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

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

The tremendous development 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 emails into different classes for the detection of fraudulent emails. The categorization of emails such as legitimate and illegitimate can be done based on the intention of it.

Illegitimate Emails

Illegitimate emails may contain unnecessary informations, phishing emails [6], [7], [8], frightening messages, or plans for bad assassinations. 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 intention.
- Intended to get essential information from the receiver. It may cause damage to receiver’s computer.
- It may forward receiver to illegal website.

II. RELATED WORK

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

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

The tremendous development 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 emails into different classes for the detection of fraudulent emails. The categorization of emails such as legitimate and illegitimate can be done based on the intention of it.

Illegitimate Emails

Illegitimate emails may contain unnecessary informations, phishing emails [6], [7], [8], frightening messages, or plans for bad assassinations. 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 intention.
- Intended to get essential information from the receiver. It may cause damage to receiver’s computer.
- It may forward receiver to illegal website.

III. PROPOSED METHOD

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.
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 used in Markovian model. It is based on the theory of Markov chains defined by Andrey Markov. 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 \( X \) and \( Y \), the random variables \( X \) and \( Z \) are conditionally independent of \( X \) and \( Y \). Here, \( Z \) separates \( X \) and \( Y \) 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 probability these three Markov properties are equivalent.

II. RELATED WORK

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 Hafitidetected 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 Askii et. al. proposed an efficient algorithm [3] to filter spam with low error rates and high efficiency using a multilayer perception model. The proposed algorithm can be implemented on a Mail Server and Mail Client to eradicate problems, like reduction of bandwidth and low efficiency, from which users frequently suffer. M. Tariq Banday et. al. discussed the effectiveness and limitations [4] of statistical spam filters in which CBART and NB classifiers based filters showed better performance. Konstantin Tretyakov et. 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 classification algorithms. 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 filter-spam messages using hybrid method. The proposed hybrid spam filtration method uses Naive 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. Naive Bayes classifier
3. Markov Random Field technique

Preprocessing

Incomplete, aggregate, noisy, and missing values are handled in this preprocessing step. Conjunction words and articles which are not useful for classification of e-mail messages are removed from email body.
Naïve Bayes classifier

After preprocessing step, Naïve Bayes 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 for the 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.

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

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 on the data. Hence the system is highly efficient (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 model demonstrated higher efficiency than existing classifiers like Naïve Bayes and Markov random field. Since users regularly suffer due to the problems such as reduced bandwidth and low efficiency the proposed hybrid method can be modeled to be implemented on a Mail Server and Mail Client in order to eliminate these issues.

REFERENCES
Hybrid Spam Filtration Method using Machine Learning Techniques