EPiC Series in Computing
Volume 81, 2022, Pages 262–268
Proceedings of 11th International Congress on Advanced Applied Informatics
EPiC Computing

Social Media Spamming on the Class Room Electronic

Handika Limanto[1], Santiaji[1], Ivan Setiawan[1], Ford Lumban Gaol[1], Harjanto Prabowo[1], Meyliana[1], Winanti[1], Fonny Hutagalung[2]
[1] Bina Nusantara University, Jakarta, Indonesia
[2] University of Malaya, Kuala Lumpur, Malaysia
handikalimanto@binus.ac.id, santiaji@binus.ac.id, fgaol@binus.edu, harprabowo@binus.edu, meyliana@binus.edu, winanti@binus.ac.id, fonny@um.edu.my

Abstract

In the current state of the world, the issue of identifying spammers has received escalating attention due to its influence towards social network security. The popularity of social networking sites made them unsurprisingly easy target for most spammers due to the ease of information sharing they provide. While having important information shared easily is a good thing, the extra bits of harmful objects including viruses or malwares are not. Not to mention the irrelevant information found across almost everywhere in the social media shared by said spammers. Spamming does not only affect social media but could also affect most websites and e-mails. Thus, an extensive action is needed to detect and counter the act of spamming, and this study attempts to review them.

1 Introduction

In this modern era, the internet has become an important part of human life with its huge influence in transportation, communication, and information sharing. In Indonesia, from approximately 264 million people, about 171,17 million or 65% of its residents are connected to the internet through either smartphones, desktops, laptops, or tablets (Pratomo, 2019). With the enormous amount of internet users at the current moment, internet could be considered to be one of the biggest marketing targets to be monitored through due to its high potential.

There have been numerous amounts of successful start-up companies doing their activities relying on the internet. Internet based marketing system is more preferred due to its wider effective area and better efficiency compared to traditional (or conventional) marketing system. One of the preferred marketing strategies, even though sometimes it is considered to be bothersome, is spamming.

Spamming is an activity of using messaging services or systems to send unwanted messages, as well as repeated use of the messaging service or system on the same platform or site. The person doing spamming is called a spammer. Spamming is highly used in internet due to the belief of spamming is a cheap and easy (Aiyar & Shetty, 2018; Baharim & Hamid, 2016).

There are several reasons keeping spammers' existence even though they are harming others depending on their motivation to spam as follows.

1.1 Marketing Purposes

There are many parties who want to market something on the internet in unethical ways, one of which is by sending spam to other people. The contents of the spam messages usually offer goods or services. Spamming can be done by anyone, either in individuals or in groups. There are numerous companies making spam messages via email and SMS as their main marketing channel (Hao & Zou, 2008; Stringhini, 2015)

Most of the time, the recipients of the spam messages feel disturbed each time they receive spam messages. Generally, promotional spam does not generate sales as easily as it used to, as people are more aware of the dangers of spam messages. Sales rarely happens, and even if sales occur, it is often a coincidence and the effects are temporary unless the goods or services you offer are extraordinary (Aiyar & Shetty, 2018)

1.2 Criminal Purposes

Spam for fraudulent purposes is also very prevalent, either through email, SMS, or other media. The fraud spamming purpose is diverse, from taking over (phishing) accounts, data theft, asking for (or even stealing) money, and so forth. As an example, a type of fraud that is very common nowadays is online phishing. Usually the spammer sends a spam email randomly where the message says that the recipient has won a prize draw, or something similarly attractive to the readers (Yu, 2015). The email contains a link to a particular site where the recipient of the spam must log in using an e-mail account. When the victim types in the e-mail and password, the data is recorded so the hacker knows the e-mail address and password of the victim's e-mail (Aiyar & Shetty, 2018).

1.3 Entertainment Purposes

Spam for entertainment purposes is a form of spam that is quite annoying to internet users. Although it does not mean to commit acts of fraud or other criminal acts, but the victim gets a loss and the spammer will benefit if the victim responds. One example of spam activities in this category is spreading certain website links in other people's blog comments. Another example, sharing indecent images or links to banned sites in the comments section of other people's social media status (Aiyar & Shetty, 2018; Baharim & Hamid, 2016; Inuwa-dutse et al., 2018).

Lots of ads are sent to the internet with the aim of promoting a product that is usually considered as spam. This happens because a lot of messages on the internet are usually fraudulent modes. Both frauds directly from person to person in the name of a particular organization, to fraud by buying a spammer to raise the rating of a particular store or organization (Stringhini, 2015). In the past there was also spam like someone who rented or paid someone else to surround his trading place to ask to make visitors look crowded and gave the impression that the goods sold were of high quality to attract buyers.

We are presenting this research paper with the aim to detect spam in general social networking sites using Naive Bayes Classifier method. Our purpose of presenting the research paper titled "Introduction to Social Media Spamming and Ways to Overcome It" is basically to detect and diagnose (Baharim & Hamid, 2016) the huge amount of spam messages available on general social media websites which is found easily by each and every one of us so that this issue can be overcome easily. (Aiyar & Shetty, 2018; Inuwa-dutse et al., 2018) Problem Formulation

1. What's is spamming?
2. What's the impact of spamming?
3. Why we have to be careful about spamming?
4. How to overcome spamming?

2 Research Technique

The technique that we use is an online survey using google form, this survey is distributed to random responders consisting of our friends either on a campus, at home, and in the neighbourhood. After we conducted the survey, in this survey we targeted 15 people. While those who responded were 16 people. Of the 16 respondents, every single one of them have received spam messages and some of them even have been tricked by spam messages.

In this survey we also learned that all respondents were exposed to spam, in this survey we also asked, "Do you agree with the use of spam as one of the advertising methods?" And many said no.

Because a lot of people feel annoyed by spam, we need a way to deal with spam. We also searched for various sources from the internet, and found a method called Naive Bayes Classifier. In this way spam messages can be recognized and blocked by the system. After finding a research method that can be used, we also looked for data about the results of the study using the Naive Bayes Classifier method.

3 Research Method

We use quantitative research methodology in order to collect information, by comparing various sources from the internet and books. Most respondents thought that spam was really irritating from the survey that we did, so we needed a way to circumvent spam, one of the ways is by the process of the Naive Bayes Classifier, or generally called the Bayesian Filter. The new tool used to identify a collection of documents is the Bayesian filter or Naive Bayes Classifier. This algorithm uses the probability and statistical methods proposed by British scientist Thomas Bayes, based on previous experience, to estimate probabilities in the future. This Bayesian statistical approach for anti-spam filter technology was introduced by two groups of researchers, one by Pantel and Lin, and the other by Microsoft Research. But the approach taken by Paul Graham is what makes this Bayesian filtering algorithm popular. The basis of the naive theorem used in programming is the following Bayes formula:

$$P(\mathrm{A}|\mathrm{B}) = (P(\mathrm{B}|\mathrm{A}) * P(\mathrm{A}))/ P(\mathrm{B}) \ldots(1)$$

The chance of occurrence A as B is determined by opportunity B when A, opportunity A, and opportunity B. In the application later this formula changes to:

$$P(\mathrm{C}_i|\mathrm{D}) = (P(\mathrm{D}|\mathrm{C}_i) * P(\mathrm{C}_i))/ P(\mathrm{D}) \ldots(2)$$

Naive Bayes Classifier or can be referred to as Multinomial Naive Bayes is a simplification modelof the Bayes algorithm that is suitable in classifying text or documents.

The formula is as follows:

$$v_{MAP} = \arg\max P(a_1, a_2, \dots, a_n|v_j) \dots(3)$$

With formula (3), formula (1) could be rewritten as follows.

$$v_{MAP} = arg \max_{v_j g V} \frac{P(a_1, a_2, \cdots, a_n|v_j) P(v_j)}{P(a_1, a_2, \cdots, a_n)} \dots(4)$$

$P(a_1, a_2, ..., a_n)$ are constant and can be ignored, thus

$$v_{MAP} = arg \max_{v_j g V} P(a_1, a_2, \cdots, a_n|v_j) P(v_j) \text{ ...(5)}$$

Because $P(a_1, a_2, \dots, a_n|v_j)$ is quite difficult to be counted, we will assume every word in thedocument is disjoint.

$$v_{MAP} = arg \max_{v_j g V} P(v_j) \prod_i P(a_i|v_j) \dots(6)$$

Explanation:

$$P(v_j) = \frac{|docs_j|}{|Example|} \text{ ...(7)}$$

$$P(w_k|v_j) = \frac{n_k + 1}{n + |vocabulary|} \dots(8)$$

Where:

$P(v_j)$: The probability of each document against a set of documents.

$P(w_k|v_j)$: The probability of occurring the word w_k in a document with the category class v_j

$|docs_j|$: frequency of documents in each category

$|Example|$: number of documents available

n_k: frequency of k-th word in each category

$|vocabulary|$: total word in test document

In equation (8) there is an addition 1 to the numerator, this is done to anticipate if there is a word in the test document that does not exist in each training data document. The Naive Bayes Classifier Algorithm is described as follows:

1. Learning

Naive Bayes is an algorithm that is included in supervised learning (Kumar et al., 2015), so early knowledge will be needed to be able to make decisions. Steps:

a. Step 1: Form a vocabulary in each training data document

b. Step 2: Calculate the probability in each category $P(vj)$

c. Step 3: Determine the frequency of each word wk in each category $P(wk|vj)$

2. Classify

a. Step 1: Calculate $P(vj) - P(ai|vj)$for each category

b. Step 2: Determine the category with the maximum value of $P(vj) - P(ai|vj)$.

4 Research Result

From our survey results, we learned that people of this era are susceptible to spamming and most of the time they felt annoyed by the spam messages they receive, and that we need a method to overcome spamming. Of the many respondents who responded, most of them felt disturbed because of spam messages received, therefore a method was needed to overcome spam. The method used in this experiment is the Naive Bayes Classifier method.

As a result of testing the Naive Bayes Classifier method (Kumar et al., 2015), the following are the table results as a result of testing the Naive Bayes method:

Table 1: Spam Received Email

Received Email	Email Clarification Result	Truth Value	Received Email	Email Clarification Result	Truth Value
Spam 1	SPAM	True	Non-Spam 1	NON-SPAM	True
Spam 2	NON-SPAM	False	Non-Spam 2	NON-SPAM	True
Spam 3	SPAM	True	Non-Spam 3	NON-SPAM	True
Spam 4	SPAM	True	Non-Spam 4	NON-SPAM	True
Spam 5	SPAM	True	Non-Spam 5	NON-SPAM	True
Spam 6	SPAM	True	Non-Spam 6	NON-SPAM	True
Spam 7	SPAM	True	Non-Spam 7	NON-SPAM	True
Spam 8	SPAM	True	Non-Spam 8	NON-SPAM	True
Spam 9	SPAM	True	Non-Spam 9	NON-SPAM	True
Spam 10	SPAM	True	Non-Spam 10	NON-SPAM	True
Spam 11	SPAM	True	Non-Spam 11	NON-SPAM	True
Spam 12	SPAM	True	Non-Spam 12	NON-SPAM	True
Spam 13	SPAM	True	Non-Spam 13	NON-SPAM	True
Spam 14	SPAM	True	Non-Spam 14	NON-SPAM	True
Spam 15	SPAM	True	Non-Spam 15	NON-SPAM	True
Spam 16	SPAM	True	Non-Spam 16	NON-SPAM	True
Spam 17	SPAM	True	Non-Spam 17	NON-SPAM	True
Spam 18	SPAM	True	Non-Spam 18	NON-SPAM	True
Spam 19	SPAM	True	Non-Spam 19	NON-SPAM	True
Spam 20	SPAM	True	Non-Spam 20	NON-SPAM	True
Spam 21	SPAM	True	Non-Spam 21	NON-SPAM	True
Spam 22	SPAM	True	Non-Spam 22	NON-SPAM	True
Spam 23	SPAM	True	Non-Spam 23	NON-SPAM	True
Spam 24	SPAM	True	Non-Spam 24	NON-SPAM	True
Spam 25	SPAM	True	Non-Spam 25	NON-SPAM	True
Spam 26	SPAM	True	Non-Spam 26	NON-SPAM	True
Spam 27	SPAM	True	Non-Spam 27	NON-SPAM	True
Spam 28	SPAM	True	Non-Spam 28	NON-SPAM	True
Spam 29	SPAM	True	Non-Spam 29	NON-SPAM	True
Spam 30	SPAM	True	Non-Spam 30	NON-SPAM	True
Spam 31	SPAM	True	Non-Spam 31	NON-SPAM	True
Spam 32	SPAM	True	Non-Spam 32	NON-SPAM	True
Spam 33	NON-SPAM	False	Non-Spam 33	NON-SPAM	True
Spam 34	SPAM	True	Non-Spam 34	NON-SPAM	True
Spam 35	SPAM	True	Non-Spam 35	NON-SPAM	True
Spam 36	SPAM	True	Non-Spam 36	NON-SPAM	True
Spam 37	SPAM	True	Non-Spam 37	SPAM	False
Spam 38	SPAM	True	Non-Spam 38	NON-SPAM	True
Spam 39	SPAM	True	Non-Spam 39	NON-SPAM	True
Spam 40	SPAM	True	Non-Spam 40	NON-SPAM	True
Spam 41	SPAM	True	Non-Spam 41	NON-SPAM	True
Spam 42	SPAM	True	Non-Spam 42	NON-SPAM	True

Spam 43	SPAM	True	Non-Spam 43	NON-SPAM	True
Spam 44	SPAM	True	Non-Spam 44	NON-SPAM	True
Spam 45	SPAM	True	Non-Spam 45	NON-SPAM	True
Spam 46	SPAM	True	Non-Spam 46	NON-SPAM	True
Spam 47	SPAM	True			

By repeating the experiment above 120 times, we obtain the table below:

Table: 2 Observed Spam Keyword

Data Count	Spam Count	Non-Spam Count	Spam Keyword Count	Non- Spam Keyword Count	Error Count
20	10	10	1132	914	1
40	20	20	4780	1587	12
60	3	30	5112	4067	3
80	4	40	5378	5043	3
93	47	46	5513	5710	3
120	60	60	7964	8227	3

From Table 1, it can be observed that there are 3 errors that have been generated, namely 2 data which are spam, classified as non-spam and 1 non spam data classified as spam. Then the percentage of error is obtained as follows:

$$Error = 3/120 * 100\% = 2.5\%$$

5 Discussion

From the results of the survey and research, we discuss whether the results obtained are in accordance with what has become the problem formulation. The results of our discussion that each of the results obtained is in accordance with our hypothesis that spam is disturbing. The results of the study also showed satisfactory results with an error ratio of only 2.5%. With these results it can be said that the error ratio of the naïve Bayes method is still in an acceptable ratio. The results of our discussion also found that today there are still many other ways that can be used to select spam.

6 Conclusion

Spam is one of the most disturbing things for most people, and for that reason spam is researched a lot recently and it turns out there are several ways to reduce spam. One way to reduce spam is using the Naive Bayes Classifier method. With this method, spam can be filtered by more than 70%. Although this method is very effective for reducing spam, it is expected that in the future there will be more advanced and effective ways to deal with spam. In the next study we also expected that with sufficient time and adequate resources a better spam filter could be generated, as well as the results of testing the Naive Bayes Classifier method that could be better explained.

7 Acknowledgment

The research is funded by the Research and Technology Transfer Office of Bina Nusantara University as a portion of Bina Nusantara University's International Research Grant contract number: No.026/VR.RTT/IV/2020 of PIB47 and agreement term: 6 April 2020.

References

Aiyar, S., & Shetty, N. P. (2018). *N-Gram Assisted Youtube Spam Comment Detection*. Procedia Computer Science, 132(Iccids), 174–182. https://doi.org/10.1016/j.procs.2018.05.181

Baharim, K. N., & Hamid, S. (2016). *Opinion Spamming in Social Media: A Brief Systematic Review Opinion Spamming in Social Media*. A Brief Systematic Review. September.

Hao, J., & Zou, J. (2008). *A Real-Time Payment Scheme for SIP Service Based on Hash Chain*. 279–286. https://doi.org/10.1109/ICEBE.2008.19

Inuwa-dutse, I., Liptrott, M., & Korkontzelos, I. (2018). *Neurocomputing Detection of spam-posting accounts on Twitter.* Neurocomputing, 315, 496–511. https://doi.org/10.1016/j.neucom.2018.07.044

Kumar, D., Singh, R., Kumar, A., & Sharma, N. (2015). *An Adaptive Method of PCA for Minimization of Classification Error Using Naïve Bayes Classifier*. Procedia Computer Science, 70, 9–15. https://doi.org/10.1016/j.procs.2015.10.018

Pratomo, Y. (2019). APJII: *Jumlah Pengguna Internet di Indonesia Tembus 171 Juta Jiwa*. Tekno.Kompas.Com. https://tekno.kompas.com/read/2019/05/16/03260037/apjii-jumlah- pengguna-internet-di-indonesia-tembus-171-juta-jiwa

Stringhini, G. (2015). *Detecting spammers on social networks Detecting Spammers on Social Networks*. December 2010. https://doi.org/10.1145/1920261.1920263

Yu, S. (2015). *Covert communication by means of email spam: A challenge for digital investigation*. Digital Investigation, 13, 72–79. https://doi.org/10.1016/j.diin.2015.04.003