International Journal of Research in Advent Technology, Vol.4, No.5, May 2016
E-ISSN: 2321-9637
Available online at www.ijrat.org

Implementation of SMS Spam Detection System

Suraj J. Warade[1], Pritish A. Tijare[2], Swapnil N. Sawalkar[3]
M.E (Pursuing)[1], Associate Professor[2], Assistant Professor[3]
Computer Engineering[1], Information and Technology[2], Computer Science and Engineering[3]
Sipna C .O. E. T, Amravati, India[1, 2, 3]
surajwarade@hotmail.com[1], pritishtijare@rediffmail.com[2], swapnil.sawalkar@gmail.com[3]

Abstract- In the today's world the use of mobile phone increases rapidly. And hence the advertisers start to use of SMS for their advertisement. At the beginning the advertisers are send their promotional messages through SMS gateways. But due to increasing number of promotional messages the service provider start the service do not disturb(DND), the DND service restrict only the SMS send through SMS gateways and hence the advertisers start to send their promotional advertisement messages through spammer's mobile phones. The approach discussed in this paper detects these messages sent through spammers mobile and restrict it from being in inbox.

Index Terms – spam filtering, SMS spam, mobile spam, text classification, SMS.

1. INTRODUCTION

Due to daily increase in mobile phone users, Short message service (SMS) is very widely use text messaging service. Firstly it was designed for Global System for Mobile Communication (GSM), but now it is also available on Code Division Multiple Access (CDMA)[8]. Hence, the popularity of SMS increases over the years. The advertiser thought it is the best way to advertise their product, the one reason behind this is, first all promotional messages such as store opening announcement, shopping discounts, credit card of bank etc. are sent over the email but as their spam detection facility is available all the promotional messages are going to the spam folder. And hence the companies and advertiser start sending their messages over the mobile phones as the SMS, at the beginning all messages are sent through the SMS gateways as this is easy way to broadcast message to multiple users [3]. On arrival of every message the user have to check inbox and hence because of these messages not only the mobile user is distracted but also it causes to quick fill of users inbox and user have to waste his valuable time to read and delete these unsolicited message hence, the communication service provider provides the service DND (do not disturb) which restricts the unsolicited spam messages sent through the SMS gateways [4].

When DND service restrict the messages over mobile phones. The advertiser comes with solution that send the promotional messages through the spammers mobile phone as there is not at all any restriction for messages sent through spammers mobile and it is easy and thanks to communication service provider companies unlimited SMS plans it is opportunity for spammers to use it for sending promotional unsolicited messages.

In this paper we provide the scheme for spam filtering by employing the two feature selection techniques chi-square (CHI2) and information gain (IG), and Bayesian based binary classification algorithm to classify message. Furthermore the one approach of mutual relation checking between sender and receiver is employed to check whether the message is from known or unknown sender. If there is mutual relation between sender and receiver and no spamming content are found then message goes to inbox. If there is no mutual relation between sender and receiver and no spamming content then message goes to the unknown sender box. If there is spamming content in the message then it goes to the spam box.

The real time windows phone application is developed for the display result of experimental study on client side. The filtering mechanism is work on SMS in English language.

2. LITRATURE SURVEY

Huang Wen-Liang, Liu Yong, Zhong Zhi-Qiang, and Shen Zhong-Ming proposed a complex-network based SMS filtering algorithm which compares an SMS network with a phone call communication network. Because such comparison can provide additional features, SMS networks and obtaining well-aligned phone-calling networks that can be aligned perfectly is difficult in practice. In this paper, author presents an efficient SMS spam detection algorithm that only considers the SMS communication network. Authors first analyze characteristics of the SMS network, and then check the properties of different sets of meta-features including static feature, network features and temporal features. Authors combine these features into an SVM classification algorithm and evaluate its performance on a real SMS dataset and a video social network benchmark dataset. They also compare the SVM algorithm and KNN based algorithm to reveal

the advantages of the former. Our experimental results demonstrate that SVM based on network features can get 7%-8% AUC (Area under the ROC Curve) improvement as compared to some other commonly used features [1].In [2]authors consider a local concentration based extraction approach. Two implementation strategies are designed for detecting the SMS spam as fixed length sliding window and variable length sliding window.

A novel framework for SMS spam filtering is proposed to be able to block unsolicited SMS messages by Uysal, S. Gunal, S. Ergin, E. Gunal. In the filtering framework, distinctive features representing SMS messages are identified using CHI2 and IG based features election methods. The selected features upsets with varying sizes are then fed into two different Bayesian based classification algorithms, namely the binary and probabilistic models, to classify SMS messages as either legitimate or spam. Additionally, the proposed SMS spam filtering scheme is employed to develop a real-time mobile application running on the mobile phones with Android operating system [3]. In [4] author examines the effectiveness of various content-less features that range from network and to time-oriented categories. He find that some intuitively appealing features are in fact not very effective, whereas a combination of temporal and network features can be very useful in training high performance classifiers for spammer detection.

Zi Chu, S. Gianvecchio, Haining Wang, and Sushil Jajodia focus on the classification of human, bot, and cyborg accounts on Twitter. Author first conduct a set of large-scale measurements with a collection of over 500,000 accounts. They observe the difference among human, bot, and cyborg in terms of tweeting behavior, tweet content, and account properties. Based on the measurement results, author proposes a classification system that includes the following four parts: an entropy-based component, a spam detection component, an account properties component, and a decision maker. It uses the combination of features extracted from an unknown user to determine the likelihood of being a cyborg, bot or human [5]. Stylistic feature that characterizes the manner in which SMS is written is introduced by authors [6]. First authors determines the style of spam messages written in Korea and they found that the most of the spam messages over Korea are sent in either English or Korean language in the same pattern and hence they provide the approach of stylistic pattern matching for detection of the spam messages. They only focus on the two languages English and Korean [6]. K. Uysal, S. Gunal, S. Ergin, and E. Sora Gunal, proposed a system for Extraction and selection on SMS spam filtering on the mobile. The author suggested the technique of data extraction of datasets like web link, alphabets, numbers, length of the message etc [7].

Artificial Immune System (IAS) of Soft Computing which Motivated by the Biological Immune System (BIS). Particularly it is based on how human immune system resist against disease and infections in the same way the mobile spam could be handled [8].Brief description about the SMS spamming methods like content matching, pattern matching and current practices for detecting spam and the data are provided in[9].

An efficient Read Aligner for next generation sequencing reads structures to detect and compare the results of web spam bot sand Viruses. This paper proposed a method of using a bio informatics pattern matching algorithm to evaluate signature-based virus/spam detection in Windows [10].

In a mobile network, viruses and malwares can cause privacy data leakage, extra charges, and remote listening. Author presented a two-layer network model for simulating and analyzing the propagation dynamics of SMS-based and BT-based viruses. This model characterizes two types of human behavior and mobile behavior, in order to observe and uncover the propagation mechanisms of mobile viruses [11].

To address the limitations of the state of research on SMS spam detection, Amir Karani and Lina Zou propose a content-based method that leverages lexical semantics. Instead of relying on individual words, proposed method uses semantic categories of words as features, which allows us to handle variations in word choices by spammers. To address the limitations of the state of research on SMS spam detection, they propose a content based method that leverages lexical semantics. Instead of relying on individual words, their proposed method uses semantic categories of words as features, which allows us to handle variations in word choices by spammers. In addition, using categories of words as features also helps to reduce the feature space, which in turn improves the efficiency of spam detection that has significant implications for SMS users. An empirical evaluation of the proposed methods has shown promising results[12].

3. FEATURE EXTRACTION

Unlike e-mail the SMS message does not contain some text along with graphics, attachments, hyperlinks [13], It only contains limited 160 character for the entire message. So, mainly classification can be done on two main parameters text classification or classification based on who is sending the message. In this paper the message is separated as legitimate, spam, or unknown sender using both of these techniques.

In this study we used classical bag-of-words approach to extract features from SMS messages. In this approach, term occurrence considered rather than term ordering. Thus, every different term in an SMS

message collection corresponds to a feature. As a result an SMS message is represented by a multi-dimensional feature vector where the elements of the feature vector are constituted by the weighted values of corresponding terms.

4. FEATURE SELECTION

The two feature selection method used in this paper to detect spam are chi-square(CHI2) and information gain(IG).

4.1 CHI2

The first feature selection method used in this paper is CHI2. This method is applied to determine the independence of two different events [3]. The two events A and B are independent if,

$$P(AB) = P(A)P(B) \qquad (1)$$

CHI2 is statistical test to measure occurrence of term against the expected number of occurrence of the term [14]. The high value of CHI2 indicates the hypothesis of independence is not true. In CHI2 the independent terms are features and dependent terms are the categories (legitimate or spam). CHI2 can be computed using

$$\text{CHI2} = \sum \frac{(O-E)^2}{E} \qquad (2)$$

Where O is observed count and E is expected count.

4.2 IG

Another feature selection method used in this paper is well known ranking measure IG (Information Gain). It is based on impact of feature on entropy of message i.e how much information the presence or absence of term contributes to make correct classification [3].

IG provides ranking to its features and when the ranking of the term reaches to certain threshold it treated the message as the particular class.

5. CLASSIFICATION

In this study, Bayesian based binary classification model is used to categorize the messages are either spam or legitimate. The binary model considers whether the terms representing the feature vector are found or not in message [3],[6]. The probability of class C is legitimate or spam is calculated using

$$P(C_i \mid X) = \sum_{j=1}^{n} \begin{cases} wD_{ij}, & \text{if } jth \text{ term occurs in SMS message} \\ -D_{ij}, & \text{otherwise} \end{cases},$$

Where D_{ij} is ratio of number of SMS in class C_i containing jth term to the total number of messages in class $C_{i,}$ n is a size of feature vector, and w is weight values.

6. PROPOSED POISON REMOVAL APPROACH

There are some spammers who try to bypass the spamming criteria of content checking so they intentionally misspelled the word or uses special characters in the place of original alphabet called as poisoning, like spell offer as 0ffer, sale as $ale. And because of these words are not in our spamming dictionaries used by our feature selection methods the spam mechanism tag them as non spam entity and spam message get placed in the inbox of user. We cannot add all combination of the spam words with special characters or probable misspelled word to the dictionary because it increases the length of the dictionary so the overhead of the feature selection and extraction increases and it cannot provide the proper solution to the problem of poisoning the dictionary.

To overcome this type of poisoning we proposed a poison removal scheme in which the words in the SMS message are divided into small strings of characters and match with our selected feature terms strings if these strings are match with the feature words up to certain threshold then the system treat it as a feature term, and process it as per selected criteria either CHI2 or IG. By using this approach the stemming and stop word removal is also carried out.

If input word contains 3 to 5 characters then the string divided into substrings of 2 characters, and matches it with the substrings of the terms presented in the spam dictionary in the forward index. If the substrings in the input words matches with the substring in spam dictionary over 50% then it treat as spam dictionary word and increase the count of spam word in the message.

Ex:

Words	**Substrings**			
0ffer	0f	Ff	Fe	er
Offer	Of	Ff	Fe	er
	X	√	√	√

Table 1: Example 1 of poison removal approach

In the above example offer is written as 0ffer according to existing content based spam detection approach it's not a spam entity but in fact it's a spam entity and spammers tries to bypass your content checking mechanism by such activities. But with our proposed poison removal approach first it divide the 0ffer into substring of 2 characters in forward indexing and match the substrings with every word in the spam dictionary and it found that 0ffer matches 75% with offer, and our determined criteria is of 50% and hence it treat it as a spam entity and increase the spam count.

During the study of poison removal approach we found that the words larger than 6 characters are not correctly identified by above substrings of 2 characters. Hence, if input word contains more than 6 characters then the string divided into substrings of 3 characters, and matches it with the substrings of the terms presented in the spam dictionary in the forward index. If the substrings in the input words matches with the substring in spam dictionary over 30% then it treat as spam dictionary word
Ex:

Words	**Substrings**					
d!scount	d!s	!sc	Sco	Cou	Oun	unt
Discount	dis	Isc	Sco	Cou	Oun	unt
	x	X	√	√	√	√

Table 2: Example 2 of poison removal approach

In the above example discount is written as d!scount according to existing content based spam detection approach it's not a spam entity but in fact it's a spam entity and spammers tries to bypass your content checking mechanism by such activities. But with our proposed poison removal approach first it divide the 0ffer into substring of 3 characters in forward indexing and match the substrings with every word in the spam dictionary and it found that d!scount matches 66.64% with discount, and our determined criteria is of 30% and hence it treat it as a spam entity and increase the spam count.

7. EXPERIMENTAL STUDY

The working of system is described in following steps,

1) First the sender and receiver have to be register over the mobile service provider to start messaging service. It is a prototype model of user's mobile phone.
2) The mobile service provider allows the connected user to communicate.
3) There are two databases are used here one for storing the spam content dictionary and other for the call log storage.
4) The result analyzer analyzes the databases with the incoming SMS message and tags them with the appropriate label spam, legitimate or unknown sender.

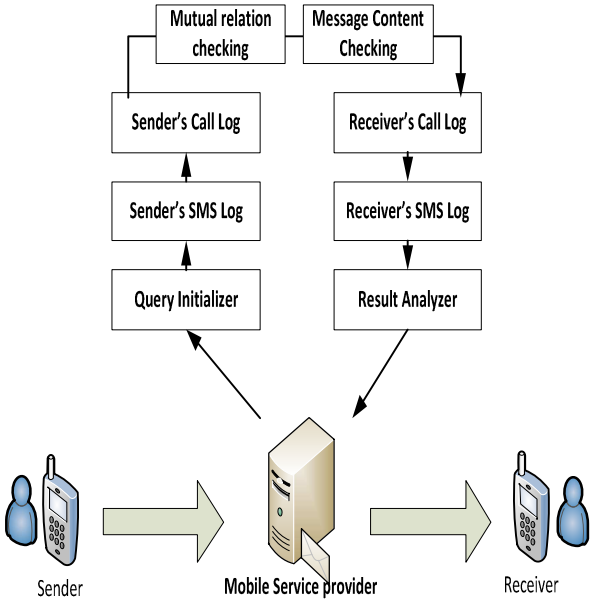

Figure 1: System Architecture

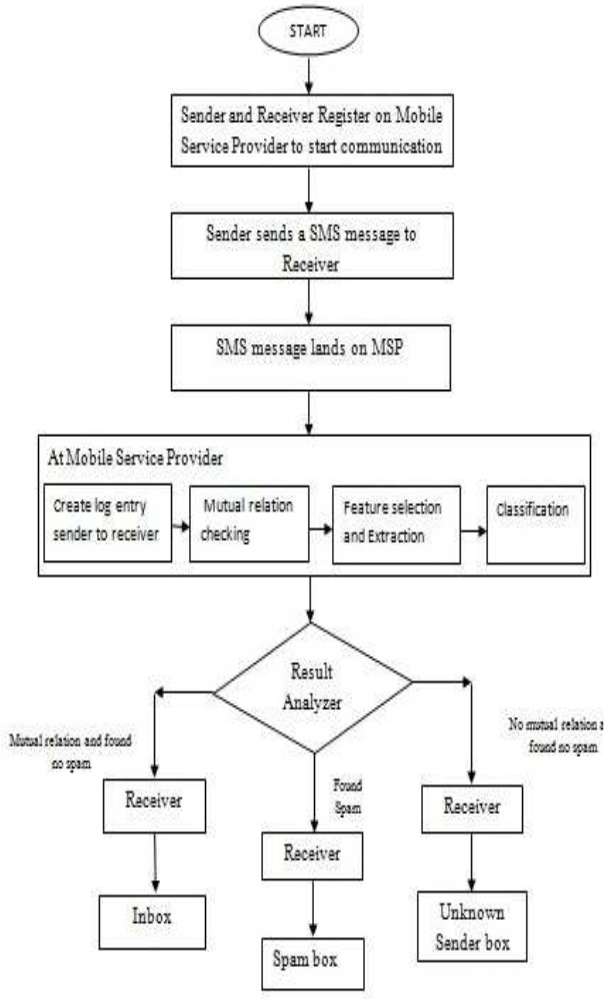

Figure 2: Systems workflow

The above figure1 and 2 describes the process or the execution of the proposed system where sender will send a text message which will land at mobile service provider server. Once the message is received by the server then server will send the sender and the receivers address to relationships analysis module which will give the concluded result in positive or the negative format. Here the relation analysis module will look in to previous SMS log between the sender and receiver and check mutual relation between sender and receiver. System will also check for the content of the message with the selected criteria as CHI2, IG, CHI2 with proposed poison removal approach and IG with proposed poison removal approach. After the successful result from result analyzer system will apply and normal or spam as a tag to message and forward it to receiver.

The criteria for inbox, unknown sender, and spam box is as,

Inbox

If there is mutual relation between sender and receiver and found not spam then the message forward to inbox by the server.

Unknown Sender

If there is no mutual relation between sender and receiver and found not spam then the message forward to Unknown sender box by the server.

Spam

If the message found to be spam by CHI2 or IG then it forward to the Spam box by the server. Then it doesn't matter there is mutual relation is present or not.

The proposed spam filtering scheme is evaluated by using publically available SMS messages in our day to day to life. Following table summarize the top 50 terms based on CHI2 and IG in descending order.

offer, prize, price, discount, bumper, deals, additional, shop, subscribe, download, demand, exclusive, activate, coupon, amazing, credit, voucher, package, limited, sale, latest, link, award, click, launch, free, store, chance, www, market, maximum, lucky, special, collection, dollar, rupee, lose, extra, claim, consult, congratulation, rate, holiday, flat, crore, cash, week, bazaar, collect, info

Table 3: Top 50 Terms Based on CHI2 and IG (Descending Order)

The results for all the techniques used are summarized in the figure 3 to 6.

Legitimate: Legitimate SMS found according to selected criteria from known and unknown sender.

Spam: Spam SMS found according to selected criteria from known and unknown sender.

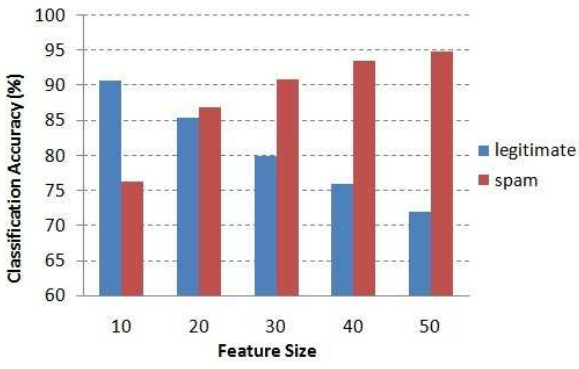

Figure 3: Accuracy for CHI2 based model

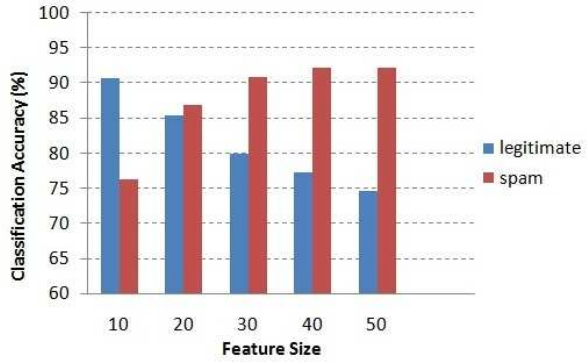

Figure 4: Accuracy for IG based model

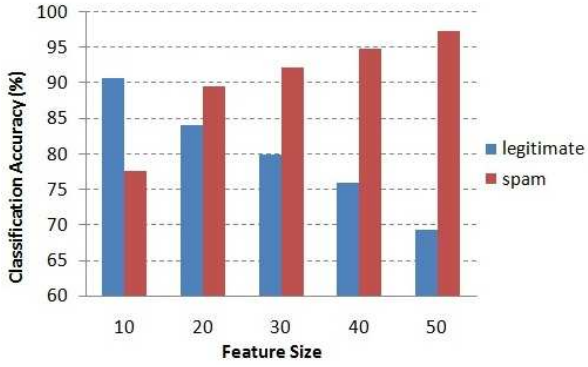

Figure 5: Accuracy for CHI2 based model with proposed poison removal approach

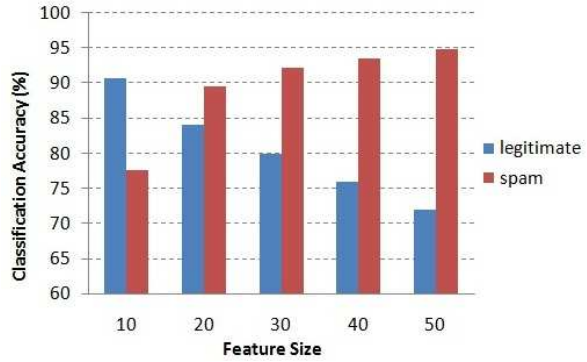

Figure 6: Accuracy for IG based model with proposed poison removal approach

The above result graph shows the high accuracy is obtained for the spam and legitimate messages especially with the lower number of features. As the feature size increases the accuracy of class spam increases but the accuracy of legitimate messages is decreases gradually. The overall average best result is obtained at the feature size 20.

8. CONCLUSION

Mobile phone spam is a form of spam directed at the text messaging or other communications services of mobile phones. By using Mobile service provider level SMS spam detection system, the system will first look up in log data and check for Spam content in the message and our proposed poison removal approach extend the dictionary spam detection ability by adding a character level spell checker in to system. So the overall accuracy of system is increased.

By using this system, the problems occur due to spam messages like balance deduction, early inbox filling, time wastage in reading and deleting spam messages are getting solved.

Acknowledgments

It is a matter of great pleasure by getting the opportunity of highlighting fraction knowledge, I acquired during my technical education through this paper. This would not have been possible without the guidance and help of many people. This is the only page where I have opportunity of expressing my emotions and gratitude from the care of my heart to them. This paper would not have been successful without enlightened ideas; timely suggestion and keen interest of my respected Guide **Prof. Pritish A. Tijare** and Co-Guide **Prof. Swapnil N. Sawalkar** without their best guidance this would have been an impossible task to complete.

Appendix

Suraj J. Warade, Pritish A. Tijare, Swapnil N. Sawalkar "An Approach for SMS Spam Detection " IJRAT-International Journal of Research in Advent Technology, Vol.2, No.12, December 2014. E-ISSN: 2321-9637

REFERENCES

[1] Huang Wen-Liang, Liu Yong, Zhong Zhi-Qiang, and Shen Zhong-Ming "Complex network based SMS filtering algorithm" China Academic Journal Electronic Publishing House 13, 2008.

[2] Y. Zhu and Y. Tan, "A local-concentration-based feature extraction approach for spam filtering", IEEE Trans. on Information Forensics and Security, vol. 6, no. 2, pp. 486 – 497, 2011.

[3] Uysal, S. Gunal, S. Ergin, E. Gunal, "A Novel Framework for SMS Spam Filtering", IEEE International journal 978-1-4673-1448-0/12 2012.

[4] Qian Xu, Evan Wei Xiang and Qiang Yang "SMS Spam Detection Using Non-Content Features" IEEE Intelligent System,2012, 10.1109/MIS.2012.3

[5] Zi Chu, S. Gianvecchio, Haining Wang, and Sushil Jajodia, "Detecting Automation of Twitter Accounts: Are You a Human, Bot, or Cyborg?"IEEE Transactions On Dependable And Secure Computing, Vol. 9, No. 6, November/December 2012.
[6] D.-N. Sohn, J.-T. Lee, K.-S. Han, and H.-C. Rim, "Content based mobile spam classification using stylistically motivated features," Pattern Recognition Letters, vol. 33, pp. 364–369, 2012.
[7] K. Uysal, S. Gunal, S. Ergin, and E. Sora Gunal, "Detection of sms spam messages on mobile phones," Proc. of IEEE 20th Signal Processing and Communications Applications Conference, 2012.
[8] Tarek M. Mohhamad, Ahemed M. Mahofouz "SMS Spam Filtering Technique Based on Artificial Immune System "IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012 ISSN (Online): 1694-081
[9] Sarah Jane Delany, Mark Buckley, Derek Greene "SMS Spam Filtering: Methods and Data", Expert Systems with Applications 39(10), p9899-9908, Elsevier 2012.
[10] M. Elloumi, P. Hayati, C IIiopoulos, J.Mirza ,S. Pissis, A. Shah, "Comparison for the Detection of Virus and Spam using Pattern Matching Tools", IEEE International journal ISBN:978-1-4673-5613-8 2013.
[11] Chao Gao and Jiming Liu, "Modeling and Restraining Mobile Virus Propagation", IEEE Transactions On Mobile Computing, Vol. 12, No. 3, March 2013
[12] Amir Karani and Lina Zou "Improving Static SMS Spam Detection by Using New Content-based Features", Twentieth Americas Conference on Information Systems, Savannah, 2014.