Hybrid Spam Filtration Method using Machine Learning Techniques

S.Jancy Sickory Daisy, A.Rijuvana Begum

Abstract: ***Electronic mail (e-mail) is one of the most prevalent approaches for online communication and transferring data through web because of its quick and easy distribution of data, low distribution cost and permanency. Despite these benefits there are certain weaknesses of e-mail. Among these, spam also known as junk e-mail tops. Spam is set of unwanted or inappropriate messages sent over the internet to a massive amount of users for the purpose of marketing, phishing, disseminating malware, etc.With the internet becoming the dominant platform anti-spam solutions are of great use today. This paper illustrates an efficient hybrid spam filtration method using Naïve Bayes algorithm and Markov Random Field technique, which detects and filters spam messages. The proposed method is evaluated based on its accurateness, meticulousness and time consumption. The results confirm that the proposed hybrid method achieves high percentage of true positive rate in finding e-mail spam messages.***

Keywords: E-mail spam, Naïve Bayes algorithm, Markov Random Field.

I. INTRODUCTION

Electronicmail (email or e-mail) is one of the most commonly used services on the Internet.People simultaneously exchange messages between them using e-mail.It is highly difficult for some of individuals to visualize their life without email. Because of these reasons, email has become a optimalagent for communication between people who may have ill intentions.

The tremendousdevelopment of internet also increases the number of email users which results in increased rate of spam emails. According to a survey 70% of the email traffic is spam. Hence it is necessary to categorize emailsinto different classes for the detection of fraudulent emails. The categorizationof emails such as legitimate and illegitimatecan be done based on theintension of it.

Illegitimate Emails

Illegitimate emails may contain unnecessary informations, phishing emails [6], [7], [8], frightening messages, or plans for bad assasinations. Generally,emails do not disclose the identity of the sender.

An illegitimate email may fall in anyone of the following categories:

- Annoys the receiver unnecessarily.
- Dishonesty is the main intension.
- Intended to getessential information from the receiver.It may cause damage to receiver's computer.
- It may forward receiver to illegalwebsite.
- In this paper an efficient hybrid spam filtration methodis employed to detect and filter fraudulent email. A fraudulent email contains unreliable messagewhich is not requested by the receiver. Some of the characteristics of such emails are as follows:
- It attracts the recipient with a deceiving subject line and appears to the recipient as an important notice. The subject line bypasses spamming filtersbyhavingnumeric characters or letters in it.
- Itconsists of only attractive messages.
- It reveals only the fake address or identity of the sender and appears as if it comes from the organisation it claimed to be.
- To receive the confidence of the recipient it looks similar to the legitimate website by having similar wordings, images, links etc.
- The hyperlink takes the recipient to a fraudulent website.
- It obtains the personal/financial information from the recipient and store it in a database for the access of the fraudsters.

Naive Bayes Classifier

The Naive Bayes algorithm identifies the frequency and combination of values [4] from a dataset and performs a simple probabilistic classification operation. For the identification of spam e-mail Naive Bayes classifier usesa collection of words. For training the document classifier the frequency of occurrence of each word is used. Spam and non-spam e-mails have particular probabilities of occurrence of words. Example,for the words such as Congratulations and 'Act Now!' Bayesian spam filters learned high spam probability and a low spam probability for words seen in non-spam e-mail, such as names of friend and family member. For this reason, Naive Bayes technique calculates the spam or non-spam e-mail probability by using Bayes theorem as shown in formula below.

$$P(sp|w) = \frac{P(sp).P(w|sp)}{\mathrm{P(sp).P(w|sp) + P(nsp).(w|nsp)}}$$

Where:

1. P(sp|w) is the conditional probability that given the e-mail is in spam itcontains particular word in it.
2. P(sp) is the probability that any given message is spam.
3. P(w|sp) is the conditional probability that the specific word appears in spam message.
4. P(nsp) is the probability that any given message is not spam.
5. P(w|nsp) is the conditional probability that the particular word appears in non-spam message.

Revised Manuscript Received on July 07, 2019.

S.Jancy Sickory Daisy, Department of CSE, PRIST University, Thanjavur-613403.

A.Rijuvana Begum, Department of ECE, PRIST University, Thanjavur-613403

Markov Random Fields

To accurately model the statistical behaviors of spam this method is used. Relative transition probability that given one word, predict what the next word will be is used in Markovian model.It is based on the theory of Markov chains defined by AndreyMarkov. Markov models are classified into two types: visible Markov model, and hidden Markov model or HMM. The current word is considered to contain the entire state of the language in a visible Markov model, while the state is hidden and presumes only that the current word is probably related to the actual internal state of the language in a hidden Markov model.

A Markov random field(MRF)is an undirected graph G=(V,E) where G is a set of vertices and E is a set of edges which could satisfy the following set of Markov properties.

- Pairwise Markovianity:The non-adjacent variables are conditionally independent when conditioned on all other variables.
- Local Markovianity:A variable is conditionally independent on all other variables given its neighborhood.
- Global Markovianity:R satisfies the global Markov property with respect to a graph G if for any disjoint vertex subsets X, Y, and Z, such that Z separates Xand Y, the random variables SX are conditionally independent of XT given XU . Here, Z separates Xand Y if every path from a node in X to a node in Y passes through a node in Z.

The Global Markov property is stronger than the Local and Pairwise Markov properties.However, for a positive probabilitythese three Markov properties are equivalent.

II. RELATEDWORK

The literature survey for spam detection is discussed in this section. Different spam detection methods proposed by researchers for finding spam in e-mail messages are discussed here.

Shalendra Chhabra et. al. discussed a Markov Random Field model [1] based approach for spam filtration which specifies the advantages of the neighborhood relationship (MRF cliques) among words in an email message. The weighting sequences define a set of clique potentials, where the neighborhood of a single word is given by the words surrounding it.Nurul Fitriah Rusland, Norfaradilla Wahid, Shahreen Kasim, Hanayanti Hafittested the performance of Naive Bayes algorithm and concluded that type of email and number of instances of dataset influence the performance of Naive Bayes. Time taken to build model is faster with less attribute.Ali Shafigh Aski etc. al. proposed an efficient algorithm [3] to filter spam with low error rates and high efficiency using a multilayer perception model. The proposedalgorithm can be implemented on a Mail Server and Mail Client to eradigate problems, like reduction of bandwidth and low efficiency, from which users frequently suffer. M. Tariq Banday etc. al discussed the effectiveness and limitations [4] of statistical spam filters in which CBART and NB classifiers based filters showed betterperformance.Konstantin Tretyakov etc. al. gives an overview [5] of some of the most popular machine learning methods (Bayesian classification, k-NN, ANNs, SVMs) and their applicability to the problem of spam-filtering. Neelam Choudhary etc.presented a novel approach [6] that can detect and filter spam messages using machine learning classificationalgorithms. The proposed approach achieved 96.5% true positive rate and 1.02% false positive rate for Random Forest classification algorithm.

An evaluation of five supervised learning methods [7] in the context of statistical spam filtering was performed by Le Zhang etc. al. Support Vector Machine, AdaBoost and Maximum Entropy Model are top performers in this evaluation, sharing similar characteristics: not sensitive to feature selection strategy, easily scalable to very high feature dimension and good performances across different datasets. Performance evaluation and robustness against attack [8] of spam filter using Naive Bayes, Support Vector Machine (SVM), and LogitBoost machine algorithms was experimented by Steve Webb etc. al. These filters were found to be highly effective in identifying spam and legitimate email, with an accuracy above 98%. Alexy Bhowmick etc. al. reviewed content-based [9] spam filtering techniques based on machine learning methods and achieved tremendous success.

Sarwat Nizamani etc. al. [10] found that by including advanced features, the detection of fraudulent detection task increases regardless of classification method used. The proposed method was extended by employing header information of the email for the task of fraudulent email detection task.A software framework for spam detection, analysis and investigation [11] was elaborated by Son Dinh etc. al. to reduce investigation efforts by consolidating spam emails into campaigns.An efficient algorithm [12] to filter spam using machine learning techniques can be modeled to be implemented on a mail server and mail client with low bandwidth and efficiency.A Hybrid E-Mail Spam Filtering [17] Technique was proposed by Subhana Khan etc. al. to improve accuracy and efficiency of classification using Bayesian classifier and back propagation neural network algorithms which lacked in terms of time consumption.

III. PROPOSED APPROACH

In this section, we present an efficient method to detect and filtere-mail spam messages using hybrid method. The proposed hybrid spam filtration method uses Naïve Bayes algorithm and Markov Random Field technique.

There are three phases in the proposed research methodology to filter spam messages as given below:

1. Preprocessing
2. Naïve Bayes classifier
3. Markov Random Field technique

Preprocessing

Incomplete, aggregate, noisy and missingvalues are handled in this preprocessing step. Conjunctionwords and articles which are not useful inclassification of e-mail messagesare removed from email body.

Naïve Bayes classifier

After preprocessing step,NaïveBayes classification algorithm is applied. The algorithm is applied initially for the header information and the classifier is trained. The content of e-mail message is tested later against Naïve Bayes algorithm to make a verification of the trained data and correspondingly the trained classifier is updated.

Markov Random Field technique

The global markovian property is applied on the designed classifier because of its strength in classifying e-mail messages.

Algorithm for the proposed hybrid e-mail spam filtering method is given below.

1. Preprocess the e-mail message entering into system as given below for useful classification of e-mail messages.
   - Remove conjunction words such as and, but, while, although etc.
   - Remove articles like a, an, the from email body
   - Handle incomplete, aggregate, noisy and missing values properly
2. IF (preprocessed data matches with the designed classifier which is available in ontology) THEN
   - Classify the e-mail message as Legitimate mail(L)/Spam mail(S) using ontology
3. ELSE
   - Train the classifier
   - by applying Naïve Bayes classification algorithm forthe header information.
   - Make a verification of the trained data using e-mail message content and update the trained classifier correspondingly.
   - Apply global markovian property on the designed classifier which results in hybrid classifier.
   - Using ant colony optimization technique store the designed hybrid classifier in ontology accordingly.

IV. RESULT AND DISCUSSIONS

The proposed hybrid spam filtration method is validated using machine learning algorithms, methodologies and preprocessing steps. Legitimate and spam e-mails were used for experimental setup. These emails were labeled as L and S in the designed classifier indicating Legitimate and Spam correspondingly.

Fig. 1 No of mails Vs Precision %

Classifiers including Naïve Bayes and Markov Random Field technique were executedusing training data. Using header information the training data were first tested. The designed classifier is stored according to ant colony optimization and is used.

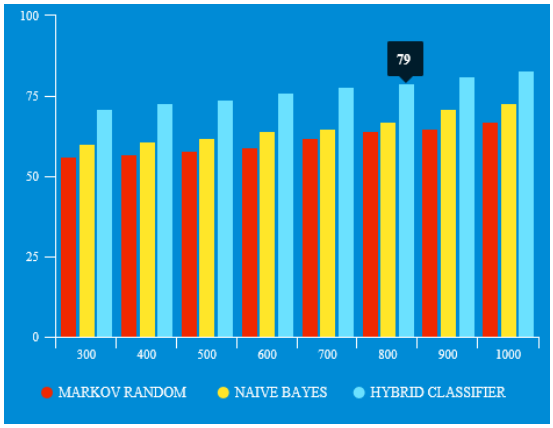

Fig. 2 No of mails Vs Accuracy %

For classification operation of new data entering into the system in future. The hybrid method is applied on the new data only when the existing classifier is not able to perform classification operation onthedata.Hencethe system is highlyefficient (90%) in terms of time consumption. Thus the proposed system is efficient in terms of accuracy, precision and time consumption.

V. CONCLUSION

Spam filtering is an important issue for secure communication using e-mail. Nowadays it is highly essential to provide an effective spam managing mechanism by considering the daily growth of spam and spammers. The proposed modeldemonstrated higher efficiency than existing classifiers like Naïve Bayes and Markov random field.Since users regularly suffer due to the problems such as reduced bandwidthand low efficiency the proposedhybrid method can be modeled to be implemented on a Mail Serverand Mail Client in order to eliminate these issues.

REFERENCES

1. Shalendra Chhabra etc. al., “Spam Filtering using a Markov Random Field Model with Variable Weighting Schemas”, Fourth IEEE International Conference on Data Mining (ICDM'04), 2004.
2. Nurul Fitriah Rusland etc. al., “Analysis of Naive Bayes Algorithm for Email Spam Filtering across Multiple Datasets”, International Research and Innovation Summit (IRIS2017) IOP Conf. Series: Materials Science and Engineering 226 (2017).
3. Ali Shafigh Aski etc. al., “Proposed efficient algorithm to filter spam using machine learning techniques”, Pacific Science Review A: Natural Science and Engineering 18 (2016) 145e149 Elsevier.
4. M. Tariq Banday etc. al., “Effectiveness and Limitations ofStatistical Spam Filters”, International Conference on “New Trends in Statistics and Optimization”.
5. Konstantin Tretyakov etc. al., “Machine Learning Techniques in Spam Filtering”, Data Mining Problem-oriented Seminar, 2004, MTAT.03.177, May 2004, pp. 60-79.
6. Neelam Choudhary etc. al., “Towards Filtering of SMS Spam Messages Using Machine Learning Based Technique”.

7. Le Zhang etc.a., "An Evaluation of Statistical Spam Filtering Techniques", Asian Transactions on Asian Language Information Processing, Vol.3, No.4, December 2004, pages:243-269.
8. Steve Webb etc. al., "An Experimental Evaluation of Spam Filter Performance and Robustness Against Attack".
9. Alexy Bhowmick etc. al., "Machine Learning for E-mail Spam Filtering: Review, Techniques and Trends", 2016.
10. Sarwat Nizamani etc. al., "Detection of fraudulent emails by employingadvanced feature abundance",2014. Egyptian Informatics Journal (2014) 15, 169–174.
11. Son Dinh etc al., "Spam campaign detection, analysis, and investigation", 2015 DFRWS 2015 Europe Digital Investigation 12 (2015) S12-S21.
12. Ali Shafigh Aski etc. al., "Proposed efficient algorithm to filter spam using machine learning Techniques", 2016 Pacific Science Review A: Natural Science and Engineering 18 (2016) 145-149.
13. Saadat Nazirova , "Survey on Spam Filtering Techniques" 2011 Communications and Network, 2011, 3, 153-160.
14. Amol G. Kakade etc. al., "Survey of Spam Filtering Techniquesand Tools, and MapReduce with SVM", 2013 Monthly Journal of Computer Science and Information Technology IJCSMC, Vol. 2, Issue. 11, November 2013, pg.91 – 98.
15. Diksha S. Jawale etc. al., "Hybrid spam detection using machine learning", 2018 International Journal of Advance Research, Ideas and Innovations in Technology (Volume 4, Issue 2).
16. Hedieh Sajedi etc. al., "SMS Spam Filtering Using Machine Learning Techniques:A Survey", 2016Machine Learning Research2016; 1(1): 1-14.
17. Subhana Khan etc. al., "A Hybrid E-Mail Spam Filtering Techniqueusing Data Mining Approach", 2016IJLTET Vol. 6 Issue 3 January 2016.