A Neural Technique for Classification of Intercepted e-mail Communications with Multilayer Perceptron using BPA with LMS Learning

K. Soundararajan 1, U Eranna 2, Shilpa Mehta 3

1 Professor Dept. Of ECE, JNTUCEA, JNTU Anantapur
2 Principal, BITM Bellary, VTU Belgaum,
3 Asst. Prof, Dept. of Instrumentation Technology, PDIT Hospet, VTU Belgaum
Email- Shilpa Mehta shilpartha@yahoo.com

Abstract- Electronic mail (e-mail) is a very convenient ways for communicating over the Internet. With the rapid growth in Internet usage, the convenience of sending emails with zero cost and almost no delay and has led to increasing popularity of emails. Unfortunately, the facility is increasingly used for planning notorious activities by negative elements. Hence, the technical community is thinking and researching deeply into surveillance of communications. Many techniques have been developed for interception at the physical layer protocol level of the OSI model, but not much at the higher layers to classify the intercepted mails. But hardly any systems have yet been developed for proper and fast classification of intercepted mails into generalized categories. So, even if untoward mails are intercepted in time, the specific ones get overlooked, and no good can be done by just intercepting them. At most they can be used at prosecution stage, but not for the prevention of the planned notorious acts. Classification techniques based on keywords have been developed, but if the list of keywords is fixed, they are easy to overcome using substituted words etc. In this paper, we propose a multilayer feed-forward neural network based e-mail classification system that applies the BPA (Backward Propagation Algorithm) using LMS approach (Least Mean Squared Error) to classify intercepted messages. The inference engine has classified the mail database into various categories, attaining up to 95% accuracy. Here we use automatic real time generation of keywords on a continuous basis, thus adapting to the changing keywords used, even though they may be cleverly substituted. Further, as the whole technique uses a self learning approach, it does not depend on the language used.

Keywords: E-mail, Neural Networks, Classification, BPA, LMS, Surveillance, Language Independent.

I. INTRODUCTION

With the rapid development of the Internet, E-mail has become one of the most efficient and convenient ways to communicate. But, the increasing popularity and low cost of sending an e-mail make it very attractive to negative elements. They use it to plan incidents. Content based filters exist for categorizing spam, but we need effective and fast techniques for identifying suspicious messages among intercepted communication at a fast enough speed to avoid the occurrence of untoward incidents. It is very simple to physically intercept mails, but the sheer numbers of such intercepted communications make it almost impossible to check each one for contents in time to prevent incidents from happening.

Existing content-based filters are categorized as rule-based, key-word based and learning based. The keyword based filters [1] use a dictionary of common phrases, and search for a particular pattern in the messages. They perform well, but need to be maintained and tuned constantly, since the characteristics of the intended messages change over time. Rule-based filters [1, 2] use a wide range of tests to recognize features and assign a score to every mail. They are quite popular, but they too need periodic update and maintenance.

As emails are stored and forwarded at numerous nodes and gateways etc. from the source machine to the destination machine, the “capturing” or the interception of the mail is not a technical challenge. A copy can be made at any one of the numerous intermediate storage locations, by acquiring the relevant legal permissions. In this work we shall not talk about the physical interception of mails. Rather, the focus here is on the fast classification (of the already intercepted communications) into various categories of interest.
A Neural Technique for Classification of Intercepted e-mail Communications with Multilayer Perceptron using BPA with LMS Learning

A lot of work has been done to separate out messages into spam or legitimate based on certain criteria. But very little work has been done for categorization of messages into general categories (like - normal, planning, personal, official or other general categories of communication). The concentration in this work is on such classification towards making sense out of intercepted communications, without human interference. The privacy rights of users need not be encroached upon while simultaneously the rising concerns about social safety and security can be managed.

Fixed classification methods are unlikely to provide good performance for all purposes. Filters are still emerging that automatically learn how to identify messages, using previously received “bad” and legitimate messages. Sahami et. al. [6] proposed a machine learning algorithm to construct a filter, to learn automatically. A large number of machine learning techniques have been used for spam detection with promising results (Decision tree [7], support vector machine [8], k nearest neighbor [9], boosting [10]). Researchers increasingly use machine learning techniques to automatically build anti-spam filters.

We use a Neural Learning Based classifier to build an automatic general classifier in this work. The motivation of using neural approach for this purpose being that it can learn and adapt to the experience like a human can, and continuously improve itself.

Fuzzy classification assumes the presence of a shared area at the boundary, within which an object has partial membership in each class. The limitation of fuzzy logic is that, while it can reason with imprecise information, they can not acquire knowledge automatically. Learning techniques will be incorporated using neuron layers to define such boundaries. For each neuron, the synaptic weights vectors are computed.

II. Design And Implementation:

The flowchart for the overall algorithm is shown in Figure 1. The algorithm is as below:

a. At first the captured messages are scanned, and words are separated. A set of total words found in the whole epoch is created, and duplicates discarded. The resultant unique words are arranged in alphabetical order, not in the order of appearance. A MATLAB column vector is created for the total words present in all the messages.

b. Next, a second table is created indicating the presence or absence of specific words in specific messages. Now extra rows are appended to calculate the weight change and effective weight of each word after every recursion.

c. Columns are appended for total number of words per message, bias, target type of message as specified in the training set and resultant type of message calculated by the neural network.

d. Multilayer Feed-forward neural net is defined.

e. It is trained with BPA algorithm by minimizing mean squared error (MSE)

f. The model is validated and results simulated and compared to targets. The results are displayed

The following points need to be considered while studying this paper

• The scope of this paper is limited to e-mails with plain text only.
• The cases of purposeful text substitutions, image information etc, are not considered in this paper.
• No privacy rights have been invaded.
• Willing colleagues were requested to send mails on a server specifically set up for this work; the dataset constructed in this way contains about 300 messages.
• The dataset made in excel sheet using relevant apps contains the sender id, receiver id, the message itself, and the type of message.
• The messages selected for sending were basically from two categories, personal and official ones.
Figure 1. Flowchart for the Overall Algorithm.

We can select all the messages sent by that sender, by searching for the target “suspicious” sender mail Id’s in the relevant column. Similar approach with the receiver column can identify messages to a particular “suspect” receiver. Scanning the “body” of the message for keywords involves dealing with the “body” column of the selected message.

In this paper, the stress is on learning and weighing keywords themselves from the training dataset. The identification itself is a secondary task. The neuron can be embedded in a node to identify the new unclassified messages, after having learnt the initial keywords, and simultaneously keep adapting and strengthening its own keyword list.
III. Model Used:

Messages of various categories generally contain many distinctive features. One point to be considered is that the messages need not be always in English. Instead, they may be in any language written in the respective scripts with Unicode. Or else, words of other languages typed out using the English alphabet (called roman typing) may also be used. The initial training dataset consists of the messages along with the target type associated with the content (e.g. “happy birthday to you” is a personal message, while “the inspection is at 10 am on 22nd august” is an official one). Another thing to be remembered is that the language is not necessarily grammatically (and spelling wise) correct. The words ‘I am good, how are you all?’ could appear as ‘i m gud, hw r u al?’ Hence, making readymade and fixed keyword lists may prove completely useless.

Our detector recognizes words as ‘bags of alphabets’ and makes a database of unique combinations, appearing in the message set. It then continues, to identify (message by message) which ‘words’ appear more often in which ‘type’ of messages. Later on, the keyword list created here may be used by fuzzy Mamdani controllers also [4]. Here we shall use the multilayer feed forward neural model to adjust the synaptic weights of the respective input words. [5]

IV. DESIGN AND IMPLEMENTATION

A. Neural Network

The network receives the Boolean values (+ or – 1) as a N-element input vector, where each input indicates the presence (+1) or absence of a word in the message at hand. The message target is +1 for official and -1 for personal messages. It is then required to create and train the MLFF network, to identify the class by responding with a N-element output vector.

The network is a two-layer feed-forward net with 4 neurons in the hidden layer. The transfer function is ‘tansig’ for hidden layer and ‘purelin’ for output layer. The Back propagation network training function is ‘trainlm’. The function selected for Back Propagation weight/ bias learning function ‘learngdm’. Finally, the Performance function is ‘mse’. The network is trained to output a ‘+10’ for an official message and a ‘-7.951’ for a personal message.

The Network architecture and the training stage are shown below in Figure 2. We have chosen a two layer feed forward network with just four neurons in the hidden layer.

The Performance plot epoch wise is as shown in Figure. 3. It is seen that the Mean Squared Error (The Blue Curve) falls steadily from the first to the seventh epoch from 10^1 to 10^-30.
Figure 2. Neural Network Architecture and Training
A Neural Technique for Classification of Intercepted e-mail Communications with Multilayer Perceptron using BPA with LMS Learning

The training states in terms of gradient, Mu and validation checks for the last (seventh) epoch are shown in Figure 4.
The network is trained on 55 sets of messages. All training is done using back propagation with the function “trainlm”.

B. System Performance

The reliability of the neural network pattern recognition system is measured by testing the network with unseen input vectors. The number of messages and the erroneous classifications are displayed in the command window as shown below in Figure. 5

(a)
The regression plots at different stages (training, testing, validation and overall) are shown in Figure. 6.
From the Output screens shown in Figure 5 above we can see that the size of the test set being 55, the errors are reduced to just 3% of the total tested messages. This Performance has been achieved with a very tiny training set, and just a single layer of hidden neurons whose depth \( m_L \) is only 4 neurons. Much better accuracies (approaching 100%) are possible with more hidden layers and / or more number of neurons in the hidden layers, which can be done for practical applications.
A Neural Technique for Classification of Intercepted e-mail Communications with Multilayer Perceptron using BPA with LMS Learning

IV. Conclusions:

A classification technique using Multilayer Feed-forward neural network with Back Propagation Learning, for fast classification of intercepted communication into generalized categories has been proposed. The method was tested with a test mail set of about 200 emails, and another test set of 55 mails. The results were found to be 97% accurate. The classifier was found to give satisfactory results within the test set.

REFERENCES