probability of each terms. The probability (Equation 3) depends on the pheromone values and also the heuristics values of each terms.

$$P_{ij} = \frac{n_{ij} . \tau_{ij}(t)}{\sum_{i=1}^{a} x_i . \sum_{j=1}^{b_i} \tau_{ij}(t)} \qquad (3)$$

Where n_{ij} is the value of a problem-dependent heuristic function for term *ij*; $\tau_{ij}(t)$ is the amount of pheromone currently available at any time t in the position *i*, *j* of the path that being followed by the ant. *a* is the number of attributes and b_i is the total number of values in the domain of attribute *i*.

After the three steps are completed, one term is found and added to the current partial rule. These steps are repeated until all attributes have been chosen. Finally, the algorithm will select a class for the rule, and the steps is repeated to generate the next rule.

The final process is model validation. In this process, the classification model is tested and validated to identify its performance and accuracy. The k-fold technique is used for validation, which is expected to achieve the maximum performance of the classifier. The 10-fold cross validation method is used in this study to calculate the predictive accuracy Eqn. 4

$$Accuracy = \left(\frac{forecast - actual}{actual}\right) * 100\% \qquad (4)$$

Later, the predictive accuracy of the classification model is compared with the result from J48 [16], an implementation of C4.5 [17] in WEKA [18]. The best result depends on the highest accuracy in a classification model.

III. RESULTS AND DISCUSSION

There are three different outputs in Ant-Miner: predictive accuracy, number of the rules, and the number of conditions for each rule.
The number of ants used was chosen randomly, 200, 400, 600, 800, 1000, and 1200. These different numbers of ants were used with a constant value of other parameters. The maximum uncovered cases are 10, rules of convergence is 10, and the number of iterations is set to 10. For cross validation, k is set to ten as suggested [19].
Table 2 depicted a different number of ants of the Ant-Miner algorithm predictive accuracy between 67.81% to 69.69% interval when a different value of ants was used. Prediction accuracy for the first 200 number of ants is 68.75%. Then, the value of accuracy was still the same for the next 400 number of ants. The accuracy value decreased for 800 number of ants, which is 67.81%, but at 1000 the number of ants, the graph increased to 69.38%. So, the highest accuracy for the minimum case per rule 1200 is 69.69%. The lowest minimum case per rule is 67.81%. Hence, 69.69% is the best percentage of accuracy.
Table 3 shows the average number of rules for a different number of ants. The highest number is 7.1 for 200 number of ants. Then, the lowest rule number is 6.6 for 800 and 1200 the number of ants. Thus, 6.6 rules number is the best choice because it has the lowest rule number. The lowest numbers of rules are 6.6 for 800 and 1200 number of ants, as shown in Table 3.
Table 4 depicted the average number of conditions for a different number of ants. It is found that the changes in the number of ants do not affect the number of conditions. In average, the number of conditions is 10.9.

Table 2: Predictive accuracy for different number of ants.

Number of Ants	Accuracy (%)
200	68.75
400	68.75
600	69.38
800	67.81
1000	69.38
1200	69.69

Table 3: Number of rules for different number of ants.

Number of Ants	Rule Number
200	7.1
400	6.8
600	6.8
800	6.6
1000	6.8
1200	6.6

Table 4: Number of conditions for different number of ants.

Number of Ants	Number of Conditions
200	11.5
400	10.7
600	11.0
800	10.3
1000	11.0
1200	10.9

In summary, the changes in the number of ants will produce a different value of predictive accuracy for the rule. The minimum number of rules produced by algorithm is 6.6 at 800, and 1200 number of ants and the minimum condition for each of the rules is 10.3 at 800 number of ants. In this study, the lowest rules number and condition numbers were chosen for this purpose. The best predictive accuracy for Ant-Miner is 69.69% when the number of ants is 1200.
WEKA is used in this study to develop a classification model using the J48 algorithm. The accuracy of the model for J48 is 66.88%. Hence, the Ant-Miner algorithm produced a classification model with better predictive accuracy, which is 69.69%. It is submitted that prediction of cyberbullying can be conducted using machine learning as described [20].

IV. CONCLUSION

This study's main objective is to develop a classification model to predict the attackers of cyberbullying using Ant Colony Optimization Algorithm for rule induction called the Ant-Miner. The Ant-Miner algorithm produced a good predictive accuracy classification model and was compared to the industry standard J48 algorithm. It can be concluded that both algorithms are competitive in terms of predictive accuracy. However, it should be noted that the Ant-Miner algorithm discovered a rule set that is more accurate than the rule set discovered by J48. Hence, for developing the classification model to predict the attackers of cyberbullying, Ant-Miner is a better algorithm. The prediction is essential to prevent cyberbullying. If cyberbullying can be prevented, it can combat psychological problems among the victims. As such, it can create a healthy environment in the electronic society.

V. FUTURE SCOPE

The study suggests a future study to improve the Ant-Miner algorithm in two ways. First, it is suggested that we find a better heuristic function to increase predictive accuracy. Second, it would be great if the algorithm can cater continuous data on the fly. It is hoped that in future, the government can fight cyberbullying by exploiting the developed model.

VI. CONFLICT OF INTEREST

No conflict of interest involved.

ACKNOWLEDGEMENTS

The authors wish to thank the Ministry of Higher Education Malaysia for funding this study under Fundamental Research Grant Scheme, FRGS/1/2018/STG06/UITM/02/3.

REFERENCES

[1]. Simon, S. (2017). Cyber Victimization: School Experience of Malaysian Cyberbullied Teenagers. *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, *11*(3): 695–702.
[2]. Lam Thye, L. (2017). On the Alert for Cyberbullying.
[3]. Hashim, H. N.M., Mahmood, A., & Suhaimi, N. S. (2018). Prevalent Culture of Online Shaming Amongst UiTM Law Students. *Voice of Academia: Academic*

Series of Universiti Teknologi MARA Kedah Special Issue: Selected Papers Form the 6th International Conference on Public Policy and Social Sciences ICOPS2017), *13*: 53–60.
[4]. Dewey, C. (2015). Can online shaming shut down the Internet's most skin—Crawly creeps? Internet culture. https://www.washingtonpost.com/news/the-intersect/wp/2015/09/16/can-online-shaming-shut-down-the-internets-most-skin-crawly-creeps/?noredirect=on&utm_term=.38dabb5fb759.
[5]. Schenk, A.M. and Fremouw, W.J. (2012). Prevalence, Psychological Impact and Coping of Cyberbully Victims Among College Students. *Journal of School Violence*, *11*: 21-37.
[6]. Nixon, C.L. (2014). Current Perspectives: The Impact of Cyberbullying on Adolescent Health. *Adolescent Health, Medicine and Therapeutics*, *5*: 143-148.
[7]. Lai, C. S., Mohamad, M. M., Lee, M. F., Salleh, K. M., Sulaiman, N. L., Rosli, D. I. and Chang, W. V. S. (2017). Prevalence of cyberbullying among students in Malaysian Higher Learning Institutions. *Advanced Science Letters*, *23*(2): 781–784.
https://www.thestar.com.my/opinion/letters/2017/04/11/on-the-alert-for-cyberbullying/.
[8]. Basyir, M. (2018). Form 2 Student Tries to Commit Suicide after being Accused of Stealing Teacher's iPhone.
https://www.nst.com.my/news/nation/2018/01/329107/form-2-student-tries-commit-suicide-after-being-accused-stealing-teachers
[9]. Cheah, B. (2018). Pro-Skateboarder Tony Hawk Comes to Fatin Syahirah's Defence.
https://www.thestar.com.my/news/nation/2018/09/01/proskateboarder-tony-hawk-comes-to-fatin-syahirahs-defence/.
[10]. Dorigo, M., Maniezzo, V. and Colorni, A. (1996). Ant System: Optimization by A Colony of Cooperating Agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *26*(1): 29–41.
[11]. Awan-Ur-Rahman. (2020). Introduction to Ant colony optimization (ACO). A Probabilistic Technique for finding Optimal Paths.
https://towardsdatascience.com/the-inspiration-of-an-ant-colony-optimization-f377568ea03f
[12]. Parpinelli, R. S., Lopes, H. S. and Freitas, A. A. (2002). Data Mining with An Ant Colony Optimization Algorithm. *IEEE Transactions on Evolutionary Computation*, *6*(4): 321–332.
[13]. Saian, R. and Ku-Mahamud, K. R. (2011). Hybrid Ant Colony Optimization and Simulated Annealing for Rule Induction. 2011 UKSim 5th European Symposium on Computer Modeling and Simulation.
[14]. Al-behadili, H. N. K., Ku-Mahamud, K. R. and Sagban, R.(2019). Annealing strategy for an enhance rule pruning technique in ACO-based rule classification. *Indones. J. Electr. Eng. Comput. Sci.*, *16*(3): 1499–1507.
[15]. Al-Behadili, H. N. K., Sagban, R. and Ku-Mahamud, K. R. (2020). Adaptive Parameter Control Strategy for Ant-Miner Classification Algorithm. *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, *8*(1): 149–162.
[16]. Witten, I. H. and Frank, E. (2002). Data mining: Practical Machine Learning Tools and Techniques with Java Implementations. *Acm Sigmod Record*, *31*(1): 76–77.
[17]. Quinlan, J. R. (2014). *C4. 5:* Programs for machine learning. California: Elsevier.
[18]. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter*, *11*(1): 10–18.
[19]. Kohavi, R. (1995). A Study of Cross-Validation And Bootstrap For Accuracy Estimation And Model Selection. *International Joint Conference on Articial Intelligence*.
[20] Al-Garadi, M. A., Hussain, M. R., Khan, N., Murtaza, G., Nweke, H. F., Ali, I., Mujtaba, G., Chiroma, H., Khattak, H. A. and Gani. A. (2019). Predicting Cyberbullying on Social Media in the Big Data Era Using Machine Learning Algorithms: Review of Literature and Open Challenges. *IEEE Access 7*(1): 70701-70718.

How to cite this article: Saian, R., Fauzi, N. S. F. M., Zakuan, Z. Z. M., Hashim, H. N. M. and Mahmood, A. (2020). Predicting Attackers of Cyberbullying using Ant Colony Optimization in Combating Psychological Effects among the Victims. *International Journal on Emerging Technologies, 11*(5): 205–208.