Hidden Markov Models And Artificial Neural Networks For Spam Detection

[1] David Ndumiyana

Bindura University of Science Education

Faculty of Sciences

Department of Computer Science

P. Bag 2010

Bindura, Zimbabwe

[2] Lucia Sakala

Bindura University of Science Education

Faculty of Sciences

Department of Computer Science

P. Bag 2010

Bindura, Zimbabwe

Abstract

The stampede by both individuals and business organizations to the Internet, the acceptance of the Internet as a strategic tool for commerce, sharing of information through communication, electronic surveys and research, entertainment and the exponential growth of the World Wide Web have all combined to rekindle the threat from email spam. The Internet has virtually removed communication barriers between the corporate world and the rest of electronic world due to its ability to share and transmits information within the shortest time possible. Corporations today deploy spam filtering systems to guard the door from the outside world against invasion by unwanted email spam into their email inboxes thereby reducing the impact of junk emails.

This paper presents a spam filtering system using hidden markov models and artificial neural networks to filter out spam where word obfuscation on the keyword is conducted to evade detection. To detect spam with word obfuscation on the keywords, we experimented with the use of hidden Markov models (HMMs) to capture the statistical properties of spam variants belonging to the same class. The results showed that our model was able to detect over 90% of spam with a false positive rate of less than 13%. The use of artificial neural network

enhanced performance measurement of our filtering system especially on the ability of the system to learn more from any new spam messages that entered the system.

Keywords: Hidden Markov models, spam keyword obfuscation, spam filter, artificial neural networks.

INTRODUCTION

Spam is defined as the bulk-mailing, usually repeated several times, of a large chunk of unsolicited email messages [4], in most cases for commercial purposes to Internet users with whom the mailing organization has had no previous business partnership and whose email addresses the mailer collects from public spaces of the Internet such as newsgroups, mailing lists and websites among others [1, 4, 5]. Unwanted, unsolicited email is annoying to its recipients because it is undesired and poses a serious security threat to the network communication field. For instance, spam may contain a link that leads to a fraudulent website that intends capturing user's login information (phishing and identity theft) or provide a connection to a uniform resource locator (URL) which installs malicious software on the client computer [2] without permission. In most cases, installed malware can be used to collect user information, send junk email, host malicious software, host phish or launch denial of service attacks. Prevention of spam transmission is a very important thing to do, but detection allows Internet users and network service providers to address the challenge today [3] before it becomes a pandemic.Today people are sending and receiving email messages on a daily basis, communicating with friends, relatives and business partners because email continues to be cost effective means of exchanging files and confidential information. Email continues to be part of people's life everywhere in the world regardless of culture, creed or geographical locations just like i-pods and mobile phones, hence there need for spam filtering as unwanted hostile bulk email brings huge set of problems in terms of time spent and resources for dealing with those spam messages [2]. A critical analysis of previous published literature reveals an upward surge in the number of junk email messages coming to users email-boxes. For example, by the end of 2002, 40% of all email traffic consisted of unwanted email. The following year, 2003, the percentage of spam was approximated to about 50% of all emails. As if that was not enough, BBC News reported in 2006 that 96% of all emails received were spam [1].

Spam filtering technique is widely used to deal with a proliferation of junk emails sent to users email-boxes and can be defined as the automatic classification of messages into spam and legitimate mail. Spam filtering algorithms can be applied on different levels of email transmission at routers, at destination mail server or in the recipient's destination mailbox. There is an advantage if filtering is done at destination port as it prevents users from wasting their time on unwanted messages.

Generally, a spam filtering system is an application that uses the following function:

$$f(m, \Theta) = \begin{cases} a_{spam}, \text{ if the result is spam} \end{cases}$$

a_{ham}, otherwise (1)

where 'm' is a message or email to be classified, Θ is a vector of parameters and a_{spam} and a_{ham} are labels assigned to the email messages. The majority of spam filters are based on a machine learning classification method.When a learning based method is used the vector of parameters θ represents the result of training the classifier on a pre-collected dataset:

$$\Theta = \theta (M) \quad (2)$$

$$M = \{(m_1, y_1),\ldots\ldots\ldots(m_n, y_n)\},\ y_i \in \{ a_{spam}, a_{ham} \},$$

Where m_1, m_2 ,....., m_n are previously collected email messages, $y_1, y_2, \ldots\ldots\ldots\ldots, y_n$ are the corresponding labels and θ is the training function.

To classify new incoming email messages, a spam filtering system analyses them separately either by checking the presence of certain keywords or in groups. Moreover, learning-based filtering systems analyse a set of labelled training data, pre-collected email messages with reliable decision. A critical analysis of methods proposed by various researchers indicated that any email can be represented in terms of characteristics with discrete values based on some statistics of the presence or absence of certain keywords based on a vector space model [2].

RELATED WORK

The use of machine learning techniques by a variety of researchers demonstrated that a measurable level of success in filtering unwanted email messages had been achieved. We look at some of the potential and results achieved by some of frequently used techniques.

Using Hidden Markov Models

Hidden Markov models (HMMs) are very appropriate for use in statistical pattern analysis and from the time of initial application to speech recognition problems in the early 1970's [6], HMMs have had many application areas such as biological sequence analysis [7] and in short message service (SMS) spam detection [8] in mobile communication industry. A HMM is defined as a state machine where the transitions between the states are characterised by fixed probabilities and every state is associated with a single probability for observing a group of observation symbols. We can create and train a HMM to represent a collection of data which is usually in the form of observation sequences [6]. The states in the trained HMM are used to represent attributes of input data, while the transition and observation probabilities stand for statistical properties of these attributes. Now, for any given observation sequence in the machine, a match against a trained HMM is produced to determine the probability of observing such a sequence. The probability computed is high if there is similarity between the sequence and trained sequences.

For example in protein modelling, HMMs are used to model a given class of proteins [9] where the states correspond to the sequence of positions in space while the observations correspond to the probability distributions of the twenty amino acids that can occur in each

position. A model for a protein family assigns high probabilities to sequence belonging to that family. A trained HMM can therefore be used to distinguish family members from non-members.

Conducting word obfuscation on keywords generates classes of spam but this evasion technique does not completely remove similarities between variants in order to keep the purpose of spam in existence. Therefore the ability by a spam filter to detect these word obfuscations enhances the filter's capability to identify similarities among the variants. Hidden Markov models can provide a method to describe statistical variations of sequences in email spam messages whose keywords have been changed. In the same line of thinking, this paper proposed the use of HMMs similar to those used in protein sequence analysis to model classes of spam messages. In spam modelling the states correspond to the attributes of email message while the observations are keywords obfuscated. As a result, a trained model should therefore be able to assign high probabilities to and thereby identifying words belonging to the same class of spam.

Using Neural Networks

An Artificial neural network is a group of linked nodes or neurons of which the human brain is the first example that can be sighted because of its complexity. The fundamental purpose for neural network use is to extract linear combinations of inputs and generated characteristics from input and model the target as a nonlinear function of these attributes [16, 19]. Artificial neural network is a collection of algorithms that has capability of classification, regression and density approximation [10]. A neural network is composed of a sophisticated group of functions that have capability to disintegrate into smaller components (neurons, processing unit) which can be represented by means of a graph as a network.

Many researchers, for example James Clark [11] presented a paper on a neural network based system for automated email classification. The same author went on to present LINGER a neural network based system for automatic email classification problem. In his research, James Clarke [11] successfully showed that neural networks can be used for automatic email into mailboxes and spam mail filtering. Not to be outdone, Levent Ozgur [12], another researcher developed anti-spam filtering for agglutinative languages in general including for Turkish to be precise. He used dynamic methods which were based on artificial neural networks (ANNs) and Bayesian Network. Levent developed algorithms which are user-specific and they could adjust themselves with the features of incoming emails. According to results of his experiments, a 90% success rate was achieved. D. Punskis [13] also added global contribution to his research by applying neural network approach to the categorization of junk mails in which his technique employed features composed of descriptive characteristics of the most evasive patterns that spammers use rather than the context or frequency of words in the message. According to his findings, he concluded that ANN is good but should not be used alone as a spam filtering tool. The Back propagation algorithm is the most widely used ANN machine learning technique which Duhong Chen [14] applied in his research. A neural network with three layers was created with the following details: The first layer is the input layer with a node number which is equal to the number of frequent

words plus one. The second layer was the hidden layer with half number of the number of input node. And lastly, the third layer was the output layer with a single node. Their results showed a success rate of over 98%, which is a remarkable achievement. Ian Stuart [15] also used artificial neural network technique on a corpus of email messages collected from one user. The category set used to define spam messages was descriptive characteristics of words and messages similar to those a reader would use to detect spam. Ian found out that neural network required fewer attributes to achieve results similar to Naïve Bayesian approach.

Using Data Mining

Data mining is used to extract hidden knowledge that may not be explicitly stored in data structures but can be derived from real world relations and processes [16]. Classification is one of the supervised learning techniques used in the construction of models that describe different classes of data. A variety of examples are separated into their respective classes according to their distinctive features. The model constructed is used to predict the class of objects whose class label is unknown. In electronic communication world, each email contains many features such as keywords, sender details, or file attached that are used to predict the class of the message [17]. Classification involves two phases: The first phase is the construction of the model using a trained dataset in which every object of the class must be pre-classified so that its label is known. The second phase is the testing of the model by assigning class labels to data objects in a dataset [16, 17]. The test data is different from training data and every element of test data is also pre-classified. The accurate performance of the model is determined by comparing normal class labels in the testing set with those assigned by the model.

Schultz et al [18] experimented with a number of data mining techniques to identify new malicious binaries where they used three training algorithms to train a set of classifiers on some publicly available malicious and benign executables. The made comparisons of their algorithms to a traditional signature-based method and recorded a higher detection rate for each of their algorithms. The unfortunate thing is that their algorithms produced higher positive rates when compared to signature-based techniques.

Background Theory and Algorithms for Hidden Markov Models

A hidden Markov model is defined as a statistical model that describes a series of observations generated by stochastic process or Markov process [6]. A Markov process is a sequence of states, where the progression to the next state depends entirely on the present state but not on the past states. The Markov process in an HMM is hidden; what we can see is the sequence of observations associated with states. The fundamental objective in spam filtering is to make use of the observation information to gain insight into various aspects of the underlying Markov process [20]. Further details can be illustrated by means of an example taken from [20]. According to [20], he supposed that if one wants to find the average annual temperature of a particular location over a preceding period of several consecutive years and suppose that there is no recording of previous temperature of any form for that location, then, since there is no way we can know year-to-year temperature directly, we

search for clues to predict temperature indirectly. To simplify the facts, consider only two possible annual temperatures of either 'hot' (H) or 'cold' (C). Suppose we know the probability of a hot year followed by another hot year is given as 0.7 and that of a cold year followed by another cold year is 0.6, then this information can be illustrated on a matrix form:

	H	C
H	0.7	0.3
C	0.4	0.6

If it is assumed that research result reports that the treering sizes of a certain kind of tree, whether small (S), medium (M) or large (L), it related to the annual temperature recorded as:

	S	M	L
H	0.1	0.4	0.5
C	0.7	0.2	0.1

Meaning to say in a hot year, the probability of a tree having small, medium or tree ring is 0.1, 0.4 and 0.5 respectively. Observing the tree ring sizes for such a tree, we can use that information to deduce the possible annual temperatures over a period of time.

According to this example, the temperatures (H and C) are the states and the transition of temperature from one year to another precisely define Markov process. Tree ring sizes (S, M, L) are the observable outcomes and the probabilities of observing the different tree ring sizes at each temperature represent the probability distribution of the observable symbols at each state. The actual states are hidden since we cannot directly observe the temperatures recorded. We are able to see the observations (tree ring sizes) and these are associated with states in a statistical manner.

RESEARCH METHODOLOGY

A research methodology is defined as the architecture used to structure, plan and implement research process. The methodology is a model that describes a group of activities that leads to a successful investigation [21]. Researchers consider the ability to select an effective and ideal research methodology, a very important phase when doing a research because it guarantees a successful research development as well as minimizing the risk of failure.

We designed a HMM_ANN Spam detector using the algorithms developed by [27] for training, implementing and testing dataset obtained from [22].In figure 1(a) where regular Markov model is shown, x_i is the state directly visible to the user and the state transition probabilities $p(x_1|x_{i-1})$ are the only parameters.

Using the Markov property, the joint distribution for a sequence of n observations under this model is defined by:

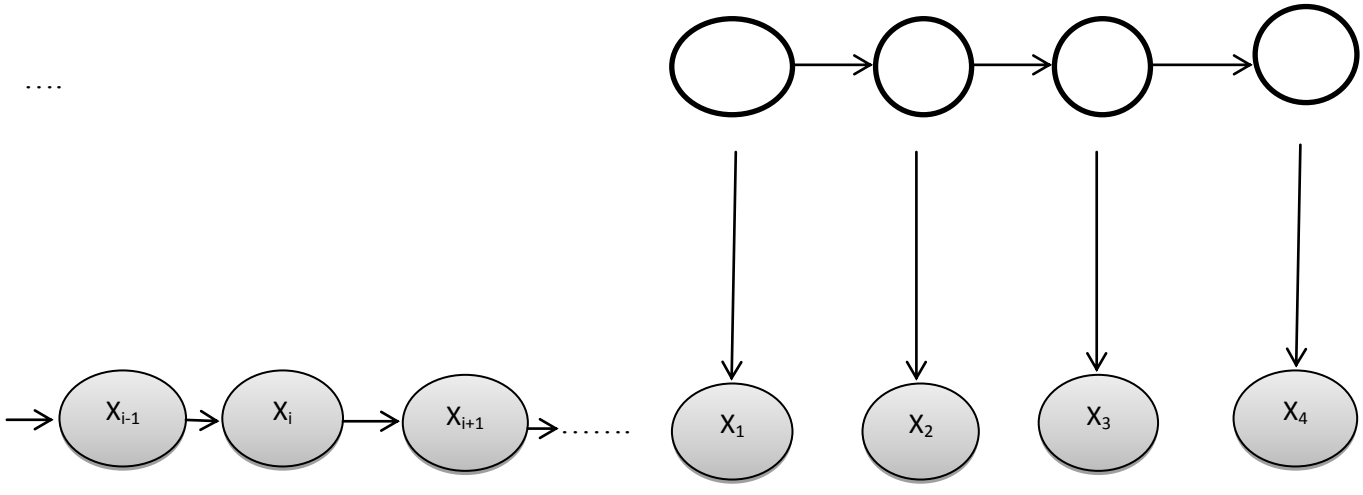
$$P(x_1 \ldots x_n) = p(x_1) \prod_{i=2}^{n} p(x_1|x_{i-1})$$

When the model is used to predict the next observation in a sequence, the distribution of predictions will depend on the value of the immediately preceding observation and shall not depend on previous observations. Figure 1 (b) shows hidden Markov models where the state y_i is not directly visible and only the output x_i is visible as it depends on the state. The hidden state space is discrete and the assumption is it consists of one of N possible values. The observations generated by Gaussian distribution are either discrete or continuous. If we suppose that the latent variables in figure 1(a) form first order Markov chain, then the random variable $y_t$is the hidden state at time t, and the random variable x_t is the observation at time t. The arrows in the figure indicate conditional dependencies. The diagram also illustrates that y_{t-1} and $y_{t+1}$are independent given y_t and hence $y_{t+1} \perp\!\!\!\perp y_{t-1}|y_t$ a key conditional independence property called Markov property. Without loss of generality the value of the observed variable x_t only depends on the value of hidden variable y_t whose joint distribution model is given by:

$$p(x_1,\ldots,x_n, y_1,\ldots,y_n) = p(y_1) \prod_{t=2}^{n} p(y_t \mid y_{t-1}) \prod_{t=1}^{n} p(x_t \mid y_t)$$

where $p(y_t \mid y_{t-1})$ is the state transition probability and $p(x_t \mid y_t)$ is the observation probability.

.



(a) Regular Markov model(RMM) (b) Hidden Markov model (HMM)

Figure 1: Showing RMM and HMM states transition Probabilities

Artificial Neural Networks

The fundamental goal of neural networks in the model was to accurately assist in identifying a classification of a given incoming mail messages even when facing a new class of input object the classifier may not have dealt with previously. Neural networks are included here because they are able to work with a huge amount of data thereby enhancing the effectiveness of our filter. This method is treated with a set of training data consisting of both spam and legitimate emails where phrases and obfuscated spam keywords are the input vectors. When the filter is eventually fed with test data the ANN finds a pattern and classifies as spam and non-spam.

The Learning Algorithms

The parameter learning task of HMM uses a set of possible states $Q_Y = \{q_1 \ldots q_N\}$and a set of possible observations $Q_X = \{o_1 \ldots o_M\}$to find the best set of state transition probabilities $A = \{a_{ij}\}$,

where $a_{ij} = p(y_{t+1} = q_j \mid y_t = q_i)$ and observation probabilities $B = \{b_i(k)\}$, where $b_i(k) = p(x_t = o_k | y_t = q_i)$ in addition to the initial state distribution $\prod = \{л_i\}$, where $л_i = p(y_0 = q_i)$ for a set of output sequences.

If $\Theta = \{A, B, \prod\}$ denote the parameters for a given HMM with static $Q_X$and Q_Y, then the basic task here is to generate the largest likelihood estimation of the parameters of HMM given the set of output sequences

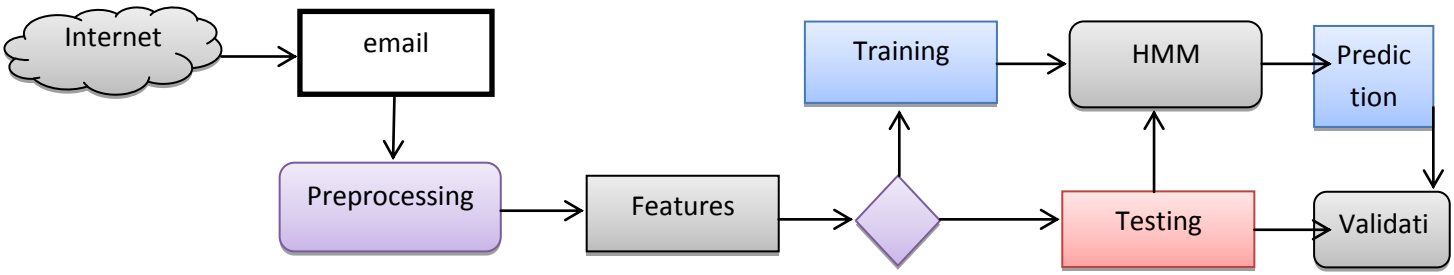


Figure 2: The Architecture of HMM

The incoming email is preprocessed and transformed into a set of keyword features. These features are put into two independent groups forming the training and testing set. The training set is used as a learning data for HMM model. The validation set is included to assess the quality performance of the model to ascertain its effectiveness when similar or new experiences are encountered.

Testing

Our research used dataset from Text Retrieval Conference 2007 (TREC 2007) Spam Track Public Corpus which had been chosen for our experiments [22]. The dataset consisted of 75 419 email messages with 25 220 email messages being legitimate while 50 199 of the messages were spam. TREC 2007 Spam Track Public Corpus is suitable for this research because of [23]: First, it is public available, making it possible for new and old researchers to verify the results or test against the same corpus. Secondly, Spam corpuses gathered from multiple email addresses provide better experimental results than when they are collected from a single address. Our spam filter classified email basing on spam keyword, hence a list of common spam keyword was gathered from this corpus for testing our dataset. This development therefore came up with the final dataset consisting of 172 mail keywords of which 46 of them were exact spam keywords, 75 of them were obfuscated spam keywords and the remaining 61 keywords were non-spam keywords.

Spam Filter Performance Measurement

Our spam filter was evaluated using a number of performance measurements, chief among them included spam recall (SR), spam precision (SP), false positive (FP), false negative (FN) and overall accuracy (A) [22]. Spam recall was used to measure the percentage of spam emails which were accurately categorized as spam while spam precision was measuring the proportion of email messages classified as spam that were indeed spam [23, 26].False positive refers to the number of legitimate messages which were inaccurately classified as spam while false negative defines the number of spam which were wrongly classified as legitimate messages. We listed the formulas of each measurement [23, 24] used to evaluate our spam filter:

$$SR = \frac{Nss}{Nss + Nsl}$$

$$SP = \frac{Nss}{Nss + Nls}$$

$$FP = \frac{Nls}{}$$

$$FN = \frac{Nsl}{Nls + Nll}$$

$$A = \frac{Nss + Nll}{Nall}$$

Nss: the numbers of spam correctly classified as spam by the filter.

Nll: the numbers of legitimate emails accurately classified as legitimate.

Nls: numbers of legitimate emails incorrectly classified as spam.

Nsl: the numbers of spam mistakenly classified as legitimate email.

N all: the total number of email messages in a dataset.

False positive generates much more severe consequences on the recipient than a false negative because information could be lost [25] when a legitimate message is prevented from reaching the user's mailbox and to make matters worse, the sending part may not be aware of this problem.

If all things go according to plan, a spam filter should have high spam precision, spam recall, accuracy with low false positive and false negative.

RESULTSAND DISCUSSION

Our experiment used a set of 172 email keywords from TREC 2007 Spam Corpus with five thresholds such as 60%, 65%, 70%, 75% and 80% meant to evaluate the best performance for our spam filter. Each group of email keyword was classified based on each of the five thresholds. The spam filter had to classify every incoming mail by calculating the mark for the email and compared the mark against the threshold. Any email was classified as spam if the mark for that particular email exceeded the chosen threshold; otherwise it was classified as a legitimate message. The number of spam keywords classified correctly and mistakenly classified was counted. In addition, the filter also counted the number of legitimate keywords which were accurately classified and those that the system wrongly categorized. The classification results are illustrated in the table 1.

Threshold	**60%**	**65%**	**70%**	**75%**	**80%**
Legitimate keywords classified accurately	32	37	44	47	49
Legitimate keywords classified inaccurately	19	14	7	4	2
Total Legitimate Keywords	51				
Spam keywords correctly classified	116	112	108	97	87
Spam keywords incorrectly classified	5	9	13	24	34
Total Spam Keywords	121				

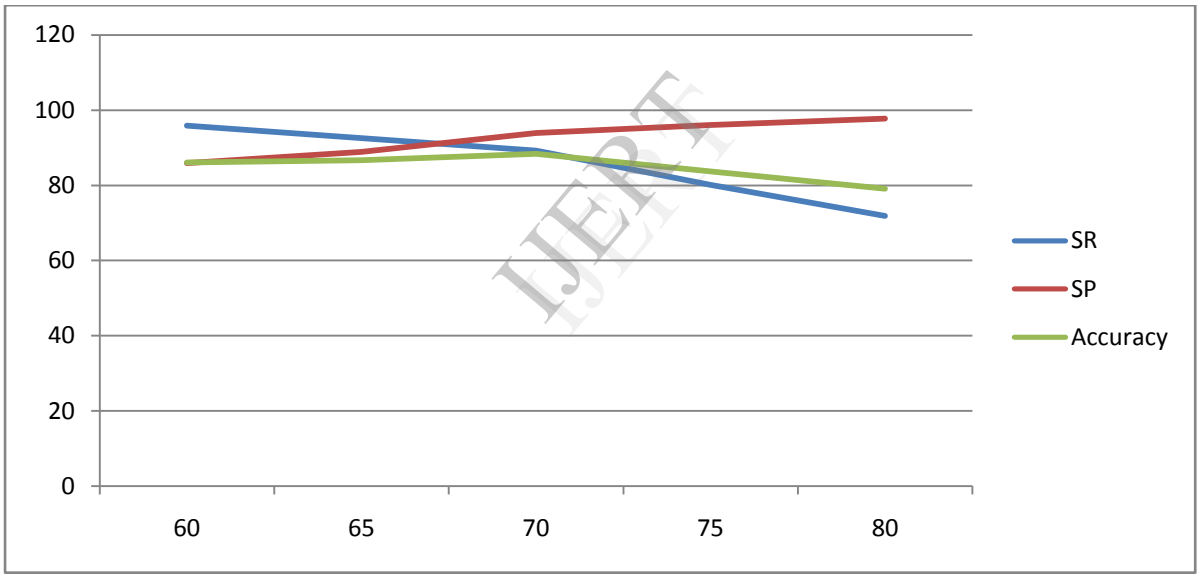
Table 1: Showing Classification Decision of five thresholds.

The classification decisions for performance measurements which are spam recall, spam precision, false positive, false negative and overall accuracy for each of the five thresholds had been calculated and determined. The results collected are displayed in table 2.

Threshold	**60%**	**65%**	**70%**	**75%**	**80%**

SR	95.87	92.56	89.26	80.17	71.90
SP	85.93	88.89	93.91	96.04	97.75
FP	37.25	27.45	13.73	7.84	3.92
FN	4.13	7.44	10.74	19.83	28.10
Accuracy	86.05	86.63	88.37	83.72	79.07

Table 3: Showing SR, SP, FP, FN and overall Accuracy for the 5 Thresholds

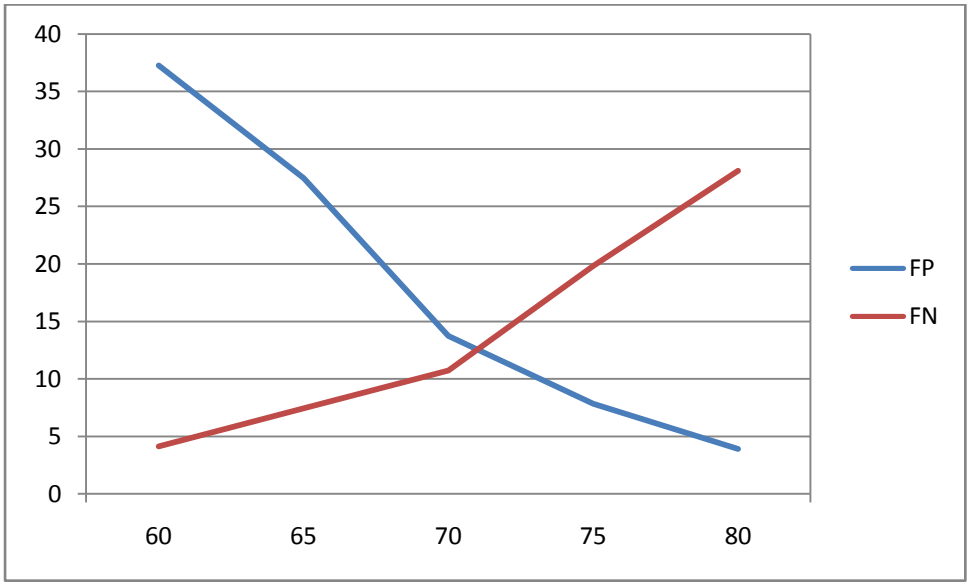


Threshold

Figure 4: Showing SR, SP and overall Accuracy Decision for the 5 threshold.

In most spam filtering system evaluation, performance evaluation using spam precision is usually preferred to spam recall and we also put performance measurement on spam precision for our filter. According to information illustrated by figure 1above, indicates that the higher the threshold the higher spam precision but the lower the spam recall becomes. Our spam filter recorded the highest spam recall of 95.87% at threshold 60% while the lowest spam recall of 71.90% at threshold 80% was recorded. The filter achieved the highest spam precision of 97.75% at threshold 80% with the lowest spam precision of 85.93% at threshold 60%. We can deduce that accuracy also increased from threshold 60% to 70% before

dropping from threshold 70% to 80%. The highest accuracy decision of 88.37% at threshold 70% was successfully achieved.



Threshold

Figure5: Showing FP andFN decisions for the 5 thresholds.

Our discussion on the results from the experiment carried out could not be complete if false positive and false negative are not discussed since they are important issues for every spam filtering system. According to information shown in figure 2, false positive rate was going down when the threshold became higher. It registered a minimum false positive of 3.92% at threshold 80%. Our observation of the same figure 2 indicates that contrary to false positive, false negative was increasing when the threshold became higherand in the process achieving the lowest false negative of 4.13% at threshold 60%. Our spam filtering system had the advantage that it could provide users with flexibility of controlling both false positive rates and false negative rates. For example, when users discovered that many of their legitimate emails were classified as junk mail, they could adjust threshold to higher score. On the other hand, they could adjust threshold to a lower target value if they discovered that they were receiving more spam than before.

From the results we can safely say our spam filter works perfectly with threshold 70% and provides a satisfactory result of 93.91% spam precision, 88.37% overall accuracy, 89.26% spam recall with false positives as low as 13.73%. We can safely say threshold 70% was defined as default threshold for this system while both thresholds 60% and 65% could not be defined as default thresholds because they produced high false positive rates. In a related decision making process, thresholds 75% and 80% were not chosen as default threshold owing to low accuracy and low spam recall even though they provided higher spam precision than threshold 70% according to results from the experiment.

CONCLUSION AND FUTURE WORK

This spam filter was able to detect all exact spam keywords accurately since results from the experiment proved that all 46 exact spam keywords in dataset had been accurately classified as spam in every threshold used. Therefore we conclude that the fundamental objective of this spam filter of detecting and preventing spam obfuscated from reaching recipient inbox was satisfactorily accomplished by achieving 93.91% spam precision and 88.37% overall accuracy. In addition, we recommend this spam filtering solution for use to enhance spam detection and classification so that concerns and risks brought by spam can be reduced.

In future research process, there is need to enhance the overall performance of the system by inserting as much spam keyword as possible in a pattern database. This suggestion is important because as the number of pattern in database increases, the chances of discovering exact spam keywords is also raised upwards.

References

[1] Gauthronet, S. & Drouard, E. (2001), "Unsolicited Commercial Communications and Data Protection",

Commission of the European Communities.

[2]Basavaraju, M. & Prabhakar, R. (2010), "A Novel Method of Spam Mail Detection using Text Based

Clustering Approach", International Journal of Computer Applications (0975-8887), 5 (4), pp. 15-25.

[3] Guha, S., Meyerson, A., Mishra, N., Motwani, R. & O'Callaghan, L. (2003), "Clustering Data Streams",

IEEE Trans.s on Knowledge and Data Engg.

[4] Mahmoud, T.M. & Mahfouz, A.M. (2012), "Sms Spam Filtering Technique Based on Artificial Immune

System", International Journal of Computer Science Issues, 9 (2), pp. 589 – 597.

[5] Goodman, J., Cormack, G.V. & Heckerman, D. (2007), "Spam and the ongoing battle for the inbox",

Communications of the ACM, 50 (2), pp. 25 – 33.

[6] Wong, W. (2006), "Analysis and Detection of Metamorphic Computer Viruses", Master Thesis, San Jose

State University.

[7] Krogh, A. (1998), "An introduction to hidden Markov models for biological sequences", Computational

Methods in Molecular Biology, pp. 45 – 63, Elsevier.

[8] Rafique, M.Z. & Farooq, M. (2010), "Sms spam detection by operating on byte-level distributions using

hidden Markov models",Virus Bulletin Conference.

[10] Tretyakov, K. (2004), "Machine Learning Techniques in Spam Filtering", Institute of Computer

Science, University of Tartu, Data Mining Problem-oriented Seminar, MTAT.03.177, pp. 60-79

[11] Clarke, J., Koprinska, I. & Poon, J. A Neural Network Based Approach to Automated Email

Classification.

[13] Puniskis, D., Laurutis, R. & Dirmeikis, R. (2006), "An Artificial Neural Nets for Spam email

Recognition", Electronics and Electrical Engineering ISSN, pp. 1215 - 1392

[14] Chen, D., Chen, T. & Ming, H. () Spam Email Filter using Naïve Bayesian, Decision Tree, Neural

Network, and AdaBoost.

[15] Stuart, I., Cha, S.H. & Tappert, C. (2004), "A Neural Network Classifier for Junk E-Mail", Proceedings

of Student/Faculty Research Day, CSIS, Pace University.

[16] Kovac, S. (2012), Master Thesis, "Suitability Analysis of Data Mining Tools and Methods", Faculty of

Informatics, Masaryk University.

[17] Appavu, S., Rajaram, R., Athiappan, G. & Muthupandian, M. (2007), "Data Mining for Suspicious

Email Detection: A Comparative Study", IADIS European Conference Data Mining.

[18] Schultz, M.G., Eskin, E., Zadok, E. & Stolfo, S.J. (2001), "Data Mining Methods for Detection of New

Malicious Executables", IEEE Symbosium on Security and Privacy, pp. 0038.

[19] Sabri, A.T., Mohammads, A.H., Al-Shargabi, B. & Hamdeh, M.A. (2010), "Developing New

Continuous Learning Approach for Spam Detection using Artificial Neural Networks", European

Journal of research, ISSN 1450-216X, 42(3), pp. 525-535.

[20] Stamp, M. (2012), "A Revealing Introduction to Hidden Markov Models", Department of Computer

Science, San Jose State University.

http://www.cs.sjsu.edu/faculty/stamp/RUA/HMM.pdf (visited October, 2012)

[21] Sommerville, I (2004), "Software Engineering 7th Edition, Addison Wesley,

[22] Text Retrieval Conference (TREC 2007). Spam Track Guidelines are available on:

http://plg.uwaterloo.ca/~gvcormack/spam/

[23] Oda, T. (2005). Master Thesis, "A Spam-Detecting Artificial Immune System", Carleton University

[24] Hunt, R. & Carpinter, J. (2006), "Current and New Development in Spam Filtering", Proceedings of

14th IEEE International Conference on Networks (ICON06), Singapore 23 – 29 April 2006,

(ISBN 0-7803-9746-0), vol. 2, pp. 1 – 6.

[25] Liu, C. & Stamm, S. (2007), "Fighting Unicode-obfuscated spam", In Proceedings of the Anti-

Phishing Working Groups 2nd Annual eCrime Researchers Summit, Pittsburgh, Pennsylvania, 4 - 5

October 2007,(ISBN 978 – 1- 59593 – 939 – 8), pp. 45 – 49.

[26] Duncan, C., Jacky, H., Kelvin, M., & Joel, S. (2006), Catching Spam Before It Arrives: Domain

Specific Blacklists Proceedings of the 2006 Australasian Workshops on Grid Computing and

E-research, Hobart, Tasmania, Australia, pp. 193 – 202.

[27] Sun, Y., Deng, H. & Han, J. (2010), "Mining Text Data: Probabilistic Models for Text Mining",

Department of Computer Science, University of Illinois at Urbana-Campaign.